Mean-field limits and beyond

Mean-field limits and beyond

Large deviations for singular interacting diffusions and variational convergence for population dynamics

Jasper Hoeksema

Printed by ADC Nederland

Cover: Design by Jeannette Koenis, with a background taken from www.freepik.com

A catalogue record is available from the Eindhoven University of Technology Library

ISBN: 978-90-386-5654-0

Mean-field limits and beyond

Large deviations for singular interacting diffusions and variational convergence for population dynamics

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op vrijdag 3 februari 2023 om 16:00 uur

door

Jasper Hoeksema

geboren te Groningen

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr. R.W. van der Hofstad
promotor:	prof.dr. M.A. Peletier
copromotor:	dr. O.T.C. Tse
leden:	dr. M. Maurelli (Università di Pisa)
	prof.dr. A. Mielke (Humboldt Universität zu Berlin)
	prof.dr. F.H.J. Redig (Technische Universiteit Delft)
	dr. R. Rossi (Università degli Studi di Brescia)
	prof.dr. A.P. Zwart

Het onderzoek dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

To Hans, my father

Abstract

This thesis consists of two parts, both focusing on mean-field limits for interacting particle systems, which describe the macroscopic behaviour of the system as the number of particles goes to infinity. In both cases, we will go further than merely investigating this limit and proving the convergence, by either describing the asymptotic probabilities of deviating from the limit, or the convergence of corresponding variational structures.

For Part I, we investigate large deviations for weakly interacting diffusions with singular drift, generalizing large deviations for singular functionals along the way. Via a uniqueness argument we then establish propagation of chaos, i.e. convergence to the McKean-Vlasov equation in a suitable way.

Next, in part II, we use a variational framework to study the Bolker–Pacala–Dieckmann–Law model used in ecology and population dynamics. By providing variational characterizations for jump processes and the expected mean-field limit, we establish convergence of these variational structures in the sense of generalized Energy Dissipation Principles. In particular, this will imply entropic propagation of chaos, describing an even stronger convergence of the particle system to the mean-field limit.

Finally, in Part III, we briefly summarize and review our results, and paint a way forward.

Contents

Chapter 1

Introduction

Throughout the years of my PhD, I fell in and out of love with a wide array of beautiful problems and worked on them with an even wider array of mathematical techniques. But for this thesis, I will restrict myself to the two subjects that I devoted most of my time to, and actually made any significant progress in, and the thread that connects them both: mean-field limits of interacting particle systems.

In essence, I will be painting a macroscopic picture that approximates the microscopic dynamics as the number of particles goes to infinity, for both a system of interacting diffusions and a particle system involving birth and death.

While I will discuss in detail the mathematical underpinnings and related literature of both subjects, let me briefly sketch these subjects and the tools I used to study them.

1.1 Weakly interacting systems and mean-field limits

In short, the dynamical systems that are considered in this thesis consist of interacting particles that are moving, wiggling, hopping, or even popping in and out of existence. For Part I, the guiding example is a set of N particles with positions $X_t^1, \ldots, X_t^N \in \mathbb{R}^d$ that, with some abuse of notation, satisfy for each particle

$$\dot{X}_t^i = \frac{1}{N} \sum_{j=1}^{N} \varphi(X_t^i - X_t^j) + noise, \tag{1.1}$$

where $\dot{X}_t^i$ is the velocity of particle i, φ is a function arising from the interaction between particles, and the noise terms are independent Gaussian variables.

Physically speaking, it can correspond to so-called overdamped Langevin dynamics, where φ is proportional to the force between each pair of particles and the noise is an approximation for the interaction between particles and some background medium, but similar equations can arise in various fields and applications.

Next, in Part II, I investigate a model stemming from ecology and population dynamics, where particles can create new particles that are subsequently spread around or can either die off due to competition for resources. In case the particles are living in $\mathbb{R}^d$ and A_x, A_y correspond to particles at positions $x, y \in \mathbb{R}^d$, the dynamics are characterized by a stochastic process such that on average

$$\begin{aligned} A_x &\to A_x + A_y, && m(x,y) \text{ times per second,} \\ A_x + A_y &\to A_y, && \tfrac{1}{n} c(x,y) \text{ times per second.} \end{aligned} \tag{1.2}$$

Here $m(x, y)$ is called, depending on the precise context, either a *dispersal* or *mutation* kernel, c is the *competition* kernel, and n is a parameter that will determine the average number of particles in the system. Because of the creation and annihilation of particles, the actual number of particles at time t, denoted as N_t, is itself a stochastic process.

Both cases can be cast as so-called *weakly* interacting particle systems, where for each particle i the interaction with other particles only depends on the empirical measure, a (possibly rescaled) discrete measure describing the positions of all particles. The interactions are scaled in such a way that as the number of particles goes to infinity, it is only their average, or *mean-field*, contributions that matter.

For example, in the case of the interacting diffusions of (1.1) and sufficiently regular φ, the empirical measure on path space,

$$z^N_{\boldsymbol{X}} := \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i},$$

with X^i now paths in $\mathbb{R}^d$, converges as $N \to \infty$ to the solution of the so-called McKean–Vlasov equation, which can be described as

$$\dot{X}_t = b(X_t, \mathrm{Law}(X_t)) + \mathit{noise}, \qquad b(x,\mu) := \int_{\mathbb{R}^d} \varphi(x-y)\, \mu(\mathrm{d}y), \tag{1.3}$$

where b is a measure-dependent *drift* vector. Moreover, for the model involving birth and death of (1.2) and for continuous and bounded m and c, the rescaled empirical measure at time t

$$\nu^n_t := \frac{1}{n} \sum_{i=1}^{N_t} \delta_{X^i_t},$$

converges as $n \to \infty$ to the mean-field limit $u_t(x)\mathrm{d}x$, where the function $u_t(x)$ satisfies the following nonlocal and nonlinear evolution equation

$$\partial_t u_t(x) = \left(\int_{\mathbb{R}^d} m(y,x)\, u_t(y) \mathrm{d}y \right) - \left(\int_{\mathbb{R}^d} c(x,y)\, u_t(y)\, \mathrm{d}y \right) u_t(x). \tag{1.4}$$

Moreover, φ will also be allowed to be *singular*, reflecting for example dynamics with very strong repulsive or attractive interactions, and the mutation/competition kernels m, c will be allowed to be non-continuous and defined over certain Polish spaces instead of $\mathbb{R}^d$.

In both settings, I will go beyond merely characterizing the limit itself. Namely, I will either describe the asymptotic probabilities of deviating from these limits, called *large deviations principles* (LDPs), or the corresponding asymptotic variational structures in terms of convergence of *Energy-Dissipation Principles* (EDPs). In either of these cases, the LDP and EDPs, together with the uniqueness of the limiting equation, imply convergence to the mean-field limit.

Although I am phrasing my work as two very different subjects, it should be noted that the variational structures for Part II.B are motivated by large deviation theory as well. In fact, we originally designed them as yet another tool to establish LDPs themselves. While that goal lies outside of the scope of this thesis, I will at the end give a sketch of how to close this loop.

1.2 Plan of the thesis

This thesis consists of two separate parts that can be read independently of each other: one leading up to large deviations for singularly interacting diffusions, Part I.B, and the other leading up to convergence of variational structures for population dynamics, Part II.B. To do so, a range of new techniques and structures will have to be introduced, which will be delegated to Parts I.A and II.A, and which I believe to be of independent interest. Finally, the results are reviewed in Part III.

Now, while Parts I.A, I.B, II.A and II.B all have their own introduction where we will describe and discuss their main results, let me briefly go over the content of each of them.

I.A Large deviations for singular functionals and Gibbs measures

This part is a continuation of the work of my Master's thesis [Hoe17] and is an adaptation of the joint work [HHMT20] together with Tom Holding, Mario Maurelli, and my supervisor and co-promotor Oliver Tse. Here we lift LDPs for non-interacting systems to interacting ones that are characterized by a singular energy functional, through an extension of Varadhan's Integral Lemma. In essence, this follows by approximation, i.e. we establish LDPs for singular functionals via a sequence of LDPs for regular functionals.

We apply this to Gibbs measures over Polish spaces, where the energy functional is the sum over pairwise interactions of the particles, and describe the various estimates and modifications necessary for Part I.B.

I.B Large deviations for singularly interacting diffusions

As the second component of the work [HHMT20], we return to our guiding example of (1.1), and investigate large deviation principles for the empirical measure $z^N_{\boldsymbol{X}}$ as the number of particles N goes to infinity. Via Girsanov's theorem, one can relate the interacting system to non-interacting ones via a change of measure. While strictly speaking this cannot be written in terms of Gibbs measures, the tools of Part I.A and approximation arguments do allow us to prove these LDPs for a large class of measure-dependent and singular drifts. In particular, after a uniqueness argument, we obtain convergence of $z^N_{\boldsymbol{X}}$ to the McKean-Vlasov equation (1.3).

II.A Variational and dissipation structures for jump processes

In order to take appropriate limits for measure-valued jump processes as will be done in Part II.B, we introduce a specific variational structure to the master equation for jump processes, the forward Kolmogorov equation, which describes the law of the process. This structure consists of a free energy functional, in our case given by the relative entropy with respect to some fixed measure π, and a dissipation potential relating the free energy functional and the evolution.

We generalize this to a scenario where π is simply a reference measure and not an invariant measure of the system. In particular, the energy is no longer dissipating along the evolution, but with some abuse of language, we refer to both cases as (generalized) Energy-Dissipation Principles. This allows us freedom in choosing the reference measure π, and to both treat the *reversible* and *irreversible* settings.

Moreover, in order to apply this framework to Part II.B we consider finite but possibly *unbounded* jump kernels.

II.B Variational convergence for population dynamics

In the final main subject of this thesis, we consider again the birth/death model described by (1.2), called the Bolker–Pacala–Dieckmann–Law (BPDL) model. It is part of ongoing work and a generalization of our work [HT22], which only considered the reversible setting.

For every value of the parameter n, we equip the forward Kolmogorov equation associated to the measure-valued process ν^n with a corresponding variational structure and show convergence of these structures as $n \to \infty$. This convergence is in the sense of EDPs, with the limiting structure corresponding to the Liouville equation, a transport equation associated with the mean-field limit (1.4). Moreover, from this convergence, we derive a variational formulation for the mean-field limit and establish convergence of the free energies, which implies an entropic propagation of chaos result for the particle system.

Part I.A

Large deviations for singular functionals and Gibbs measures

Chapter 2

Introduction

For this part, we consider large deviations for interacting particle systems that are determined by some singular energy functional. To do so we extend a classical tool of large deviation theory, called Varadhan's Integral Lemma, by approximating the singular energy functional by regular ones.

As the name might suggest, large deviations characterize the asymptotic probabilities of a system acting *far away* from its expected behaviour. With the precise definitions following in Section 3.1, let us say a family of variables $\{z^N\} \subset \mathcal{X}$ with laws P^N satisfy a large deviation principle (LDP) with rate function I and rate N if as N goes to infinity we have asymptotically

$$P^N(z^n \approx \mu) \asymp e^{-NI(\mu)}, \qquad \mu \in \mathcal{X}, \tag{2.1}$$

in some suitable way. It can be shown that if I has a unique minimizer μ^* the sequence z^N converges almost surely to μ^*, see Lemma 3.4, and hence in that case I determines the exponential rate at which the laws P^N *concentrate* around μ^*.

This is for example the case for the particle systems we investigate in Part I.B, where we show large deviations for a system of interacting diffusions and establish the uniqueness of the rate function, under suitable assumptions.

The fact that the particle systems we consider in this thesis satisfy a meaningful large deviation principle *at all* is not a trivial one, and has deep connections to statistical physics, as summarized in detail in the work [Tou09]. Although I will not go into the enormous amount of literature on large deviations and its historical roots, let me still highlight a few personal favourites: [dH08] as an excellent introduction; the unbelievably extensive [DZ10] for which it is said that entire articles have been written as answers to some of its exercises; the work [Pel14] which motivated me to study large deviations in the first place; and finally, the PhD thesis [Sch21] of my dear colleague, who spent considerable time and effort easing the reader into the world of large deviations, an approach that will not be taken in this thesis.

Now, let $\mathcal{X}$ be Polish, with $\mathcal{B}(\mathcal{X})$ its σ-algebra. The particle systems we investigate induce sequences of laws Q^N such that

$$\frac{dQ^N}{dP^N}(\mu) := \frac{1}{Z_N} e^{-N\mathcal{E}(\mu)} \qquad \text{for } P^N\text{-almost every } \mu \in \mathcal{X}, \tag{2.2}$$

where P^N are the laws for some *non-interacting* system, and $\mathcal{E}$ is an energy functional. In the cases we study a large deviation principle for P^N is already known and established, and therefore we would like to lift this to a large deviation principle for Q^N.

For regular enough $\mathcal{E}$, this is accomplished by Varadhan's Integral Lemma (or Varadhan's Lemma for short), which allows one to transfer an LDP through a change of measure and is in fact a natural extension of Laplace's method to infinite dimensional spaces. Namely, with again the precise definitions left to Chapter 3, the classical Varadhan's Lemma (cf. [Var84, DMZ03, DZ10]) reads as follows.

Proposition 2.1 (Varadhan's Integral Lemma)**.** *Suppose P^N satisfies an LDP with rate function $I : \mathcal{X} \to [0, \infty]$, and let $\mathcal{E} : \mathcal{X} \to \mathbb{R}$ be any continuous and bounded function. Then the family of measures $\{Q^N\}$ defined by*

$$\frac{dQ^N}{dP^N}(\mu) := \frac{1}{Z_N} e^{-N\mathcal{E}(\mu)}, \qquad \textit{for } P^N\textit{-almost every } \mu \in \mathcal{X}, \tag{2.3}$$

with normalizing constants Z_N, satisfies an LDP with rate function

$$F(\mu) := (I + \mathcal{E})(\mu) - \inf_{\nu \in \mathcal{X}} (I + \mathcal{E})(\nu). \tag{2.4}$$

Here the assumption of continuity of $\mathcal{E}$ is crucial, and for example cannot be replaced by arbitrary bounded measurable functions. Yet, various extensions have been developed to relax this assumption: for example to deal with singular functionals or contractions for Gibbs measures on $\mathbb{R}^d$ [BG99, CGZ14, DLR20, HLSS18], or on abstract spaces [Léo87, ES98, ES02, DMZ03, Ber18, GZ19, LW20]. Here we refer to the background paragraphs in Chapters 3 and 4 below for a more detailed discussion, and only highlight a few points here.

A common thread in some of the extensions above are various approximation arguments. For example, as outlined in [DZ10] on *exponential approximations*, the family $\{Q^N\}$ satisfies an LDP if there exists another family $\{Q^N_\lambda\}$ which satisfies an LDP for each $\lambda > 0$ and approximate Q^N in some exponentially good way (as $\lambda \to 0$). Liu and Wu [LW20] make use of techniques involving exponential approximations and prove LDPs for Gibbs measures with singular potential. However, in the setting of Part I.B, we cannot rely on their result since the associated $\mathcal{E}$ is not actually in the form of a Gibbs energy. Hence, we have developed the following extension (see Theorem 3.8 for the precise statement):

Theorem 2.2 (Extended Varadhan Integral Lemma)**.** *Let* $P^N = \mathrm{Law}(z^N)$ *be a family satisfying an LDP with rate function* I*, and* $\mathcal{E}, \mathcal{E}^N : \mathcal{X} \to [-\infty, \infty]$ *measurable functions. Moreover, let* $\mathcal{E}_\lambda$*,* $\mathcal{E}_\lambda^N : \mathcal{X} \to [-\infty, \infty]$*,* $\lambda > 0$ *be measurable functions such that, for each* $\lambda > 0$*, the family of measures* $\{Q_\lambda^N\}_N$ *defined by*

$$\frac{dQ_\lambda^N}{dP^N} := \frac{1}{Z_N^\lambda} e^{-N\mathcal{E}_\lambda^N(\mu)}, \qquad \text{for } P^N\text{-almost every } \mu \in \mathcal{X},$$

with normalizing constants Z_N^λ*, satisfy an LDP with rate function*

$$\mathcal{F}_\lambda(\mu) := (I + \mathcal{E}_\lambda)(\mu) - \inf_{\nu \in \mathcal{X}} (I + \mathcal{E}_\lambda)(\nu).$$

Suppose that for some $\gamma > 1$*,*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-\gamma N \mathcal{E}_\lambda^N(z^N)} \right] < +\infty,$$
$$\inf_{\mu \in D(I)} (I + \gamma \mathcal{E}_\lambda)(\mu) > -\infty,$$

for every $\lambda > 0$*, and that a constant* $K \in \mathbb{R}$ *exists, such that for every* $\beta \in \mathbb{R}$*,*

$$\limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N(\mathcal{E}^N - \mathcal{E}_\lambda^N)(z^N)} \right] \le K,$$
$$\limsup_{\lambda\to 0} \sup_{\mu \in D(I)} \big(\beta(\mathcal{E}_\lambda - \mathcal{E}) - I\big)(\mu) \le K.$$

Then the family $\{Q^N\}$ *defined by*

$$\frac{dQ^N}{dP^N} := \frac{1}{Z_N} e^{-N\mathcal{E}^N(\mu)}, \qquad \text{for } P^N\text{-almost every } \mu \in \mathcal{X},$$

satisfies an LDP with rate function

$$\mathcal{F}(\mu) := (I + \mathcal{E})(\mu) - \inf_{\nu \in \mathcal{X}} (I + \mathcal{E})(\nu).$$

We provide a self-contained proof that only relies on basic large deviation theory and elementary convexity estimates. The two main points of this extension are that we do not require $\mathcal{E}$ to be continuous and that we also do not require the approximating energies $\mathcal{E}_\lambda^N$ to be continuous, but merely to induce an LDP in the sense described above. The latter point is an essential tool in establishing large deviations for interacting diffusions, as will be discussed in Part I.B.

Outline In Chapter 3 we recall basic definitions and results in large deviation theory and provide an extension of Varadhan's Lemma in Theorem 3.8. Next, in Chapter 4, we use this to prove LDPs for empirical measures of mean-field Gibbs systems, where the log-densities $\mathcal{E}_V^N : \mathcal{P}(S) \to [-\infty, \infty]$ are parameterized by a family of Borel functions $V^N : S^k \to [-\infty, \infty]$, $k \in \mathbb{N}$ (see (4.1) for a precise definition). In Theorem 4.4, we provide sufficient conditions, in terms of suitable approximations of $\{V^N, V\}$, under which $\{\mathcal{E}_V^N, \mathcal{E}_V\}$ induces an LDP—this result is in the same spirit as [LW20]. We further prove a result (Theorem 5.1) that generalizes Theorem 4.4 to include 'Gibbs-like' measures—which are not of Gibbs form but such that the error $(\mathcal{E}_V^N - \mathcal{E}_{V_\lambda}^N)$ can be essentially bounded by Gibbs measures—that simplifies our task in proving LDPs for weakly interacting diffusions in Part I.B.

Chapter 3

An Extension of the Varadhan Integral Lemma

In this chapter, we extend the classical Varadhan's Integral Lemma (cf. [DZ10, Var84]), to allow for establishing LDPs via a change of measure with possibly discontinuous density.

3.1 Notations and preliminary results

We introduce some notation that we will use throughout Part I. The space $\mathcal{X}$ is a Polish space endowed with its Borel σ-algebra $\mathcal{B}(\mathcal{X})$. The symbol $\mathcal{P}(\mathcal{X})$ denotes the set of all probability measures on $\mathcal{X}$ and we use the letters P, Q, and P^N, Q^N, ... for probability measures on $\mathcal{X}$. With a little abuse of notation, we use μ both for a generic element of $\mathcal{X}$ and for the canonical random variable on $\mathcal{X}$ ($\mu(x) = x$ for all x in $\mathcal{X}$); this notation is unusual, but it will be convenient in the next sections, where $\mathcal{X}$ will be itself a space of probability measures. We often consider (without loss of generality) P^N as the law of an $\mathcal{X}$-valued random variable z^N, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ (independent of N); $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$.

We recall a definition of a large deviation principle (LDP).

Definition 3.1. A family of measures $\{Q^N\} \subset \mathcal{P}(\mathcal{X})$ satisfies an LDP with *rate function* $\mathcal{F} : \mathcal{X} \to [0, \infty]$ if (1) $\mathcal{F}$ is lower semi-continuous, if (2) for every Borel set A,

$$- \inf_{\mu \in A^o} \mathcal{F}(\mu) \leq \liminf_{N \to \infty} \frac{1}{N} \log Q^N(A) \leq \limsup_{N \to \infty} \frac{1}{N} \log Q^N(A) \leq - \inf_{\mu \in \bar{A}} \mathcal{F}(\mu),$$

and if (3) the family $\{Q^N\}$ is *exponentially tight*, i.e. there is a sequence of compact sets $K_M \subset \mathcal{X}$ such that

$$\limsup_{M\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log Q^N(\mathcal{X} \setminus K_M) = -\infty.$$

□

We denote the domain of $\mathcal{F}$ by $D(\mathcal{F}) := \{\mu \in \mathcal{X} \mid \mathcal{F}(\mu) < \infty\}$.

Remark 3.2. As shown in [DZ10, p. 8, 120], Definition 3.1 in Polish spaces is equivalent to stating that (2) holds with a *good rate function* $\mathcal{F}$, i.e. $\mathcal{F}$ having compact sublevel sets. □

Remark 3.3. Let $\{z^N\}$ be a family of $\mathcal{X}$-valued random variables such that $Q^N = \mathrm{Law}(z^N) \in \mathcal{P}(\mathcal{X})$ satisfies an LDP with rate function $\mathcal{F}$. If the minimizer $\mu^* \in \mathcal{X}$ of $\mathcal{F}$ is unique, then the LDP implies the convergence $Q^N \to \delta_{\mu^*}$ weakly. In fact, by a standard argument, we obtain a stronger result: almost sure convergence of the random variables z^N to μ^*, as stated below. For a proof, see for example [PS19, Theorem A.2].

Lemma 3.4. *Suppose Q^N satisfies an LDP with rate function $\mathcal{F}$, and that $\mathcal{F}$ has a unique minimizer μ^*. Then z^N converges $\mathbb{P}$-almost surely to μ^*.*

□

Now, let $P^N = \mathrm{Law}(z^N)$ be probability measures on $\mathcal{X}$ satisfying a large deviation principle with rate function $I : \mathcal{X} \to [0, \infty]$. We consider pairs $(\{\mathcal{E}^N\}, \mathcal{E})$ (denoted $(\mathcal{E}^N, \mathcal{E})$ for short) of a sequence of Borel functions $\mathcal{E}^N : \mathcal{X} \to \overline{\mathbb{R}}$ and a Borel function $\mathcal{E} : \mathcal{X} \to \overline{\mathbb{R}}$, and study whether an LDP may be established for the induced measures Q^N,

$$\frac{dQ^N}{dP^N}(\mu) := \frac{1}{Z_N} e^{-N\mathcal{E}^N(\mu)}, \qquad \text{for } P^N\text{-almost every } \mu \in \mathcal{X}, \tag{3.1}$$

where the normalization constants Z_N are assumed to be finite for all $N \in \mathbb{N}$.

Precisely, we define J and $\mathcal{F}$ as follows:

$$J(\mu) := \begin{cases} I(\mu) + \mathcal{E}(\mu), & \mu \in D(I), \\ +\infty, & \mu \notin D(I), \end{cases}$$

and, if $\inf_{\mu\in\mathcal{X}} J(\mu)$ is finite, we define

$$\mathcal{F}(\mu) := J(\mu) - \inf_{\mu\in\mathcal{X}} J(\mu).$$

Finally, note that by construction for any Borel set A

$$\inf_{\mu \in A \cap D(I)} (\mathcal{E} + I)(\mu) = \inf_{\mu \in A} J(\mu).$$

Then the property we investigate is given in the following definition.

Definition 3.5. We say that $(\mathcal{E}^N, \mathcal{E})$ induces an LDP if $\inf_{\mu \in \mathcal{X}} J(\mu)$ is finite, $\{Q^N\}$ (defined as in (3.1)) satisfies an LDP with rate function $\mathcal{F}$ and satisfies the so-called *Laplace principle*,

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)}\right] = -\inf_{\mu \in \mathcal{X}} J(\mu). \tag{3.2}$$

□

The following lemma provides a characterization in terms of an *unnormalized* LDP, in the sense that we leave out the normalization constants Z_N.

Lemma 3.6. *The following statements are equivalent:*

(i) The pair $(\mathcal{E}^N, \mathcal{E})$ induces an LDP (according to Definition 3.5);

(ii) $\inf_{\mu \in \mathcal{X}} J(\mu) \in \mathbb{R}$, J is lower semi-continuous, the family $\{Q^N\}$ is exponentially tight, and for every Borel set A,

$$\begin{aligned} -\inf_{\mu \in A^o} J(\mu) &\leq \liminf_{N \to \infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)} 1_A\right] \\ &\leq \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)} 1_A\right] \leq -\inf_{\mu \in \bar{A}} J(\mu). \end{aligned} \tag{3.3}$$

Proof. Suppose $(\mathcal{E}^N, \mathcal{E})$ induces an LDP. Exponential tightness of Q^N follows from the definition of an LDP, and since $\mathcal{F}$ is lower semi-continuous J is as well. By the Laplace principle (3.2),

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)}\right] = -\inf_{\mu \in \mathcal{X}} J(\mu),$$

where the right-hand side is assumed to be finite. For any $A \in \mathcal{B}(\mathcal{X})$, we have that

$$\begin{aligned} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)} 1_A\right] &= \frac{1}{N} \log \mathbb{E}\left[\frac{1}{Z^N} e^{-N\mathcal{E}^N(z^N)} 1_A\right] + \frac{1}{N} \log Z^N \\ &= \frac{1}{N} \log Q^N(A) + \frac{1}{N} \log Z^N. \end{aligned}$$

Therefore, by (3.2) and the LDP of Q^N, we then obtain

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)} 1_A \right] = \limsup_{N\to\infty} \frac{1}{N} \log Q^N(A) - \inf_{\mu\in\mathcal{X}} J(\mu)$$
$$\leq -\inf_{\mu\in\bar{A}} \mathcal{F}(\mu) - \inf_{\mu\in\mathcal{X}} J(\mu) = -\inf_{\mu\in\bar{A}} J(\mu).$$

The lower bound follows similarly, and hence (3.3) is satisfied.

Conversely, assume that J is lower semi-continuous, $\inf_{\mu\in\mathcal{X}} J(\mu)$ is finite and that (3.3) holds. Then $\mathcal{F}$ is lower semi-continuous as well, and by (3.3) applied to $A = \mathcal{X}$,

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)}\right] = -\inf_{\mu\in\mathcal{X}} J(\mu),$$

where the right-hand side is assumed to be finite. Now normalizing by Z_N and proceeding as above it follows that for any $A \in \mathcal{B}(\mathcal{X})$,

$$-\inf_{\mu\in A^o} \mathcal{F}(\mu) \leq \liminf_{N\to\infty} \frac{1}{N} \log Q^N(A) \leq \limsup_{N\to\infty} \frac{1}{N} \log Q^N(A) \leq -\inf_{\mu\in\bar{A}} \mathcal{F}(\mu).$$

Along with the exponential tightness of Q^N, this implies that $(\mathcal{E}^N, \mathcal{E})$ induces an LDP. □

Remark 3.7. Notice that the classical Varadhan's Integral Lemma is recovered when $\mathcal{E}$ is continuous and bounded, and $\mathcal{E}^N = \mathcal{E}$ for all $N \in \mathbb{N}$. □

3.2 Main result

Now we present the main result of this chapter, where we consider a family of pairs $\{(\mathcal{E}_\lambda^N, \mathcal{E}_\lambda)\}_\lambda$ approximating the pair $(\mathcal{E}^N, \mathcal{E})$ as $\lambda \to 0$. Recall that $(\mathcal{E}^N, \mathcal{E})$ is shorthand notation for $(\{\mathcal{E}^N\}_N, \mathcal{E})$.

Theorem 3.8 (Extended Varadhan Integral Lemma)**.** *Let $P^N = \mathrm{Law}(z^N)$ be a family satisfying an LDP with rate function I. Moreover, let $(\mathcal{E}_\lambda^N, \mathcal{E}_\lambda)$ be pairs inducing an LDP for all $\lambda > 0$. Suppose that the pair $(\mathcal{E}^N, \mathcal{E})$ is such that for some $\gamma > 1$,*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-\gamma N\mathcal{E}_\lambda^N(z^N)} \right] < +\infty, \tag{3.4a}$$

$$\inf_{\mu\in D(I)} (I + \gamma\mathcal{E}_\lambda)(\mu) > -\infty, \tag{3.4b}$$

for every $\lambda > 0$, and that a constant $K \in \mathbb{R}$ exists, such that for every $\beta \in \mathbb{R}$,

$$\limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N(\mathcal{E}^N - \mathcal{E}_\lambda^N)(z^N)} \right] \leq K, \tag{3.5a}$$

$$\limsup_{\lambda\to 0} \sup_{\mu\in D(I)} \big(\beta(\mathcal{E}_\lambda - \mathcal{E}) - I\big)(\mu) \leq K. \tag{3.5b}$$

Then the family $\{Q^N\}$ *defined by* (3.1) *satisfies an LDP with rate function* $\mathcal{F}$. *In particular, the pair* $(\mathcal{E}^N, \mathcal{E})$ *also induces an LDP.*

Remark 3.9. We give some comments on the assumptions:

(i) It was shown in [Hoe17] that under (3.4b), condition (3.5b) is equivalent to the uniform convergence of $\mathcal{E}_\lambda$ to $\mathcal{E}$ on the sublevel sets $\{\mu \mid I(\mu) \leq M\}$ of I for any $M \in \mathbb{R}$, and that (3.5a) implies:

$$\text{For any } \delta > 0: \quad \lim_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log P^N\left(|\mathcal{E} - \mathcal{E}_\lambda|(z^N) > \delta \right) = -\infty. \tag{3.6}$$

(ii) Condition (3.5b) could appear redundant for those who are familiar with LDPs: if $(\mathcal{E}^N - \mathcal{E}^N_\lambda, \mathcal{E} - \mathcal{E}_\lambda)$ is *a priori* known to induce an LDP, (3.5b) and (3.5a) are equivalent. But in the proof we need to approximate $\mathcal{E}^N$ and $\mathcal{E}$ separately, which requires us to have both conditions. Nevertheless, we do expect from this reasoning that bounds for (3.5a) are also bounds for (3.5b). We will see that this is indeed the case for the interacting particle systems in Chapter 4.

(iii) Note that condition (3.5a) implies (cf. Lemma A.4)

$$\lim_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\beta\mathcal{E}^N_\lambda(z^N)}\, 1_A\right] = \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\beta\mathcal{E}^N(z^N)}\, 1_A\right],$$

and

$$\lim_{\lambda\to 0} \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\beta\mathcal{E}^N_\lambda(z^N)}\, 1_A\right] = \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\beta\mathcal{E}^N(z^N)}\, 1_A\right].$$

for any Borel set $A \in \mathcal{B}(\mathcal{X})$ and $\beta \in \mathbb{R}$, which is considerably stronger than simply inequalities for open and closed sets. An open question is whether this is stronger than, equivalent to, or weaker than (or none of the former) the statement that the induced measures Q^N_λ *exponentially approximates* Q^N as $\lambda \to 0$ (cf. [DZ10, p. 130]).

(iv) An alternative approach to proving the type of results in Theorem 3.8 could be to get an LDP for z^N in a larger space with a stronger topology, where $\mathcal{E}^N$ is a continuous function, and then apply the classical Varadhan Lemma; the LDP in the stronger topology could be obtained by exponential approximation, as in [DZ10]. This strategy is used, for example in [ES02, LW20], in the context of certain singular Gibbs measures.

□

We now briefly explain the strategy to prove Theorem 3.8. For a Borel set $A \in \mathcal{B}(\mathcal{X})$, we define the following functionals on the space of Borel functions $\mathcal{E}$ on $\mathcal{X}$,

$$\begin{cases} \phi_A(\mathcal{E}) := -\inf_{\mu \in A \cap D(I)} (I + \mathcal{E})(\mu), \\ \phi_A^N(\mathcal{E}) := \dfrac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}(z^N)} 1_A \right], \qquad n \in \mathbb{N}. \end{cases} \tag{3.7}$$

We will show in Lemma 3.10 that ϕ_A^N and ϕ_A are convex and bounded from above (in A) by $\phi_{\mathcal{X}}^N$ and $\phi_{\mathcal{X}}$ respectively. Moreover, the fact that $(\mathcal{E}^N, \mathcal{E})$ induces an LDP can be read as a set of variational inequalities for ϕ_A^N and ϕ_A for each $A \in \mathcal{B}(\mathcal{X})$, i.e.

$$\phi_{A^o}(\mathcal{E}) \leq \liminf_{N\to\infty} \phi_A^N\left(\mathcal{E}^N\right) \leq \limsup_{N\to\infty} \phi_A^N\left(\mathcal{E}^N\right) \leq \phi_{\bar{A}}(\mathcal{E}).$$

Finally, the convergence of (3.5) over $\mathcal{X}$ will be seen to imply corresponding statements over every set $A \in \mathcal{B}(\mathcal{X})$, which implies bounds on $\phi_A^N(\mathcal{E}^N - \mathcal{E}_\lambda^N)$ and $\phi_A(\mathcal{E} - \mathcal{E}_\lambda)$. Hence the extended Varadhan Integral Lemma is morally equivalent to a type of stability of variational inequalities for convex functionals, for which we can use Theorem A.1 in Appendix A.

Here are the convexity properties and bounds for ϕ_A and ϕ_A^N:

Lemma 3.10. *For any Borel set $A \in \mathcal{B}(\mathcal{X})$ and any $N \in \mathbb{N}$, the functionals ϕ_A and ϕ_A^N defined in* (3.7) *are convex and bounded from above by $\phi_{\mathcal{X}}$ and $\phi_{\mathcal{X}}^N$ respectively (that is, $\phi_A \leq \phi_{\mathcal{X}}$ and $\phi_A^N \leq \phi_{\mathcal{X}}^N$ for every $A \in \mathcal{B}(\mathcal{X})$ and any $N \in \mathbb{N}$).*

Proof. For any $N \in \mathbb{N}$ and any $\alpha \in (0, 1)$, for any Borel functions $\mathcal{E}_1, \mathcal{E}_2$, it holds by Hölder's inequality (with exponents $1/\alpha$ and $1/(1-\alpha)$)

$$\begin{aligned} \log \mathbb{E}\left[e^{-N(\alpha\mathcal{E}_1+(1-\alpha)\mathcal{E}_2)(z^N)}\, 1_A\right] &= \log \mathbb{E}\left[e^{-\alpha N\mathcal{E}_1(z^N)}\, 1_A\, e^{-(1-\alpha)N\mathcal{E}_2(z^N)}\, 1_A\right] \\ &\leq \alpha \log \mathbb{E}\left[e^{-N\mathcal{E}_1(z^N)}\, 1_A\right] + (1-\alpha) \log \mathbb{E}\left[e^{-N\mathcal{E}_2(z^N)}\, 1_A\right]. \end{aligned} \tag{3.8}$$

Convexity of ϕ_A^N follows by dividing (3.8) by N. The bound $\phi_A^N \leq \phi_{\mathcal{X}}^N$ follows from the positivity of the exponential and the monotonicity of the logarithm. Finally, for any $\alpha \in (0, 1)$,

$$\begin{aligned} \inf_{\mu \in A\cap D(I)} I + (\alpha\mathcal{E}_1 + (1-\alpha)\mathcal{E}_2) &= \inf_{\mu \in A\cap D(I)} [\alpha(I + \mathcal{E}_1) + (1-\alpha)(I + \mathcal{E}_2)] \\ &\geq \alpha \inf_{\mu \in A\cap D(I)} (I + \mathcal{E}_1) + (1-\alpha) \inf_{\mu \in A\cap D(I)} (I + \mathcal{E}_2), \end{aligned}$$

which gives convexity of ϕ_A. The bound $\phi_A \leq \phi_{\mathcal{X}}$ is easily verified. □

We can now prove Theorem 3.8.

Proof of Theorem 3.8. Recall, for any Borel set $A \in \mathcal{B}(\mathcal{X})$ and Borel function $\mathcal{G}$, we have by definition

$$\phi_{A^o}(\mathcal{G}) = -\inf_{\mu \in A^o \cap D(I)} (I + \mathcal{G})(\mu), \qquad \phi_{\bar{A}}(\mathcal{G}) = -\inf_{\mu \in \bar{A} \cap D(I)} (I + \mathcal{G})(\mu),$$

and

$$\phi_A^N(\mathcal{G}) = \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{G}(z^N)} 1_A\right].$$

Lemma 3.10 gives that ϕ_{A^o}, $\phi_{\bar{A}}$ and ϕ_A^N are convex for every $N \in \mathbb{N}$. Moreover, by the bounds $\phi_A^N \leq \phi_{\mathcal{X}}^N$ and $\phi_{\bar{A}} \leq \phi_{\mathcal{X}}$, assumptions (3.4) imply, for some $\gamma > 1$ (independent of λ),

$$\left.\begin{aligned} \limsup_{N\to\infty} \phi_A^N(\gamma \mathcal{E}_\lambda^N) < +\infty, \\ \phi_{\bar{A}}(\gamma \mathcal{E}_\lambda) < +\infty, \end{aligned}\right\} \quad \text{for every } \lambda > 0,$$

while assumptions (3.5) imply

$$\left.\begin{aligned} \limsup_{\lambda\to 0} \limsup_{N\to\infty} \phi_A^N(\beta(\mathcal{E}^N - \mathcal{E}_\lambda^N)) \leq K, \\ \limsup_{\lambda\to 0} \phi_{\bar{A}}(\beta(\mathcal{E} - \mathcal{E}_\lambda)) \leq K, \end{aligned}\right\} \quad \text{for every } \beta \in \mathbb{R}.$$

By Lemma 3.6, the assumption that $(\mathcal{E}_\lambda^N, \mathcal{E}_\lambda)$ induces an LDP is characterized by

$$\phi_{A^o}(\mathcal{E}_\lambda) \leq \liminf_{N\to\infty} \phi_A^N\left(\mathcal{E}_\lambda^N\right) \leq \limsup_{N\to\infty} \phi_A^N\left(\mathcal{E}_\lambda^N\right) \leq \phi_{\bar{A}}(\mathcal{E}_\lambda). \tag{3.9}$$

We are now in a position to apply Theorem A.1, which implies that (3.9) also holds for $(\mathcal{E}^N, \mathcal{E})$ (cf. (A.5)), i.e.

$$\phi_{A^o}(\mathcal{E}) \leq \liminf_{N\to\infty} \phi_A^N\left(\mathcal{E}^N\right) \leq \limsup_{N\to\infty} \phi_A^N\left(\mathcal{E}^N\right) \leq \phi_{\bar{A}}(\mathcal{E}).$$

Moreover, for every $\gamma' \in (0, \gamma)$, both $\limsup_{N\to\infty} \phi_A^N(\gamma' \mathcal{E}^N) < +\infty$ and $\phi_{\bar{A}}(\gamma' \mathcal{E}) < +\infty$ (cf. (A.4a) and (A.4b)). In particular, for $A = \mathcal{X}$, we have that $-\phi_{\mathcal{X}}(\mathcal{E})$ is finite, and

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-\gamma N \mathcal{E}^N(z^N)}\right] < \infty. \tag{3.10}$$

By Lemma 3.6, we can conclude that $(\mathcal{E}^N, \mathcal{E})$ induces an LDP provided we show lower semi-continuity of J and exponential tightness of Q^N.

First, from (3.5b) we can conclude that for every $K' > K$ and $\beta \geq 0$ there exists a large enough $\lambda^*(K', \beta)$ such that

$$|\mathcal{E}_\lambda - \mathcal{E}|(\mu) \leq \frac{K' + I(\mu)}{\beta}, \quad \forall \mu \in \mathcal{X}, \forall \lambda \geq \lambda^*(K', \beta).$$

In particular, we derive that $\mathcal{E}_\lambda$ converges pointwise to $\mathcal{E}$ on $D(I)$ (in fact, the convergence is uniform on sublevel sets of I). Next, note that a Borel function J is lower semi-continuous if and only if for every $\mu \in \mathcal{X}$

$$\liminf_{\epsilon \to 0} \inf_{\nu \in B_\epsilon(\mu)} J(\nu) \geq J(\mu),$$

which in the case of

$$J_{\mathcal{G}}(\mu) := \begin{cases} I(\mu) + \mathcal{G}(\mu), & \mu \in D(I), \\ +\infty, & \mu \notin D(I), \end{cases}$$

for a Borel function $\mathcal{G}$ can be rewritten as

$$\limsup_{\epsilon \to 0} \phi_{B_\epsilon(\mu)}(\mathcal{G}) \leq -J_{\mathcal{G}}(\mu). \tag{3.11}$$

To show this for $\mathcal{G} = \mathcal{E}$, fix μ, and note that by convexity for any $\alpha \in [0, 1)$, λ, ϵ,

$$\begin{aligned} \phi_{B_\epsilon(\mu)}(\alpha\mathcal{E}) &\leq \alpha\phi_{B_\epsilon(\mu)}(\mathcal{E}_\lambda) + (1-\alpha)\phi_{B_\epsilon(\mu)}\left(\alpha(1-\alpha)^{-1}(\mathcal{E} - \mathcal{E}_\lambda)\right) \\ &\leq \alpha\phi_{B_\epsilon(\mu)}(\mathcal{E}_\lambda) + (1-\alpha)\phi_{\mathcal{X}}\left(\alpha(1-\alpha)^{-1}(\mathcal{E} - \mathcal{E}_\lambda)\right). \end{aligned}$$

Since the left-hand side is independent of λ, taking subsequently limits in ϵ and λ and using the lower semi-continuity of $I + \mathcal{E}_\lambda$ we derive

$$\begin{aligned} \limsup_{\epsilon \to 0} \phi_{B_\epsilon(\mu)}(\alpha\mathcal{E}) \leq \alpha \liminf_{\lambda \to 0} \limsup_{\epsilon \to 0} \phi_{B_\epsilon(\mu)}(\mathcal{E}_\lambda) \\ + (1-\alpha) \limsup_{\lambda \to 0} \phi_{\mathcal{X}}\left(\alpha(1-\alpha)^{-1}(\mathcal{E} - \mathcal{E}_\lambda)\right) \\ \leq -\alpha \limsup_{\lambda \to 0} J_{\mathcal{E}_\lambda}(\mu) + (1-\alpha)K. \end{aligned}$$

Since $g(\alpha) := \limsup_{\epsilon \to 0} \phi_{B_\epsilon(\mu)}(\alpha\mathcal{E})$ is convex and bounded from above around $\alpha = 1$ we conclude by Lemma A.3 after letting $\alpha \to 1$,

$$\limsup_{\epsilon \to 0} \phi_{B_\epsilon(\mu)}(\mathcal{E}) \leq -\limsup_{\lambda \to 0} J_{\mathcal{E}_\lambda}(\mu).$$

Now, to establish (3.11) for $\mathcal{G} = \mathcal{E}$, note that either $I(\mu) = +\infty$, in which case both sides of (3.11) are equal to $-\infty$, or we have $\mu \in D(I)$ and thus we employ the pointwise convergence of $\mathcal{E}_\lambda$ to $\mathcal{E}$.

Finally, to prove exponential tightness, fix an arbitrary $M \geq 1$ and let K_M be a compact set in $\mathcal{X}$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log P^N(\mathcal{X} \setminus K_M) < -M.$$

By Hölder inequality we derive, for every α in $(0, 1)$,

$$\frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)} 1_{\mathcal{X}\setminus K_M}\right] \leq \frac{\alpha}{N} \log \mathbb{E}\left[e^{\alpha^{-1}N\mathcal{E}^N(z^N)}\right] + \frac{1-\alpha}{N} \log P^N\big(\mathcal{X} \setminus K_M\big).$$

After taking the limit supremum in N, the first term on the right-hand side is independent of M and finite by (3.10), provided $\alpha^{-1} < \gamma$. Therefore,

$$\limsup_{M\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-N\mathcal{E}^N(z^N)} 1_{\mathcal{X}\setminus K_M}\right] = -\infty,$$

which proves the exponential tightness of Q^N and concludes the proof. □

Background We are only aware of one paper extending Varadhan's lemma for discontinuous log-densities in a general framework, namely [DMZ03], which however uses different assumptions. Instead, extensions of Varadhan's lemma for particular contexts have been proven, often involving conditions like (3.6). For example, [ES98] proves an extended contraction principle which is closely related to exponential approximations (cf. [DZ10]); a localized version of (3.6) is used in [BG99] for their concept of quasi-continuity to prove LDPs for vortex systems; the papers [ES02, LW20] use an alternative strategy to get an extension of the classical Varadhan Lemma in the context of singular Gibbs measures, see Remark 3.9(iv).

Chapter 4

Gibbs measures

4.1 Notations and preliminary results

In this chapter, we consider the setting for Gibbs measures over weakly interacting particle systems.

Let S be a Polish space, endowed with its Borel σ-algebra $\mathcal{B}(S)$. Let $\mu_0 \in \mathcal{P}(S)$ be a given reference measure. For simplicity, we will suppose in the following that μ_0 is *non-atomic*, i.e. it has no atoms (cf. Remark 4.1). We are given i.i.d. random variables ω_i, $i \in \mathbb{N}$, defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in S and common law μ_0; we can think of ω_i as non-interacting particles. We denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$.

To describe our interacting particle system, we fix k in $\mathbb{N}$, $k \geq 2$, and take a k-particle interaction potential V^N on S, i.e. a Borel function $V^N : S^k \to \overline{\mathbb{R}}$. We define the energy $E_V^N : S^N \to \overline{\mathbb{R}}$ of an N-particle configuration ($N \gg k$) by

$$E_V^N(x_1, \dots, x_N) := \frac{1}{N^k} \sum_{i_1,\dots i_k \text{ distinct}} V^N(x_{i_1}, \dots, x_{i_k}),$$

(when the N-dependence is made explicit in the superscript, with a little abuse of notation we use E_V^N instead of $E_{V^N}^N$). Then the interacting particle system is described by the probability measure $\mathbb{Q}_V^N$ on $(\Omega, \mathcal{A})$ defined by

$$\mathbb{Q}_V^N = \frac{1}{Z_V^N} e^{-N E_V^N(\omega_1,\dots\omega_N)} \mathbb{P},$$

where Z_V^N is the normalizing constant, assumed to be finite. Under $\mathbb{Q}_V^N$, the particles ω_i are subject to interaction via the potential V^N. Note that the energy E_V^N is invariant under permutation, or, in other words, depends only on the positions of the particle ω_i and not on their label i, via a fixed interaction potential V^N—this is a *mean-field* interaction. Note also that particle configurations $(x_1, \dots, x_N)$ are more likely, according to $\mathbb{Q}_V^N$, if V^N assumes lower values in these configurations.

Our main interest is in large deviations for the empirical measures associated with ω_i under the probability measure $\mathbb{Q}_V^N$. For this reason, we consider the state space $\mathcal{X} = \mathcal{P}(S)$, equipped with the weak topology (w.r.t. continuous and bounded functions on S), which turns $\mathcal{P}(S)$ into a Polish space [DZ10, Theorem D.8]. For each $N \in \mathbb{N}$, we denote by $z_\bullet^N : S^N \to \mathcal{P}(S)$ the continuous map

$$S^N \ni (x_1, \dots, x_N) =: \boldsymbol{x} \mapsto z_{\boldsymbol{x}}^N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \in \mathcal{P}(S).$$

We denote by $P^N \in \mathcal{P}(\mathcal{P}(S))$, resp. $Q_V^N \in \mathcal{P}(\mathcal{P}(S))$ the law of the empirical measure $z_{\boldsymbol{\omega}}^N$ for $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)$ under $\mathbb{P}$ (where ω_i $i = 1, \dots N$ are i.i.d. with common law μ_0), resp. under $\mathbb{Q}_V^N$. We aim at giving an LDP for Q_V^N.

We further define the function $\mathcal{E}_V^N : \mathcal{P}(S) \to \overline{\mathbb{R}}$ by

$$\mathcal{E}_V^N(\mu) := \begin{cases} \int_{(S^k)'} V^N \, d\mu^{\otimes k}, & \text{if } V^N \in L^1(\mu^{\otimes k}), \\ +\infty, & \text{otherwise}, \end{cases} \tag{4.1}$$

where $(S^k)'$ is S^k but with the diagonals removed, i.e.

$$(S^k)' := \left\{ (x_1, \dots, x_k) \in S^k \,\middle|\, x_i \neq x_j, \, \forall i, j \in \{1, \dots, k\} \text{ with } i \neq j \right\}.$$

Notice that when ω_i, $i = 1, \dots, N$ are i.i.d. random variables with common law μ_0, then

$$E_V^N(\omega_1, \dots, \omega_N) = \int_{(S^k)'} V^N d(z_{\boldsymbol{\omega}}^N)^{\otimes k} = \mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N) \qquad \mathbb{P}^{\otimes N}\text{-a.e.}$$

due to the non-atomic property of the reference measure μ_0. Hence, by construction and the mean-field interaction property, the interacting particle system may then be recast as a change of measure in $\mathcal{P}(\mathcal{P}(S))$, namely

$$Q_V^N := \frac{1}{Z_V^N} e^{-N \mathcal{E}_V^N(\mu)} P^N. \tag{4.2}$$

Given a Borel function $V : S^k \to \overline{\mathbb{R}}$, we define similarly the function $\mathcal{E}_V : \mathcal{P}(S) \to \overline{\mathbb{R}}$ by

$$\mathcal{E}_V(\mu) := \begin{cases} \int_{S^k} V \, d\mu^{\otimes k}, & \text{if } V \in L^1(\mu^{\otimes k}), \\ +\infty, & \text{otherwise}, \end{cases} \tag{4.3}$$

The functions $\mathcal{E}_V^N$ and $\mathcal{E}_V$ are Borel maps on $\mathcal{P}(S)$ (cf. Appendix C, in particular Lemma C.2 and Corollary C.4) and are defined identically except for the N dependence in V and the domain of integration, i.e., $(S^k)'$ instead of S^k.

Remark 4.1. A few comments on the assumption that μ_0 is non-atomic:

(i) The energy E_V^N as defined above does not rule out self-interaction, i.e., two particles occupying the same position ($x_i = x_j$ for $i \neq j$). While E_V^N is meaningful for bounded potentials V^N, this may cause an issue when V^N is singular on the diagonal. The non-atomicity of μ_0 resolves this issue: indeed, if ω_i are i.i.d. random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ with common law μ_0, then $\mathbb{P}(\omega_i = \omega_j,\ i \neq j) = 0$, which then allows $E_V^N(\omega_1, \dots, \omega_N)$ to be defined $\mathbb{P}$-a.s..

(ii) However, if the reference measure μ_0 is atomic but the energy for the N-particle system is still given by E_V^N there is also an alternative method. Namely, fix N and note that $E_V^N(\boldsymbol{x}) = E_V^N(\mathbf{y})$ for any $\mathbf{x}, \mathbf{y} \in S^N$ with $z_{\boldsymbol{x}}^N = z_{\mathbf{y}}^N$. Hence the function $\mathcal{E}_V^N : \mathcal{P}(S) \to \overline{\mathbb{R}}$

$$\mathcal{E}_V^N(\mu) := \begin{cases} E^N((z_\bullet^N)^{-1}(\mu)), & \mu \in z_\bullet^N(S^N), \\ +\infty, & \text{otherwise,} \end{cases}$$

is well-defined. Moreover, it is easy to verify that $z_\bullet^N(S^N)$ is closed and for $V^N \in C_b(S^N)$ the map $E^N((z_\bullet^N)^{-1}(\mu))$ is continuous on $z_\bullet^N(S^N)$, hence $\mathcal{E}_V^N(\mu)$ is Borel. By a monotone class argument similar as in Appendix C one can extend this to all Borel V^N, and we can then proceed as in the rest of this chapter. However, note that $\mathcal{E}_V^N$ might no longer be of integral form as in (4.1).

□

Now we give the classical results concerning LDPs. Recall that the relative entropy of ν with respect to μ is defined as

$$R(\nu\|\mu) := \begin{cases} \displaystyle\int_S \log\left(\frac{d\nu}{d\mu}\right) d\nu, & \text{if } \nu \ll \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

We start with the non-interacting case, namely Sanov's theorem:

Theorem 4.2 (Sanov's theorem)**.** *The family $\{P^N\}$ of laws of the non-interacting particle system satisfies an LDP with rate function $I : \mathcal{X} \to \overline{\mathbb{R}}$, where*

$$I(\mu) := R(\mu\|\mu_0).$$

For the LDP for the interacting particle systems, we introduce the following notation:

$$J_V(\mu) := \begin{cases} \mathcal{E}_V(\mu) + I(\mu), & \text{for } \mu \in D(I), \\ +\infty, & \text{otherwise,} \end{cases}$$

with $D(I) := \{\mu \mid R(\mu\|\mu_0) < \infty\}$, and, if $\inf_{\mathcal{X}} J_V > -\infty$,

$$\mathcal{F}_V(\mu) := J_V(\mu) - \inf_{\nu\in\mathcal{X}} J_V(\nu). \tag{4.4}$$

We now give an LDP in the case when $V^N = V$ is in $C_b(S^k)$. In this case, $\mathcal{E}_V$ is also continuous and bounded, and so the LDP for Q_V^N is essentially a consequence of the classical Varadhan Lemma. The only (and technical) difference with the classical Varadhan Lemma comes from the missing diagonal in $(S^k)'$, which in general causes $\mathcal{E}_V^N$ to not be continuous.

Lemma 4.3. *Suppose $V : S^k \to \mathbb{R}$ is continuous and bounded. Then $(\mathcal{E}_V^N, \mathcal{E}_V)$ induces an LDP (in the sense of Definition 3.5). In particular, the family $\{Q_V^N\}$ given by* (4.2) *satisfies an LDP with rate function $\mathcal{F}_V$.*

Proof. By applying Lemma C.2 k-times, we get that, for any continuous and bounded V, the function $\mathcal{E}_V$ is continuous and bounded on $\mathcal{P}(S)$. Hence, by the classical Varadhan Lemma (cf. Proposition 2.1), the couple $(\mathcal{E}_V, \mathcal{E}_V)$ induces an LDP in the sense of Definition 3.1.

Then, approximating $(\mathcal{E}_V^N, \mathcal{E}_V)$ with the constant family $(\mathcal{E}_V, \mathcal{E}_V)$, we have that $(\mathcal{E}_V^N, \mathcal{E}_V)$ induces an LDP by Theorem 3.8, provided we show that, for some $\gamma > 1$ and $K \in \mathbb{R}$,

$$\begin{aligned} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-\gamma N \mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N)} \right] &< +\infty, \\ \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N |\mathcal{E}_V^N - \mathcal{E}_V|(z_{\boldsymbol{\omega}}^N)} \right] &\le K \qquad \text{for every } \beta \ge 0. \end{aligned} \tag{4.5}$$

The first limit follows easily from the boundedness of V. For the second limit, we can bound away all the self-interactions to obtain

$$\begin{aligned} \left|\mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N) - \mathcal{E}_V(z_{\boldsymbol{\omega}}^N)\right| &= \frac{1}{N^k} \left| \sum_{i_1,\dots,i_k \text{ all distinct}} V(\omega_{i_1}, \dots, \omega_{i_k}) - \sum_{i_1,\dots,i_k} V(\omega_{i_1}, \dots, \omega_{i_k}) \right| \\ &\le \frac{1}{N^k} \frac{k(k-1)}{2} N^{k-1} \|V\|_\infty = \frac{k(k-1)}{2N} \|V\|_\infty. \end{aligned} \tag{4.6}$$

In the inequality above, we used that the number of k-tuples $(i_1, \dots i_k)$ with at least two equal indices is bounded by $N^{k-1}k(k-1)/2$. The second limit in (4.5) follows easily. □

Thus, we are in the same setting as in the previous section, i.e., we have created a large class of family of functions (V^N, V) such that their induced interacting particle systems satisfy a certain LDP. Hence, next, we will show how to extend this class by approximation.

4.2 Main results

We give our main result for this chapter, which serves as a tool for LDPs for Gibbs measures with a possibly discontinuous interaction potential. The result brings the general LDP result of Theorem 3.8 into the Gibbs measure context. In the following, $(f)^-$ denotes the negative part of the function f.

Theorem 4.4. *Let (V_λ^N, V_λ) be a family of Borel functions on S^k such that $(\mathcal{E}_{V_\lambda}^N, \mathcal{E}_{V_\lambda})$ induces an LDP. Let (V^N, V) be a family of Borel functions on S^k and assume that, for some $\gamma > 1$,*

$$\left.\begin{aligned} \limsup_{N\to\infty} \log \int_{S^k} e^{\gamma k |(V_\lambda^N)^-|}\, d\mu_0^{\otimes k} &< +\infty, \\ \log \int_{S^k} e^{\gamma k |(V_\lambda)^-|}\, d\mu_0^{\otimes k} &< +\infty, \end{aligned}\right\} \quad \text{for every } \lambda > 0, \tag{4.7}$$

and that, for some $K \in \mathbb{R}$,

$$\left.\begin{aligned} \limsup_{\lambda\to 0} \limsup_{N\to\infty} \log \int_{S^k} e^{\beta |V^N - V_\lambda^N|}\, d\mu_0^{\otimes k} &\le K, \\ \limsup_{\lambda\to 0} \log \int_{S^k} e^{\beta |V - V_\lambda|}\, d\mu_0^{\otimes k} &\le K. \end{aligned}\right\} \quad \text{for every } \beta \ge 0. \tag{4.8}$$

Then $(\mathcal{E}_V^N, \mathcal{E}_V)$ induces an LDP. In particular, the family of induced interacting systems Q_V^N satisfies an LDP with normalized rate function $\mathcal{F}_V$ given in (4.4).

Similarly to Theorem 3.8 for general LDPs, informally this result states that Q_V^N satisfy an LDP if there exists interacting potentials V_λ^N and V_λ which approximate V^N and V in an exponentially good way and whose corresponding interacting systems $Q_{V_\lambda}^N$ satisfy an LDP, for each λ. Again, this allows for the following generalization:

1. It allows for V to be discontinuous.
2. The only requirement on LDPs for approximants is that $(\mathcal{E}_{V_\lambda}^N, \mathcal{E}_{V_\lambda})$ induces an LDP, not that V_λ are continuous.

3. We can allow for the potential V^N in the interacting particle system to depend on $N \in \mathbb{N}$ (cf. Remark 4.7 below).

In the case where the sequence of functions V^N is constant and equal to V, an LDP follows whenever V satisfies the appropriate exponential moment condition.

Corollary 4.5. *Suppose that $V^N = V$ for all $N \in \mathbb{N}$, with V such that*

$$\int_{S^k} e^{\beta|V|}\, d\mu_0^{\otimes k} < \infty, \qquad \text{for all } \beta \geq 0. \tag{4.9}$$

Then $(\mathcal{E}_V^N, \mathcal{E}_V)$ induces an LDP.

Proof. The result follows from Lemma 4.3 and Theorem C.5. Indeed, by Theorem C.5, with the choice $\mu = \mu_0^{\otimes k}$, $X = S^k$, there exists a sequence $(V_\lambda) \subset C_b(S^k)$ such that

$$\lim_{\lambda \to 0} \log \int_{S^k} e^{\beta|V - V_\lambda|}\, d\mu_0^{\otimes k} = 0, \qquad \text{for any } \beta \geq 0.$$

Since the family V_λ induces an LDP by Lemma 4.3, both (4.7) and (4.8) are satisfied, and we can apply Theorem 4.4. □

Remark 4.6. Recall that in the definition of the energy E_V^N for the N-particle configuration we have excluded self-interaction, due to possible singularities in V. However, if V is bounded (but not necessarily continuous), Corollary 4.5 remains valid when the energy does include self-interactions, i.e.,

$$E_V^N(x_1, \ldots, x_N) := \frac{1}{N^k} \sum_{i_1, \ldots, i_k} V(x_{i_1}, \ldots, x_{i_k}).$$

In this case, $(\mathcal{E}_V^N, \mathcal{E}_V)$ induces an LDP with

$$\mathcal{E}_V^N(\mu) := \mathcal{E}_V(\mu) = \int_{S^k} V\, d\mu^{\otimes k}, \qquad \text{for all } \mu \in \mathcal{P}(S^k).$$

Indeed, this holds simply due to the estimate (4.6). □

Remark 4.7. The setting of Theorem 4.4 includes also the case of interactions among different number of particles. For example, let (N-independent) interaction potentials $U_k : S^\ell \to \mathbb{R}$, $k = 1, 2, 3$, be given and assume that the energy function E_U^N is the sum of these three interactions, i.e.

$$E_U^N(x_1, \ldots x_N) = \frac{1}{N^3} \sum_{i_1, i_2, i_3 \text{ distinct}} U_3(x_{i_1}, x_{i_2}, x_{i_3}) + \frac{1}{N^2} \sum_{i_1 \neq i_2} U_2(x_{i_1}, x_{i_2}) + \frac{1}{N} \sum_{i_1} U_1(x_{i_1}).$$

Therefore, by taking

$$V^N(x_1, x_2, x_3) = U_3(x_1, x_2, x_3) + \frac{N}{N-2} U_2(x_1, x_2) + \frac{N^2}{(N-1)(N-2)} U_1(x_1),$$

we see that $E_U^N = E_V^N$ (for $x_1, \dots x_N$ all distinct) and we are in the previous setting. □

Remark 4.8 (Chaoticity). Suppose for simplicity that $V^N = V$ for some potential V satisfying (4.9), and that there exists a unique minimizer $\bar{\mu} \in \mathcal{X}$ of $\mathcal{F}_V(\mu)$. Then clearly Q_V^N is $\bar{\mu}$-*chaotic*, in the sense that $Q_V^N \to \delta_{\bar{\mu}}$ weakly as $N \to \infty$ or, equivalently, that the k-marginal distributions weakly converge to $\bar{\mu}^{\otimes k}$ for every $k \in \mathbb{N}$. In Part I.B this property holds as well and, in particular, also for the time-marginals of the process, see Remark 9.5 for further details.

Even stronger, one can show that in this case *entropic chaoticity* holds.

Proposition 4.9. *Suppose that $V^N = V$ for some potential V satisfying* (4.9)*, and that there exists a unique minimizer $\bar{\mu} \in \mathcal{X}$ of $\mathcal{F}_V(\mu)$. Then*

$$\lim_{N\to\infty} \frac{1}{N} R(Q_V^N | P_{\bar{\mu}}^N) = 0, \tag{4.10}$$

where $P_{\bar{\mu}}^N$ is the law of the empirical measure $z_{\mathbf{X}^N}^N$ of N i.i.d. particles $X_1^N, \dots, X_N^N$ with common law $\bar{\mu}$.

We postpone the proof to after the proof of Theorem 4.4. Now, by a subadditivity argument, one can show that the relative entropies for the k-marginals with respect to $\bar{\mu}^{\otimes k}$ vanish as well. Moreover, in the case that $S = \mathbb{R}^d$ and V is sufficiently regular, the convergence (4.10) can even be made quantitative, see [Lac22].

We will see a similar phenomenon in Part II.B, where we establish that a property similar to (4.10) propagates in time, as seen in Theorem 15.9. □

We come now to the proof of Theorem 4.4, which relies heavily on Theorem 3.8 and the following bounds. These bounds convert the approximation properties for V into those for $\mathcal{E}_V$, needed to apply Theorem 3.8.

Lemma 4.10. *For any $k, N \in \mathbb{N}$ with $N > 1$, and nonnegative Borel functions $V, V^N : S^k \to \overline{\mathbb{R}}$, the following inequalities hold true:*

$$\frac{1}{N} \log \mathbb{E}\left[e^{N\mathcal{E}_V^N(z_\omega^N)}\right] \le \frac{1}{k} \log \int_{S^k} e^{k\frac{N}{N-1} V^N} d\mu_0^{\otimes k}, \tag{4.11a}$$

$$\sup_{\mu\in D} \big(\mathcal{E}_V(\mu) - R(\mu\|\mu_0)\big) \le \frac{1}{k} \log \int_{S^k} e^{kV} d\mu_0^{\otimes k}, \tag{4.11b}$$

where $D = \{\nu \in \mathcal{P}(S) \mid R(\nu\|\mu_0) < +\infty\}$.

Proof. For the proof of (4.11a), we use the Hoeffding decomposition [Hoe63] for $\mathcal{E}_V^N$, which reads

$$\mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N) = \frac{N!}{N^k(N-k)!}\frac{1}{N!}\sum_{\sigma\in\mathcal{S}_N}\frac{1}{[N/k]}\sum_{j=1}^{[N/k]}V^N(\omega_{\sigma(jk-k+1)},\dots\omega_{\sigma(jk)}),$$

where $\mathcal{S}_N$ is the group of permutations of $\{1,\dots N\}$ and $[N/k]$ denotes the integer part of N/k. The point of this decomposition is to group the elements $V^N(\omega_{i_1},\dots\omega_{i_k})$ into an average (over possible permutations σ) of averages (over j) of $O(N)$ independent elements (that is, for fixed σ, the elements within the nested average are independent).

By Jensen's inequality, applied to the exponential function and the average over σ,

$$\begin{aligned}\mathbb{E}\left[e^{N\mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N)}\right] &= \mathbb{E}\left[\exp\left[\frac{1}{N!}\sum_{\sigma\in\mathcal{S}_N}\frac{N!}{N^k(N-k)!}\frac{N}{[N/k]}\sum_{j=1}^{[N/k]}V^N(\omega_{(\sigma(jk-k+1))},\dots\omega_{\sigma(jk)})\right]\right]\\ &\le \frac{1}{N!}\sum_{\sigma\in\mathcal{S}_N}\mathbb{E}\left[\exp\left[\frac{N!}{N^k(N-k)!}\frac{N}{[N/k]}\sum_{j=1}^{[N/k]}V^N(\omega_{(\sigma(jk-k+1))},\dots\omega_{\sigma(jk)})\right]\right]\\ &= \mathbb{E}\left[\exp\left[\frac{N!}{N^k(N-k)!}\frac{N}{[N/k]}\sum_{j=1}^{[N/k]}V^N(\omega_{jk-k+1},\dots\omega_{jk})\right]\right],\end{aligned}$$

where we used the fact that ω_i are exchangeable (as i.i.d.) in the last line. Since the random variables $(\omega_{jk-k+1},\dots\omega_{jk})$ are independent in j, we have therefore

$$\begin{aligned}\mathbb{E}\left[e^{N\mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N)}\right] &\le \mathbb{E}\left[\prod_{j=1}^{[N/k]}\exp\left[\frac{N!}{N^k(N-k)!}\frac{N}{[N/k]}V^N(\omega_{jk-k+1},\dots\omega_{jk})\right]\right]\\ &= \prod_{j=1}^{[N/k]}\mathbb{E}\left[\exp\left[\frac{N!}{N^k(N-k)!}\frac{N}{[N/k]}V^N(\omega_{jk-k+1},\dots\omega_{jk})\right]\right]\\ &= \mathbb{E}\left[\exp\left[\frac{N!}{N^k(N-k)!}\frac{N}{[N/k]}V^N(\omega_1,\dots\omega_k)\right]\right]^{[N/k]},\end{aligned}$$

where we used again that ω_i are exchangeable in the second equality. Since $N! \le N^k(N-k)!$ and $N/[N/k] \le Nk/(N-1)$ for every $k, N \in \mathbb{N}$, $N > 1$, we then obtain

$$\mathbb{E}\left[e^{N\mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N)}\right] \le \mathbb{E}\left[\exp\left[\frac{N}{N-1}k\,V^N(\omega_1,\dots\omega_k)\right]\right]^{[N/k]}.$$

Taking the logarithm and noting that $[N/k] \le N/k$ for every $k, N \in \mathbb{N}$ yields

$$\frac{1}{N}\log\mathbb{E}\left[e^{N\mathcal{E}_V^N(z_{\boldsymbol{\omega}}^N)}\right] \le \frac{1}{k}\log\mathbb{E}\left[\exp\left[\frac{N}{N-1}k\,V^N(\omega_1,\dots\omega_k)\right]\right],$$

which concludes the proof of (4.11a).

To prove (4.11b), we first recall the additivity property of the relative entropy, i.e.

$$R(\mu^{\otimes k}\|\mu_0^{\otimes k}) = kR(\mu\|\mu_0).$$

Hence, for any $\mu \in D$, we have that

$$k\Big(\mathcal{E}_V(\mu) - R(\mu\|\mu_0)\Big) = \int_{S^k} kV\,d\mu^{\otimes k} - R(\mu^{\otimes k}\|\mu_0^{\otimes k}) \le \sup_{\nu\in D^k}\left\{\int_{S^k} kV\,d\nu - R(\nu\|\mu_0^{\otimes k})\right\},$$

where $D^k = \{\nu \in \mathcal{P}(S^k) \mid R(\nu\|\mu_0^{\otimes k}) < +\infty\}$. By Lemma B.1 we have that

$$\sup_{\nu\in D^k}\left\{\int_{S^k} kV d\nu - R(\nu\|\mu_0^{\otimes k})\right\} = \log\int_{S^k} e^{kV}\,d\mu_0^{\otimes k}.$$

Therefore, if the right-hand side is finite, the desired estimate (4.11b) follows. □

Proof of Theorem 4.4. In order to apply Theorem 3.8, we have to show that there exists some $\gamma > 1$, such that for every $\lambda > 0$,

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}\left[e^{-\gamma N\mathcal{E}^N_{V_\lambda}(z^N_\omega)}\right] < +\infty, \tag{4.12a}$$

$$\inf_{\mu\in D}\big(R(\mu\|\mu_0) + \gamma\mathcal{E}_{V_\lambda}(\mu)\big) > -\infty. \tag{4.12b}$$

and that, for some $K \in \mathbb{R}$, the following holds true for all $\beta \ge 0$:

$$\limsup_{\lambda\to 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}\left[e^{\beta N|\mathcal{E}^N_V - \mathcal{E}^N_{V_\lambda}|(z^N_\omega)}\right] \le K, \tag{4.13a}$$

$$\limsup_{\lambda\to 0}\sup_{\mu\in D}\big(\beta|\mathcal{E}_V - \mathcal{E}_{V_\lambda}|(\mu) - R(\mu\|\mu_0)\big) \le K. \tag{4.13b}$$

Due to linearity and (4.11a) of Lemma 4.10, we have for some $\gamma' \in (1,\gamma)$:

$$\frac{1}{N}\log\mathbb{E}\left[e^{-\gamma' N\mathcal{E}^N_{V_\lambda}(z^N_\omega)}\right] \le \frac{1}{N}\log\mathbb{E}\left[e^{\gamma' N\mathcal{E}^N_{\gamma'|(V_\lambda)^-|}(z^N_\omega)}\right] \le \frac{1}{k}\log\int_{S^k} e^{k\frac{N}{N-1}\gamma'|(V^N_\lambda)^-|}\,d\mu_0^{\otimes k}.$$

Since $\gamma' N/(N-1) \le \gamma$ for all $N \ge N_\gamma := \gamma/(\gamma-\gamma')$, we obtain from assumption (4.7) the finiteness of the right-hand side uniformly in N for $N \ge N_\gamma$. Hence, taking the lim sup yields (4.12a).

As for (4.12b), we apply (4.11b) of Lemma 4.10 to any $\mu \in \mathcal{P}(S)$ with $R(\mu\|\mu_0) < +\infty$, to obtain

$$R(\mu\|\mu_0)+\gamma'\mathcal{E}_{V_\lambda}(\mu) \ge R(\mu\|\mu_0)-\mathcal{E}_{\gamma'|(V_\lambda)^-|}(\mu) \ge -\frac{1}{k}\log\int_{S^k} e^{\gamma' k|(V_\lambda)^-|}\,d\mu_0^{\otimes k} > -\infty.$$

Similarly, we apply Lemma 4.10 to obtain

$$\frac{1}{N}\log\mathbb{E}\left[e^{\beta N|\mathcal{E}_V^N-\mathcal{E}_{V_\lambda}^N|(z_\omega^N)}\right]\le\frac{1}{k}\log\int_{S^k}e^{k\frac{N}{N-1}\beta|V^N-V_\lambda^N|}\,d\mu_0^{\otimes k},$$

$$\sup_{\mu\in D}\big(\beta|\mathcal{E}_V-\mathcal{E}_{V_\lambda}|-I\big)(\mu)\le\frac{1}{k}\log\int_{S^k}e^{k\beta|V-V_\lambda|}\,d\mu_0^{\otimes k},$$

which by assumption (4.8) yields (4.13a) and (4.13b). The conclusion follows applying Theorem 3.8. □

Proof of Proposition 4.9. The convergence (4.10) is clear for bounded continuous V and bounded continuous $\log d\bar{\mu}/d\mu_0$, since

$$\frac{1}{N}R(Q_V^N|P_{\bar{\mu}}^N)=-\int_{\mathcal{X}}\left(\mathcal{E}_V^N(\mu)+\langle\log\frac{d\bar{\mu}}{d\mu_0},\mu\rangle\right)Q_V^N(d\mu)-\frac{1}{N}\log Z_V^N,$$

$Q_V^N\to\delta_{\bar{\mu}}$ weakly as $N\to\infty$, and

$$-\lim_{N\to\infty}\frac{1}{N}\log Z_V^N=I(\bar{\mu})+\mathcal{E}_V(\bar{\mu}).$$

The case for V satisfying (4.9) now follows from Corollary 4.5, approximation arguments, and $L^p(\mu_0)$ estimates for $d\bar{\mu}/d\mu_0$. To be precise, note that for the convergence

$$\lim_{N\to\infty}\int_{\mathcal{X}}\mathcal{E}_V^N(\mu)\,Q_V^N(d\mu)=\mathcal{E}_V(\bar{\mu}),$$

it is sufficient to find a family of approximating potentials $\{V_\lambda\}_\lambda\subset S^k$ such that $\mathcal{E}_{V_\lambda}$ converges pointwise to $\mathcal{E}_V$ on the domain of I as $\lambda\to 0$, and

$$\lim_{\lambda\to 0}\limsup_{N\to\infty}\int_{\mathcal{X}}\left|\mathcal{E}_V^N(\mu)-\mathcal{E}_{V_\lambda}^N(\mu)\right|Q_V^N(d\mu)=0.$$

It can be shown that the family used in the proof of Corollary 4.5, satisfying

$$\lim_{\lambda\to 0}\log\int_{S^k}e^{\beta|V-V_\lambda|}\,d\mu_0^{\otimes k}=0,\qquad\text{for any }\beta\ge 0,$$

is a suitable candidate, adapting arguments from the proof of Theorem 4.4, and the fact that for all $\beta\ge 0$ we have the inequality

$$\int_{\mathcal{X}}\left|\mathcal{E}_V^N(\mu)-\mathcal{E}_{V_\lambda}^N(\mu)\right|Q_V^N(d\mu)\le\frac{1}{\beta N}\log\int_{\mathcal{X}}e^{\beta N|\mathcal{E}_V^N(\mu)-\mathcal{E}_{V_\lambda}^N(\mu)|}Q_V^N(d\mu).$$

Next, for the convergence

$$\lim_{N\to\infty}\int_{\mathcal{X}}\langle\log\frac{d\bar{\mu}}{d\mu_0},\mu\rangle\,Q_V^N(d\mu)=\mathcal{E}_V(\bar{\mu}),$$

we can employ another approximation argument, but only if $|\log d\bar{\mu}/\mu_0|$ is strongly exponentially integrable with respect to μ_0, or equivalently, if $d\bar{\mu}/d\mu_0 \in L^p(\mu_0)$ for all $p \in \mathbb{R}$. We will establish this for $k = 2$, and the case for arbitrary k follows in a similar way.

It is straightforward to verify that if the minimizer $\bar{\mu}$ of $\mathcal{F}_V$ is unique it is the fixed point of the map F defined by the relation

$$F(\mu)(dx) = \frac{1}{\int_S e^{-2\int_S V(x',y)\mu(dy)}\mu_0(dx')} e^{-2\int_S V(x,y)\mu(dy)}\mu_0(dx).$$

Suppose that V is bounded. One can then establish by Jensen's inequality, and duality of $\frac{1}{p}z^p$ and $\frac{1}{p'}z^{p'}$, that

$$\|d\bar{\mu}/d\mu_0\|^p_{L^p(\mu_0)} \leq \frac{1}{p}\|d\mu/d\mu_0\|^p_{L^p(\mu_0)} + \frac{1}{p'}\int_S e^{4pp'|V|(x,y)}\mu_0(dx)\mu_0(dy),$$

for any $p > 1$. This implies that for the fixed point $\bar{\mu}$ we have the bound

$$(1 - 1/p)\|d\bar{\mu}/d\mu_0\|^p_{L^p(\mu_0)} \leq \frac{1}{p'}\int_S e^{4pp'|V|(x,y)}\mu_0(dx)\mu_0(dy),$$

and corresponding bounds exist for any $p \in \mathbb{R}$. The desired estimates for arbitrary V now follow by approximation. □

Background As mentioned in Chapter 3, various extension principles to singular functionals or contractions have arisen. See for example [ES98] on various generalizations of Sanov's theorem, in which a stronger topology is used—defined by the property that all functions $\mathcal{E}_V$ for which $k = 1$ and V satisfies (4.9) are continuous with respect to this topology. This was subsequently generalized in [ES02] to the case $k \geq 2$. Note that an application of the classical Varadhan Lemma to the result in [ES02] should yield a very similar result to our Corollary 4.5; this type of argument has been used in [EZ03, Lemma 2.4] in the proof of a moderate deviation principle for a bounded interaction kernel.

More recently, in [LW20], a similar setting as this chapter was studied, i.e., large deviations for mean-field Gibbs measures on Polish spaces, involving singular potentials. In this work they assume so-called *strong exponential integrability* of the negative part V^- of the interaction potential V, and require an entropic inequality involving the positive part V^+ of the interaction potential. It is straightforward to verify that strong exponential integrability of $|V|$ is in fact a stronger assumption and is equivalent to (4.9), and, in particular, their results include the case of Corollary 4.5. In a sense this allows them to treat stronger repulsive singularities than the ones considered here. It should be noted that while both their work and Corollary 4.5 are not sufficient for the general setting of interacting diffusions as in Part

I.B, a similar asymmetric generalization involving variational or weak convergence techniques for Theorem 8.1 might be possible, which we leave for future work.

Other LDP results for Gibbs measures with singular potentials have been proven, see e.g. [Ber18, CGZ14, DLR20, Rey18], with locally compact base space S, [GZ19] with potentials satisfying a bound from below and a lower semi-continuity assumption. In particular, in [Ber18] the case where (4.11) only holds for some β instead of all—which, in $\mathbb{R}^d$, allows for potentials V with a logarithmic singularity—is considered, on compact Polish spaces and assuming V lower semi-continuous.

It should be noted that inequalities related to (4.11a), which in our case is derived from the Hoeffding decomposition, have also arisen in different contexts and under different names. For example, in [ES02] a similar inequality stems from the existence of regular partitions of complete hypergraphs due to Baranyai, and in [LW20] modifications of decoupling inequalities of de la Peña were used. Hoeffding decomposition has been used directly in some generalizations to [ES02], for example [Eic04].

Chapter 5

Extension to Gibbs-like potentials

The previous results use the Gibbs structure (4.3) of the potential $\mathcal{E}_V$ to reduce the LDP problem to the context of Chapter 3. However, for this reduction to hold, only some Gibbs-like bounds are needed. This allows us to prove LDPs for empirical measures not coming from Gibbs laws, as soon as Gibbs-type bounds are possible. As we will see in the example of Section 8.2.3, this is the case of interacting diffusions where the drift depends nonlinearly on the empirical measure and/or on its k-times tensor product.

As before, let S be a Polish space with its Borel σ-algebra $\mathcal{B}(S)$, $\mu_0 \in \mathcal{P}(S)$ be a given reference measure. The state space for the LDP is $\mathcal{X} = \mathcal{P}(S)$, equipped with the weak topology. As before, P^N denotes the law of the empirical measure $z^N_{\boldsymbol{\omega}}$, where ω_i, $i = 1, \dots N$ are i.i.d. random variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in S and common law μ_0; $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$. Now let $\mathcal{E}^N : \mathcal{P}(S) \to \overline{\mathbb{R}}$, $N \in \mathbb{N}$ and $\mathcal{E} : \mathcal{P}(S) \to \overline{\mathbb{R}}$ be Borel functions, let $Q^N \ll P^N$ be the probability measure on $\mathcal{P}(S)$ given by

$$Q^N = \frac{1}{Z^N} e^{-N\mathcal{E}^N(\mu)} P^N,$$

where Z^N is the renormalization constant, assumed to be finite. We further recall the notations $\mathcal{E}^N_V(\mu)$ and $\mathcal{E}^N_V(\mu)$ given in (4.1) and (4.3) for any Borel function $V : S^k \to \overline{\mathbb{R}}$. As before, we denote $D = \{\nu \in \mathcal{P}(S) \mid R(\nu \| \mu_0) < \infty\}$.

Our main result concerns the case where the difference between $(\mathcal{E}^N, \mathcal{E})$ and the approximating family $(\mathcal{E}^N_\lambda, \mathcal{E}_\lambda)$ can be bounded in a suitable way by Gibbs-type potentials.

Theorem 5.1. *Let* $(\mathcal{E}^N_\lambda, \mathcal{E}_\lambda)$ *be a family of LDP inducing pairs for all* $\lambda > 0$ *such that* (3.4) *holds for some* $\gamma > 1$ *(independent of* λ*). Assume that, for every* $\lambda > 0$

*and every $\beta \in \mathbb{R}$, for every μ in **D**, there holds*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N(\mathcal{E}^N - \mathcal{E}^N_\lambda)(z^N_\omega)}\right] \leq C + C \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{c_\beta N \mathcal{E}^N_{G^N_\lambda}(z^N_\omega)}\right], \quad (5.1a)$$

$$|\beta(\mathcal{E} - \mathcal{E}_\lambda)(\mu)| - R(\mu\|\mu_0) \leq C + C \log \int_S \exp\left(c_\beta \int_{S^{k-1}} G_\lambda(x,y)\, d\mu^{\otimes(k-1)}(y)\right) d\mu_0(x), \quad (5.1b)$$

for some constant $C > 0$ independent of β, λ and μ, some $c_\beta \geq 0$ independent of λ and μ, and some nonnegative Borel functions $G^N_\lambda, G_\lambda : S^k \to \overline{\mathbb{R}}$ (independent of β and μ). Assume also that G^N_λ and G_λ satisfy, for some $K \in \mathbb{R}$ independent of β, for every $\beta \geq 0$,

$$\limsup_{\lambda\to 0} \limsup_{N\to\infty} \int_{S^k} e^{\beta G^N_\lambda}\, d\mu_0^{\otimes k} \leq K, \quad (5.2a)$$

$$\limsup_{\lambda\to 0} \int_{S^k} e^{\beta G_\lambda} d\mu_0^{\otimes k} \leq K. \quad (5.2b)$$

Then the pair $(\mathcal{E}^N, \mathcal{E})$ induces an LDP with the normalized rate function $\mathcal{F}_V$ in (4.4).

Remark 5.2. A simple condition for the inequalities (5.1a) and (5.1b) to hold is if

$$\left|\mathcal{E}^N - \mathcal{E}^N_\lambda\right|(\mu) \leq \mathcal{E}^N_{G^N_\lambda}(\mu), \qquad \left|\mathcal{E} - \mathcal{E}_\lambda\right|(\mu) \leq \mathcal{E}_{G_\lambda}(\mu) \qquad \text{for all } \mu \in \mathcal{P}(\mathcal{X}).$$

□

Proof of Theorem 5.1. The proof is similar to that of Theorem 4.4. The result follows from Theorem 3.8 provided we verify condition (3.5). For condition (3.5a), by assumption (5.1a) and Lemma 4.10, we have

$$\begin{aligned}\limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N(\mathcal{E}^N - \mathcal{E}_\lambda)(z^N_\omega)}\right] &\leq C + C \limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{c_\beta N \mathcal{E}^N_{G_\lambda}(z^N_\omega)}\right] \\ &\leq C + C \limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{k} \log \int_{S^k} e^{k \frac{N}{N-1} c_\beta G^N_\lambda}\, d\mu_0^{\otimes k},\end{aligned}$$

for all $\beta \in \mathbb{R}$. Hence, by assumption (5.2a), we get

$$\limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N(\mathcal{E}^N - \mathcal{E}^N_\lambda)(z^N_\omega)}\right] \leq C + C \frac{\log K}{k}.$$

For condition (3.5b), by assumption (5.1b) and Lemma B.4, we have, for every μ in D, for every $\lambda > 0$ and every $\beta \in \mathbb{R}$,

$$
\begin{aligned}
&|\beta(\mathcal{E} - \mathcal{E}_\lambda)(\mu)| - (1 + C(k-1))R(\mu\|\mu_0) \\
&\le -C(k-1)R(\mu\|\mu_0) + C + C \log \int_S \exp\left(c_\beta \int_{S^{k-1}} G_\lambda(x,y)\, d\mu^{\otimes(k-1)}(y)\right) d\mu_0(x) \\
&\le -C(k-1)R(\mu\|\mu_0) + C + C(k-1)R(\mu\|\mu_0) + C \log \int_{S^k} e^{c_\beta G_\lambda} d\mu_0^{\otimes k}
\end{aligned}
$$

Hence, taking the sup over μ in D and then the lim sup over λ, by assumption (5.2b) we get

$$
\limsup_{\lambda\to 0} \sup_{\mu \in D} |\beta(\mathcal{E} - \mathcal{E}_\lambda)(\mu)| - (1 + C(k-1))R(\mu\|\mu_0) \le C + C \log K.
$$

Condition (3.5b) follows by simply dividing by $1+C(k-1)$. The proof is complete. □

Remark 5.3. As the above proof shows (and as a consequence of Lemmas 4.10 and B.4), the assumptions (5.2a) and (5.2b) imply, respectively, the following two bounds:

$$
\limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{c_\beta N \mathcal{E}^N_{G_\lambda}(z^N_\omega)}\right] \le \frac{\log K}{k} < \infty,
$$

$$
\limsup_{\lambda\to 0} \log \int_S \exp\left(c_\beta \int_{S^{k-1}} G_\lambda(x,y)\, d\mu^{\otimes(k-1)}(y)\right) d\mu_0(x) \le (k-1)R(\mu\|\mu_0) + \log K < \infty.
$$

□

Part I.B

Large deviations for singularly interacting diffusions

Chapter 6

Introduction

In this part we turn to the limiting behaviour of weakly interacting, or mean-field, diffusions, where the interaction depends only on the empirical measures of the particles. Aside from the simple example of (1.1) we consider for every $N \in \mathbb{N}$ the particle system defined by the coupled stochastic differential equations (SDEs)

$$\begin{cases} dX_t^{N,i} = b_t\left(X_t^{N,i}, \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}\right) dt + dW_t^i, \\ X_0^{N,i} \text{ i.i.d with law } \rho_0. \end{cases} \tag{6.1}$$

Here $W^1, \dots, W^N$ are independent d-dimensional Brownian motions, ρ_0 is a given initial distribution, and b is a measure-dependent *drift* vector. To retrieve a time-dependent version of (1.1) the drift vector b takes the form

$$b_t(x, \mu) = \int_{\mathbb{R}^d} \varphi(t, x - y)\, d\mu(y), \tag{6.2}$$

for some interaction kernel φ. Such types of drifts commonly appear in models of classical physical systems, biological systems such as the collective motion of micro-organisms (bacteria, cells, etc.), and flocking and swarming behaviour of animals, granular media, as well as models in opinion formation.

When b is sufficiently regular, the limiting behaviour of the particle system for a large particle number is well understood. For example, when b is Lipschitz and bounded the empirical measure

$$z_{\boldsymbol{X}}^N := \frac{1}{N}\sum_{i=1}^N \delta_{X^{N,i}},$$

converges to the law of the McKean–Vlasov equation [Szn91, Tan84]

$$\begin{cases} dX_t = b_t(X_t, \mathrm{Law}(X_t))\,dt + dW_t, \\ X_0 \text{ with law } \rho_0. \end{cases} \tag{6.3}$$

Moreover, a large deviation principle (LDP) holds for the empirical measure $z^N_{\boldsymbol{X}}$, as $N \to \infty$, see e.g. [DG87, BDF12, CDFM20].

The case of a *singular* interaction, that is irregular b, has been widely studied too. The convergence of the system (6.1) to the corresponding McKean–Vlasov equation has been shown for various examples of singular drifts, most of them of the form (6.2) with singular interaction kernel φ, e.g. [FHM14, GQ15, BO19, JW18, BJW19, Jab19] (see Chapter 9 for detailed explanations).

However, establishing LDPs for these singular drifts has remained unsolved. Apart from the work by Fontbona in [Fon04], where an LDP for the time-marginals of $(z^N_{\boldsymbol{X}})$ was shown for a repulsive kernel $\varphi(x) = 1/x$, little is known to our knowledge. We aim to fill this gap, by providing LDP results and new tools for a large class of singular measure-dependent drifts.

As the main example, we consider the following drift

$$b_t(x, \mu) := \psi\left(x, \mu, \int_{\mathbb{R}^d} \varphi(t, x - y)\,d\mu(y)\right), \tag{6.4}$$

where $\psi : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$, with $\mathcal{P}(\mathbb{R}^d)$ the space of probability measures equipped with the bounded Lipschitz metric d_{BL} given by (8.31). We will show an LDP when ψ is Lipschitz and the interaction kernel is in an appropriate L^p space. More precisely, combining several key statements throughout Part I.B (see Proposition 8.12, Remark 8.14 and Proposition 9.6), we obtain the following result:

Theorem 6.1. *Suppose that*

(i) (Lipschitz property) $\psi : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ *is jointly globally Lipschitz, i.e. there exists some constant* $M > 0$*, such that for all* $w, x, y, z \in \mathbb{R}^d$ *and* $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$*:*

$$|\psi(x, \mu, z) - \psi(y, \nu, w)| \le M\Big(|x - y| + d_{BL}(\mu, \nu) + |z - w|\Big);$$

(ii) (Linear growth) there exists a constant $L > 0$ *such that for all* $x, z \in \mathbb{R}^d$*,* $\mu \in \mathcal{P}(\mathbb{R}^d)$*:*

$$|\psi|(x, \mu, z) \le L(1 + |z|);$$

(iii) (Exponential moment) for all $\beta > 0$*, the initial distribution* ρ_0 *satisfies*

$$\int_{\mathbb{R}^d} e^{\beta|x|}\,d\rho_0(x) < \infty;$$

(iv) (Regularity) for $p, q \in [2, \infty]$ *with* $d/p + 2/q < 1$,

$$\varphi \in L^q\big((0,T), L^p(\mathbb{R}^d)\big) + L^\infty((0,T) \times \mathbb{R}^d).$$

Then the family $\{Q^N\}$ *of laws of empirical measures* $z^N_{\boldsymbol{X}}$ *for* $\boldsymbol{X} = (X^{N,1}, \dots, X^{N,N})$ *satisfying* (6.1)*, with drift* b *as in* (6.4)*, has an LDP with rate function*

$$\mathcal{F}(\mu) = \begin{cases} R(\mu \| \mathbb{W}^\mu), & \text{if } R(\mu\|\mathbb{W}) < \infty, \\ +\infty, & \text{otherwise.} \end{cases}$$

Here $R(\mu\|\nu)$ *is the relative entropy of* μ *w.r.t.* ν *defined by*

$$R(\nu\|\mu) := \begin{cases} \displaystyle\int \log\left(\frac{d\nu}{d\mu}\right) d\nu, & \text{if } \nu \ll \mu, \\ +\infty, & \text{otherwise,} \end{cases}$$

$\mathbb{W}^\mu$ *is the law of a process* X^μ_t *satisfying the SDE*

$$dX^\mu_t = b(X^\mu_t, \mu_t)\, dt + dW_t,$$

and $\mathbb{W} = \mathrm{Law}(W)$*, where* W *is a Brownian motion with initial law* ρ_0 *(and* μ_t *is the time-marginal of* μ *at time* t*).*

Furthermore, $z^N_{\boldsymbol{X}}$ *converges almost surely to the unique minimizer of* $\mathcal{F}(\mu)$*, which is the unique law of the solution to the McKean–Vlasov SDE* (6.3)*.*

An application of Theorem 6.1 with $q = \infty$ is the case of a drift b of the form (6.2), with

$$\varphi(t,z) = |z|^\alpha g\left(\frac{z}{|z|}\right) 1_{|z|\le R} + h(z) 1_{|z|>R},$$

with $g : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ and $h : \mathbb{R}^d \to \mathbb{R}^d$ both Borel bounded, $R > 0$, and exponent α satisfying

$$\alpha > -1 \ \text{ for } \ d \ge 2, \qquad \text{and} \qquad \alpha > -1/2 \ \text{ for } \ d = 1.$$

Here the inequality for $d = 1$ arises from the fact that we require $|z|^\alpha \in L^p_{loc}(\mathbb{R}^d)$ for some p with $p \ge 2$, while for $d > 1$ the constraint $p > d$ comes into play.

In fact, we prove an LDP and convergence to the McKean–Vlasov equation for systems with singular drifts under more general assumptions, see Theorem 8.1 and the examples in Section 8.2, 9.2, where we include many-particle interaction (that is, dependence of b on $\mu^{\otimes k}$) and interaction kernels φ merely satisfying

$$\mathbb{E}\left[e^{\beta \int_0^T |\varphi|^2(t, W^1_t, W^2_t)\, dt}\right] < \infty, \qquad \forall \beta \in \mathbb{R},$$

where W^1, W^2 are independent Brownian motions with common initial law ρ_0.

Note that even the convergence result to the McKean-Vlasov SDE in Theorem 6.1 is new: while some works do show convergence for the class of drifts (6.2) with even more singular φ [FHM14, JW18, BJW19], we are not aware of a result that covers drifts of the form (6.4) under our assumptions.

Our proof of the LDP relies on using a singular change of measure via Girsanov's theorem and an approximation by regular drifts. Indeed, Girsanov's theorem gives at least formally

$$\frac{dQ^N}{dP^N}(\mu) = e^{-N\mathcal{E}(\mu)}, \quad \mu \in \mathcal{P}(C([0,T];\mathbb{R}^d)), \tag{6.5}$$

where Q^N is the law of the empirical measure $z^N_{\boldsymbol{X}}$ of the interacting particle system $\boldsymbol{X} = (X^{N,1}, \dots, X^{N,N})$ satisfying (6.1), P^N is the law of the empirical measures $z^N_{\boldsymbol{W}}$ of N independent d-dimensional Brownian motions $\boldsymbol{W} = (W^1, \dots W^N)$ and

$$\begin{aligned} \mathcal{E}(\mu) &= \int V(x,\mu)\,\mu(dx), \\ V(x,\mu) &= -\int_0^T b_t(x_t,\mu_t)\cdot\, dx_t + \frac{1}{2}\int_0^T |b_t(x_t,\mu_t)|^2\, dt. \end{aligned} \tag{6.6}$$

By Sanov's theorem, $\{P^N\}$ satisfies an LDP with rate function $R(\cdot\|\mathbb{W})$. We need to transfer this LDP to an LDP for $\{Q^N\}$ and, to deal with this, we employ our Extended Varadhan's Lemma, Theorem 3.8, or to be precise, the Gibbs-like setting of Theorem 5.1.

Their use lies in the fact that even for regular drifts b, the change of measure (6.5) provided via Girsanov's theorem is not necessarily continuous, while there are classical results on LDPs for interacting particle systems with regular drifts. Hence we can prove Theorem 6.1 by approximating the singular drift b in (6.6) with regular drifts b_λ and applying then the extended Varadhan's Lemma.

Other extensions and applications of Varadhan's Lemma have also been developed to deal with this, for example to prove LDPs for weakly interacting diffusions with regular drifts [PdH96, DMZ03, DFMS18]. Moreover, in [DFMS18], Mario Maurelli and others developed an enhanced version of Sanov's theorem in the rough path setting, which allows for Varadhan's Integral Lemma to be applied. However, to my knowledge, none of them have been used to prove LDPs for weakly interacting diffusions with *singular* interaction.

It should be noted that one cannot expect to establish an LDP via a change of measure for every singular drift. This is indeed the case when the law of the interacting particle system is not absolutely continuous with respect to the law of the non-interacting system. A particular example in which this occurs is the Keller-Segel model with $\varphi(t, z) = -\nabla \log z$, for which a notion of propagation of chaos was established in [BJW19] but an LDP result remains open.

However, even in the large class of drifts for which the system can be described via a change of measure, there is still a gap between those for which there is a known LDP and propagation of chaos, and those for which others have merely shown propagation of chaos. We believe that not only are our results such as Theorem 6.1 a sizable step in closing this gap but that the general tools that are provided will help to close it even further.

Outline With the results of Part I.A at hand, establishing an LDP for a system of weakly interacting diffusions amounts to (1) proving a (change-of-measure) representation formula (Girsanov's formula) for the laws $\{Q^N\}$ of the empirical measures $z^N_{\boldsymbol{X}}$ associated to the solution $\boldsymbol{X} = (X^{N,1}, \dots, X^{N,N})$ of (6.1); and (2) proving the existence of a family $\{b^N_\lambda\}$ of "exponentially good" approximations for b (cf. Theorem 8.1 and the concrete examples in Section 8.2), which implies Theorem 6.1 considered above with a drift b specified by (6.2). For drifts b satisfying the assumption of Theorem 6.1, and for all the examples in Sections 8.2, we further show that the rate function $\mathcal{F}$ associated to $\{Q^N\}$ attains a unique minimizer (see Chapter 9 for the general case), which implies almost sure convergence.

Chapter 7

Preliminaries and representation results

In order to study LDPs for systems of mean-field interacting diffusions with singular drift, we will put the problem in the framework of Gibbs-like structure of Chapter 5). We will give representation results for the log-densities and the rate functional, and provide an LDP for a class of suitably regular drifts.

7.1 Notations and preliminary results

For $N \in \mathbb{N}$, we consider the following system of interacting SDEs

$$\begin{cases} dX_t^{N,i} = b_t^N \left(X_t^{N,i}, \frac{1}{N} \sum_{j=1}^N \delta_{X_t^{N,j}} \right) dt + dW_t^i, \quad i = 1, \dots N, \\ X_0^{N,i} \text{ i.i.d with law } \rho_0. \end{cases} \tag{7.1}$$

Here the drift $b^N : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ is a Borel map, where $\mathcal{P}(\mathbb{R}^d)$ is endowed with the (metrizable, complete and separable) topology of weak convergence. The processes W^i, $i \in \mathbb{N}$, are independent d-dimensional Brownian motions on a standard filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_t, \mathbb{P})$, $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$ is the law of the i.i.d. initial data $X_0^{N,i}$.

We now set $S := C([0,T]; \mathbb{R}^d)$, the space of continuous paths in $\mathbb{R}^d$, and note that $X^{N,i} \in S$ for each $i = 1, \dots, N$. Moreover, we set $\tilde{Q}^N_{b^N} \in \mathcal{P}(S^N)$ to be the law of $\boldsymbol{X}^N := (X^{N,1}, \dots, X^{N,N})$ defined by the system (7.1). We set $\mathbb{W} \in \mathcal{P}(S)$ to be the Wiener measure with ρ_0 as marginal at time 0 and $\tilde{P}^N \in \mathcal{P}(S^N)$ the law of $N \in \mathbb{N}$ independent Brownian motions, i.e. $\tilde{P}^N := \mathbb{W}^{\otimes N}$. With a little abuse of notation, we will consider W^i to be d-dimensional independent Brownian motions with initial law ρ_0, unless differently specified; similarly, we will use W for a d-dimensional Brownian motion with initial law ρ_0 and $\boldsymbol{W} := (W^1, \dots, W^N)$, unless differently specified.

Consider the *empirical process* $z^N_{\boldsymbol{X}} \in \mathcal{P}(S)$,

$$z^N_{\boldsymbol{X}} := \frac{1}{N} \sum_{i=1}^{N} \delta_{X^{N,i}},$$

and let $Q^N_{b^N} \in \mathcal{P}(\mathcal{P}(S))$ be the laws of the random variable $z^N_{\boldsymbol{X}}$ induced by $\tilde{Q}^N_{b^N}$. Similarly, let P^N be the law for the empirical process $z^N_{\boldsymbol{W}}$ of the non-interacting system induced by $\tilde{P}^N$. We will use $z^N_{\boldsymbol{x}}$, for $\boldsymbol{x} \in S^N$, to denote the empirical measure associated with $\boldsymbol{x}$, and z^N for the canonical process (that is, the identity process) on $\mathcal{P}(S)$. We will often use the notation $\langle f, \mu \rangle$ to denote $\int f \, d\mu$.

Recall that, as in Chapter 4, the sequence P^N satisfies an LDP with rate function

$$I : \mathcal{P}(S) \to \overline{\mathbb{R}}; \qquad I(\mu) := R(\mu \| \mathbb{W}),$$

where $R(\mu\|\mathbb{W})$ is the relative entropy of μ with respect to the Wiener measure $\mathbb{W}$.

For the particle system, a Gibbs-like representation holds. Indeed, for a Borel map $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$, we introduce the potentials

$$V^N_b(x,\mu) := -V^{N,2}_b(x,\mu) + \frac{1}{2} V^1_b(x,\mu) \quad \text{with} \quad \begin{cases} V^1_b(x,\mu) := \displaystyle\int_0^T |b_t(x_t,\mu_t)|^2 dt, \\ V^{N,2}_b(x,\mu) := \displaystyle\int_0^T b_t(x_t,\mu_t) \cdot dx_t, \end{cases} \tag{7.2}$$

where $V^{N,2}_b : S \times \mathcal{P}(S) \to \mathbb{R}$ is defined as stochastic integral under the law of $(W^i, z^N_{\boldsymbol{W}})$ on (x,μ). We define the corresponding log-density as

$$\mathcal{E}^N_b(\mu) := \begin{cases} \int_S V^N_b(x,\mu)\, d\mu(x), & \text{if } V^1_b(\cdot,\mu), V^{N,2}_b(\cdot,\mu) \in L^1(\mu), \\ 0, & \text{otherwise}, \end{cases} \qquad \mu \in \mathcal{P}(S).$$

See Section 7.2 for the precise definition and measurability properties of $\mathcal{E}^N_b$. Similarly we define V_b and $\mathcal{E}_b$ by replacing $V^{N,2}_b$ with

$$V^2_b(x,\mu) := \int_0^T b_t(x_t,\mu_t) \cdot dx_t,$$

now as stochastic integral at a deterministic μ. See again Section 7.2 for the precise definition of $\mathcal{E}_b$ and its measurability properties. The Gibbs-like representation of the particle system is given by the following theorem, which is essentially a consequence of Girsanov's theorem and the mean-field form of the interaction (b^N depending on $z^N_{\boldsymbol{X}}$), with the details of the proof are postponed to Section 7.2.

Theorem 7.1. *Fix $N \in \mathbb{N}$. Assume that*

$$\mathbb{E}\left[\exp\left(\frac{N}{2}\int_S\int_0^T |b_t^N(x_t, z^N_{\boldsymbol{W},t})|^2 dt\, dz^N_{\boldsymbol{W},t}(x)\right)\right] < \infty.$$

Then there exists a weak solution to the system (7.1)*, which is unique under the constraint*

$$\int_0^T |b_t^N(x_t, z_t^N)|^2 dt \in L^1(S, z^N) \qquad \textit{for } Q^N_{b^N}\textit{-a.e. } z^N. \tag{7.3}$$

For this law, we have the following representations:

$$\frac{d\tilde{Q}^N_{b^N}}{d\tilde{P}^N}(x^1, \dots x^N) = e^{-N\mathcal{E}^N_{b^N}(z^N_{\boldsymbol{x}})}, \qquad \frac{dQ^N_{b^N}}{dP^N}(\mu) = e^{-N\mathcal{E}^N_{b^N}(\mu)}. \tag{7.4}$$

7.2 Log-densities and energy functionals

Here we give the precise definition of $\mathcal{E}^N_b$ and $\mathcal{E}_b$ in Section 7.1.

We recall that $S = C([0,T];\mathbb{R}^d)$ and $\mathbb{W}$ is the Wiener measure on S. For $b : [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ a Borel function, we define

$$V_b^1(x,\mu) := \int_0^T |b_t(x_t,\mu_t)|^2 dt, \quad x \in S,\ \mu \in \mathcal{P}(S).$$

Note that V_b^1 is a Borel map (by Fubini theorem applied to the map $(t,x,\mu) \mapsto |b_t(x_t,\mu_t)|^2$). For μ in $\mathcal{P}(S)$, if $V_b^1(x,\mu)$ is finite for $\mathbb{W}$-a.e. x, then we can define the stochastic integral

$$\int_0^T b_t(x_t,\mu_t)\cdot dx_t, \tag{7.5}$$

on the space $(S, (\mathcal{G}_t)_t, \mathbb{W})$, where $\mathcal{G}_t$ is the σ-algebra generated on S by the $\mathbb{W}$-negligible sets and by the projection $\pi_{[0,t]} : S \to S_t = C([0,t];\mathbb{R}^d)$ on $[0,t]$; in particular, $\mathcal{G}_T$ is the completion of $\mathcal{B}(S)$ with respect to $\mathbb{W}$. Hence we can define a map $V_b^2(\cdot,\mu) : S \to \mathbb{R}$ which is a representative of the stochastic integral (7.5) and is measurable with respect to $\mathcal{B}(S)$.

Now we define for every $\mu \in \mathcal{P}(S)$

$$\mathcal{E}_b(\mu) := \begin{cases} \int_S \left(\frac{1}{2}V_b^1(x,\mu) - V_b^2(x,\mu)\right) d\mu, & \text{if } V_b^1(\cdot,\mu), V_b^2(\cdot,\mu) \in L^1(S,\mu),\\ 0, & \text{otherwise}\end{cases}$$

(for μ which is absolutely continuous with respect to $\mathbb{W}$, $\mathcal{E}(\mu)$ does not depend on the specific choice of $V_b^2(\cdot,\mu)$). Note that, if $R(\mu\|\mathbb{W})<\infty$ and $\mathbb{E}[e^{\frac{1}{2}V_b^1(W,\mu)}]<\infty$, then $V_b^1(\cdot,\mu)$ and $V_b^2(\cdot,\mu)$ are in $L^1(\mu)$: indeed, by Lemma B.1,

$$\frac{1}{2}\int_S V_b^1(x,\mu)\,d\mu \le R(\mu\|\mathbb{W}) + \log\int_S e^{\frac{1}{2}V_b^1(x,\mu)} d\mathbb{W} < \infty,$$

and, by Lemmas B.1 and D.2 (using $e^{|a|}\le e^a+e^{-a}$),

$$\begin{aligned}\frac{1}{2}\int_S |V_b^2(x,\mu)|d\mu &\le R(\mu\|\mathbb{W}) + \log\left(\int_S e^{\frac{1}{2}V_b^2(x,\mu)}d\mathbb{W} + \int_S e^{-\frac{1}{2}V_b^2(x,\mu)}d\mathbb{W}\right)\\ &\le R(\mu\|\mathbb{W}) + \frac{1}{2}\log\left(4\int_S e^{\frac{1}{2}V_b^1(x,\mu)}d\mathbb{W}\right) < \infty.\end{aligned}$$

In Lemma 7.2 we show that $\mathcal{E}_b$ is Borel (at least on the set $\{\mu \mid R(\mu\|\mathbb{W})<\infty\}$).

Coming to $\mathcal{E}_b^N$, we recall that W^i, $1\le i\le N$, are independent d-dimensional Brownian motions on some filtered probability space $(\Omega,(\mathcal{F}_t)_t,\mathbb{P})$ (under the standard assumption) and $z_{\boldsymbol{W}}^N$ is the empirical measure associated with W^i. We assume on b that

$$\int_S V_b^1(x,z_{\boldsymbol{W}}^N)\,dz_{\boldsymbol{W}}^N(x) < \infty \qquad \mathbb{P}\text{-a.s..}$$

Under this assumption, we can define the stochastic integral

$$\int_0^T b_t(x_t,\mu_t)\cdot dx_t \tag{7.6}$$

on the space $(\bar\Omega,(\mathcal{H}_t)_t,\bar P)$. Here $\bar\Omega = S\times\mathcal{P}(S)$ and $\bar P$ is the law of $W^1\otimes z_{\boldsymbol{W}}^N$, or equivalently of $W^i\otimes z_{\boldsymbol{W}}^N$ for any $1\le i\le N$, under $\mathbb{W}$. Also $\mathcal{H}_t$ is the σ-algebra on $\bar\Omega$ generated by the $\bar P$-negligible sets and by $\pi_{[0,t]}\otimes(\pi_{[0,t]})_\#$, where $\pi_{[0,t]}: S\to S_t = C([0,t];\mathbb{R}^d)$ is the projection on time $[0,t]$ and $(\pi_{[0,t]})_\#:\mathcal{P}(S)\to\mathcal{P}(S_t)$ is the corresponding image measure map; in particular, $\mathcal{H}_T$ is the completion under $\bar P$ of $\mathcal{B}(S)\otimes\mathcal{B}(\mathcal{P}(S))$. Hence we can define a map $V_b^{2,N}: S\times\mathcal{P}(S)\to\mathbb{R}$ which is a representative of the stochastic integral (7.6) and is measurable with respect to $\mathcal{B}(S)\otimes\mathcal{B}(\mathcal{P}(S))$. Note that, for every $1\le i\le N$,

$$V_b^{2,N}(W^i,z_{\boldsymbol{W}}^N) = \int_0^T b_t(W_t^i,z_{\boldsymbol{W},t}^N)\,dW_t^i \qquad \mathbb{P}\text{-a.s.} \tag{7.7}$$

and that, $\mathbb{P}$-a.s., $V_b^{2,N}(\cdot,z_{\boldsymbol{W}}^N)$ is in $L^1(z_{\boldsymbol{W}}^N)$.

Now we define for every $\mu \in \mathcal{P}(S)$.

$$\mathcal{E}_b^N(\mu) := \begin{cases} \int_S \left(\frac{1}{2}V_b^1(x,\mu) - V_b^{2,N}(x,\mu)\right) d\mu & \text{if } V_b^1(\cdot,\mu), V^{2,N}(\cdot,\mu) \in L^1(S,\mu), \\ 0 & \text{otherwise.} \end{cases}$$

By Theorem C.1, the function $\mathcal{E}_b^N$ is Borel.

Lemma 7.2. *Assume that* $\mathbb{E}[e^{\beta V_b^1(W,\mu)}] < \infty$ *for every* $\beta > 0$. *Then the map*

$$\mathcal{P}(S) \ni \mu \mapsto \mathcal{E}_b(\mu) 1_{R(\mu\|\mathbb{W})<\infty} \quad \textit{is Borel.}$$

Proof. The map

$$\mu \mapsto \int_S \frac{1}{2} V_b^1(x,\mu)\, d\mu$$

is Borel by Theorem C.1, so it is enough to show Borel measurability of

$$\begin{aligned} F : \mu \mapsto & \left(\int_S \frac{1}{2} V_b^2(x,\mu)\, d\mu\right) 1_{R(\mu\|\mathbb{W})<\infty} \\ &= \mathbb{E}\left[\int_0^T b_t(x_t,\mu_t)\cdot dW_t \frac{d\mu}{d\mathbb{W}}(W)\right] 1_{R(\mu\|\mathbb{W})<\infty}. \end{aligned}$$

We start with the case of b in $C_b([0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d))$ ($\mathcal{P}(\mathbb{R}^d)$ being endowed with the weak topology). We take a sequence Π^n of partitions $0 = t_0 < t_1 < \ldots < t_n = T$ on $[0,T]$ with size tending to 0 in n. For each n, we call $b^n : [0,T]\times S\times\mathcal{P}(S) \to \mathbb{R}$ the Borel function

$$b^n(t,x,\mu) = \sum_i b(t_i, x_{t_i}, \mu_{t_i}) 1_{t\in[t_i,t_{i+1})}$$

and $I^n : S\times\mathcal{P}(S) \to \mathbb{R}$ the Borel function defined by

$$I^n(\gamma,\mu) = \sum_i b(t_i,\gamma_{t_i},\mu_{t_i})\cdot(\gamma_{t_{i+1}} - \gamma_{t_i}).$$

Again by Theorem C.1, the map

$$F_n : \mu \mapsto \mathbb{E}[I^n(W,\mu)\frac{d\mu}{d\mathbb{W}}(W)] 1_{R(\mu\|\mathbb{W})<\infty}$$

is Borel. Now, for each μ with $R(\mu\|\mathbb{W}) < \infty$, we have by Lemmas B.1 and D.2, for every $\beta > 0$,

$$\begin{aligned} \beta|F(\mu) - F_n(\mu)| &\le \log \mathbb{E}[e^{\beta|V_b^2(W,\mu) - I^n(W,\mu)|}] + R(\mu\|\mathbb{W}) \\ &= \log \mathbb{E}[e^{\beta|\int_0^T (b_t(W_t,\mu_t) - b^n(t,W,\mu))\cdot dW_t|}] + R(\mu\|\mathbb{W}) \\ &\le \frac{1}{2}\log\left(\mathbb{E}[e^{2\beta^2\int_0^T |b_t(W_t,\mu_t) - b^n(t,W,\mu)|^2\cdot dt}]\right) + R(\mu\|\mathbb{W}). \end{aligned}$$

Since b is continuous, $b_t(W_t,\mu_t) - b^n(t,W,\mu)$ tends to 0 for every t in $[0,T)$ and every W and μ. Hence, for every fixed $\beta > 0$, by dominated convergence theorem and boundedness of b, $\mathbb{E}[e^{2\beta^2 \int_0^T |b_t(W_t,\mu_t)-b^n(t,W,\mu)|^2 \cdot dt}]$ tends to 1 and so

$$\limsup_n |F(\mu) - F_n(\mu)| \leq \frac{1}{\beta} R(\mu\|\mathbb{W}).$$

By arbitrariness of β, F is the pointwise limit of the Borel functions F_n, hence F is Borel (for b continuous and bounded).

The case of b Borel bounded follows from the case of b continuous bounded via a monotone class argument (cf. Theorem C.3): the stability assumption needed for the monotone class theorem can be verified as in the proof of convergence of F_n to F. Finally, the case of general b (satisfying $\mathbb{E}[e^{\beta V_b^1(W,\mu)}] < \infty$ for every β) follows approximating b with bounded b^n and proceeding as in the proof of convergence of F_n to F. The proof is complete. □

We now turn to the proof of the representation result Theorem 7.1.

Proof of Theorem 7.1. The SDE (7.1) is an SDE on $\mathbb{R}^{dN}$ for the vector $\boldsymbol{X}^N$, where the i-th component of the drift is $(t,\boldsymbol{x}) \mapsto b_t^N(x_i, z_{\boldsymbol{x}}^N)$. Note that, for this SDE, Novikov's condition is satisfied, indeed

$$\mathbb{E}\left[e^{\frac{1}{2}\sum_{i=1}^N \int_0^T |b_t^N(W_t^i, z_{\boldsymbol{W},t}^N)|^2 dt}\right] = \mathbb{E}\left[\exp\left(\frac{N}{2}\int_S \int_0^T |b_t^N(x_t, z_{\boldsymbol{W},t}^N)|^2 dt\, dz_{\boldsymbol{W}}^N(x)\right)\right] < \infty.$$

Girsanov's theorem gives then the existence of a weak solution $\boldsymbol{X}$, with law $\tilde{Q}_{b^N}^N$. The uniqueness in law condition (D.3) reads here

$$\sum_{i=1}^N \int_0^T |b_t^N(x_t^i, z_{\boldsymbol{x},t}^N)|^2 dt < \infty \quad \tilde{Q}_{b^N}^N\text{-a.s.},$$

which is equivalent to (7.3). The representation formula of the law $\tilde{Q}_{b^N}^N$ in Girsanov's theorem reads here (recall the definition of V_b^1 and $V_b^{2,N}$ in (7.2))

$$\begin{aligned}
\frac{d\tilde{Q}_{b^N}^N}{d\tilde{P}^N}(\boldsymbol{W}) &= \exp\left(-\frac{1}{2}\sum_{i=1}^N \int_0^T |b_t^N(W_t^i, z_{\boldsymbol{W},t}^N)|^2\, dt + \sum_{i=1}^N \int_0^T b_t^N(W_t^i, z_{\boldsymbol{W},t}^N)\cdot dW_t^i\right) \\
&= \exp\left(-\frac{1}{2} N \int_S V_{b^N}^1(x, z_{\boldsymbol{W}}^N) dz_{\boldsymbol{W}}^N(x) + N \int_S V_{b^N}^{2,N}(x, z_{\boldsymbol{W}}^N) dz_{\boldsymbol{W}}^N(x)\right) \\
&= \exp\left(-N\mathcal{E}_{b^N}^N(z_{\boldsymbol{W}}^N)\right),
\end{aligned}$$

where we used (7.7) for the stochastic integral. The first formula in (7.4) is proved. The second formula (for the law $Q_{b^N}^N$ of the empirical measure) follows from the first one by a standard argument from measure theory. □

7.3 Large deviations for regular drifts

Now we provide a class of drifts (b^N, b) which induce an LDP for the law $Q^N_{b^N}$. Morally, this is the class of Lipschitz drifts (with respect to the 1-Wasserstein distance). The class of $\mathcal{F}$Lip-inducing pairs in the definition below allows for some margin to include the case of drifts without self-interactions, similarly to Lemma 4.3. In the following, $\mathcal{P}_1(\mathbb{R}^d)$ denotes the subset of $\mathcal{P}(\mathbb{R}^d)$ of all probability measures on $\mathbb{R}^d$ with finite first moment; W_1 denotes the 1-Wasserstein distance on $\mathcal{P}_1(\mathbb{R}^d)$.

Definition 7.3. The class $\mathcal{F}$Lip consists bounded Borel functions $b : [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ such that the map $(x, \mu) \mapsto b(t, x, \mu)$ is globally Lipschitz continuous on $\mathcal{P}_1(\mathbb{R}^d)$ with respect to the 1-Wasserstein distance, uniformly in $t \in [0,T]$, i.e., for some $M_b \geq 0$,

$$|b(t, x, \mu) - b(t, y, \nu)| \leq M_b\big(|x - y| + W_1(\mu, \nu)\big),$$

for all $t \in [0,T]$, $(x, y) \in \mathbb{R}^d$ and $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$.

Moreover, the pair $(\{b^N\}, b)$ (in short (b^N, b)) is called $\mathcal{F}$Lip-inducing (subjected to $\{P^N\}$) if it satisfies the following conditions:

1. $b \in \mathcal{F}$Lip;
2. The sequence b^N is uniformly bounded, i.e. $\sup_{N\in\mathbb{N}} \|b^N\|_\infty < \infty$;
3. There exists a sequence c_N with $c_N \to 0$ as $N \to \infty$ such that

$$\int_0^T \big\langle |b_t - b_t^N|^2(\cdot, z_t^N), z_t^N \big\rangle\, dt \leq c_N, \qquad \text{for } P^N\text{-almost every } z^N. \quad (7.8)$$

□

Lemma 7.4. *Assume that the initial law ρ_0 satisfies*

$$\int_{\mathbb{R}^d} e^{\beta|x|}\, d\rho_0(x) < \infty, \qquad \forall \beta > 0. \quad (7.9)$$

Assume that (b^N, b) is $\mathcal{F}$Lip-inducing. Then $(\mathcal{E}^N_{b^N}, \mathcal{E}_b)$ induces an LDP in the sense of Definition 3.5.

Remark 7.5. The exponential condition (7.9) is required only to apply [CDFM20, Theorem 34]: that result needs (7.9) because it works with the 1-Wasserstein topology instead of the weak topology (the LDP in weak topology follows then by the contraction principle). For this reason, we suspect that in the space $\mathcal{P}(S)$ with the weak topology, the condition (7.9) is not necessary. □

Proof of Lemma 7.4. The LDP induced by $(\mathcal{E}_b^N, \mathcal{E}_b)$ follows from [CDFM20, Theorem 34] (see also [Fis14]). There an LDP is proved for Q_b^N with rate function $R(\mu\|\mathbb{W}^\mu)$, where $\mathbb{W}^\mu$ is defined as in Theorem 8.1 (as the law of the solution to (8.5)). By Lemma 7.6, we have

$$R(\mu\|\mathbb{W}^\mu) = R(\mu\|\mathbb{W}) - \mathcal{E}_b(\mu).$$

In particular, the infimum of the $R(\cdot\|\mathbb{W}) - \mathcal{E}_b$ is the infimum of the left-hand side above, that is 0. The Laplace principle is then trivially satisfied as $e^{-N\mathcal{E}_b}(\mu)$ is the density of the Girsanov's transform.

Hence, note that by Theorem 3.8 it is enough to show (3.4) and (3.5a) for $\mathcal{E}_\lambda^N :=$ $\mathcal{E}_{b^N}^N$ and $\mathcal{E}_\lambda = \mathcal{E}^N = \mathcal{E} := \mathcal{E}_b$. To derive (3.4a), we have for any $\gamma \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}\left[e^{\gamma N\langle V_{b^N}^{N,2}(\cdot,z_{\boldsymbol{W}}^N),z_{\boldsymbol{W}}^N\rangle}\right] &= \mathbb{E}\left[e^{\gamma\sum_{i=1}^N\int_0^T b_t^N(W_t^i,z_{\boldsymbol{W},t}^N)\cdot dW_t^i}\right]\\ &\le \mathbb{E}\left[e^{2\gamma^2\sum_{i=1}^N\int_0^T |b_t^N(W_t^i,z_{\boldsymbol{W},t}^N)|^2dt}\right]^{\frac{1}{2}} \le e^{\gamma^2NT\|b^N\|_\infty^2}.\end{aligned} \tag{7.10}$$

where Lemma D.2 was applied in the first inequality. Since $V_{b^N}^1$ is trivially bounded by $T\|b^N\|_\infty^2$, we then obtain

$$\mathbb{E}\left[e^{-\gamma N\mathcal{E}^N(z_{\boldsymbol{W}}^N)}\right] \le e^{(\gamma^2+\gamma/2)NT\|b^N\|_\infty^2}.$$

Since $\sup_{N\in\mathbb{N}}\|b^N\|_\infty := M < \infty$, (3.4a) is satisfied.

Moreover, using Lemma B.1 we have for any $\gamma > 1$ and any $\mu \in \mathcal{P}(S)$ with $R(\mu\|\mathbb{W}) < \infty$, that

$$\gamma\mathcal{E}_b(\mu) \le R(\mu\|\mathbb{W}) + \log\mathbb{E}\left[e^{\gamma|V_b|(W,\mu)}\right].$$

In a similar fashion as before, we can estimate the second term on the right-hand side uniformly in μ to obtain

$$R(\mu\|\mathbb{W}) + \gamma\mathcal{E}_b(\mu) \ge -a,$$

for some constant $a_0 > 0$ independent of $\mu \in \mathcal{P}(S)$, which gives (3.4b).

For (3.5a), we will consider the part of $\mathcal{E}_{b^N}^N - \mathcal{E}_b^N$ determined by V^1 and $V^{N,2}$ separately. First, note that $V_{b^N}^{N,2} - V_b^{N,2} = V_{b^N-b}^{N,2}$ and similar to the argument of (7.10) we have for all $\beta \in \mathbb{R}$

$$\mathbb{E}\left[e^{\beta N\langle V_{b^N-b}^{N,2}(\cdot,z_{\boldsymbol{W}}^N),z_{\boldsymbol{W}}^N\rangle}\right] \le \mathbb{E}\left[e^{2\beta^2\sum_{i=1}^N\int_0^T |b_t^N-b_t|^2(W_t^i,z_{\boldsymbol{W},t}^N)dt}\right]^{1/2} \le e^{c_N N\beta^2}.$$

Secondly, since

$$\beta \left| |b^N|^2 - |b|^2 \right| \leq \beta(|b^N| + |b|)|b^N - b| \leq \tfrac{1}{2}(M + \|b\|_\infty)^2 + \tfrac{\beta^2}{2}|b^N - b|^2,$$

we derive

$$\begin{aligned}
\mathbb{E}\left[e^{\beta N\left(\langle V^1_{b^N}(\cdot, z^N_{\boldsymbol{W}}), z^N_{\boldsymbol{W}}\rangle - \langle V^1_b(\cdot, z^N_{\boldsymbol{W}}), z^N_{\boldsymbol{W}}\rangle\right)}\right] &= \mathbb{E}\left[e^{\beta \sum_{i=1}^N \int_0^T \left(|b^N_t|^2 - |b_t|^2\right)(W^i_t, z^N_{\boldsymbol{W},t}) dt}\right] \\
&\leq \mathbb{E}\left[e^{\frac{TN}{2}(M + \|b\|_\infty) + \frac{\beta^2}{2}\sum_{i=1}^N \int_0^T \left(|b^N_t - b_t|^2\right)(W^i_t, z^N_{\boldsymbol{W},t}) dt}\right] \\
&\leq e^{\frac{TN}{2}(M + \|b\|_\infty) + \frac{c_N N \beta^2}{2}}.
\end{aligned}$$

Finally, via Cauchy-Schwarz, we conclude that for all $\beta \in \mathbb{R}$

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta\left(\mathcal{E}^N_{b^N}(z^N_{\boldsymbol{W}}) - \mathcal{E}^N_b(z^N_{\boldsymbol{W}})\right)}\right] \leq \tfrac{T}{2}(M + \|b\|_\infty).$$

□

7.4 Representation of the rate functional

In this section, we give another representation of the rate functional $\mathcal{E}_b(\mu) + I(\mu)$, in the form of the relative entropy with respect to a process depending on μ itself.

Lemma 7.6 (Relative entropy representation)**.** *Assume that*

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |b_t(W_t, \mu_t)|^2 dt}\right] < \infty, \quad \forall \mu \textit{ with } R(\mu \| \mathbb{W}) < \infty. \tag{7.11}$$

Then we have the following representation formula:

$$R(\mu\|\mathbb{W}) + \mathcal{E}_b(\mu) = \begin{cases} R(\mu\|\mathbb{W}^\mu), & \textit{if } R(\mu\|\mathbb{W}) < \infty, \\ +\infty, & \textit{otherwise}, \end{cases} \tag{7.12}$$

where $\mathbb{W}^\mu$ is the law of the process X^μ satisfying the SDE

$$dX^\mu_t = b_t(X^\mu_t, \mu_t)\, dt + dW_t$$

with initial law ρ_0 (the law $\mathbb{W}^\mu$ exists and is uniquely determined by Girsanov's theorem D.1).

*Moreover, when b is in the class $\mathcal{F}$*Lip*, the restriction $R(\mu\|\mathbb{W}) < \infty$ in* (7.12) *may be dropped.*

Proof. When $R(\mu\|\mathbb{W}) = \infty$, the representation formula holds trivially (recall that $\mathcal{E}_b(\mu)$ is finite for every μ). Now fix μ with $R(\mu\|\mathbb{W}) < \infty$. By the condition (7.11), we can apply Girsanov's theorem, which gives the formula

$$\begin{aligned} \frac{d\mathbb{W}^\mu}{d\mathbb{W}}(W) &= \exp\left(-\frac{1}{2}\int_0^T |b_t(W_t,\mu_t)|^2\,dt + \int_0^T b_t(W_t,\mu_t)\cdot dW_t\right) \\ &= \exp\left(-\frac{1}{2}V_b^1(W,\mu) + V_b^2(W,\mu)\right). \end{aligned} \tag{7.13}$$

In particular, $\mathbb{W}$ and $\mathbb{W}^\mu$ are equivalent and so μ is absolutely continuous also with respect to $\mathbb{W}^\mu$. Hence we can compute the relative entropy

$$\begin{aligned} R(\mu\|\mathbb{W}^\mu) &= \int \log\frac{d\mu}{d\mathbb{W}^\mu}d\mu = \int \log\frac{d\mu}{d\mathbb{W}}d\mu - \int \log\frac{d\mathbb{W}^\mu}{d\mathbb{W}}d\mu \\ &= R(\mu\|\mathbb{W}) - \int\left(-\frac{1}{2}V_b^1(W,\mu) + V_b^2(W,\mu)\right)d\mu(W) \\ &= R(\mu\|\mathbb{W}) + \mathcal{E}_b(\mu), \end{aligned}$$

which is the desired representation formula.

For b in $\mathcal{F}$Lip we have to prove that for every μ,

$$R(\mu\|\mathbb{W}^\mu) < \infty \iff R(\mu\|\mathbb{W}) < \infty. \tag{7.14}$$

Note that $\mathbb{W}^\mu$ is well-defined for every μ when b is in $\mathcal{F}$Lip. Now, fix any μ, and note that in particular b is bounded and hence Girsanov's formula (7.13) still holds. Moreover, for every $\beta > 0$ and $i = 1, 2$, by boundedness of b and Lemma D.2 we have,

$$\mathbb{E}\left[e^{\beta V_b^i(W,\mu)}\right] < \infty.$$

Applying Corollary B.3 to the measures $\mathbb{W}$ and $\mathbb{W}^\mu$ we easily deduce (7.14). $\square$

Chapter 8

Large deviations for singular drift

Now we give the main result of Part I.B, which states the LDP for the system (7.1). The assumptions may seem involved at first glance, but their meaning is not difficult: we have an LDP for the system (7.1) as soon as we can approximate in a suitable way the drift b by regular drifts b_λ, along the Brownian empirical measure $z^N_{\boldsymbol{W}}$ and the couple of Brownian path W and measure μ with finite relative entropy.

Theorem 8.1. *Assume the condition* (7.9) *on the initial law* ρ_0. *Suppose there exists a sequence of* $\mathcal{F}$Lip*-inducing drifts* $(b^N_\lambda, b_\lambda)_{\lambda>0}$ *and a sequence* $(g_\lambda)_{\lambda>0}$ *of Borel functions* $g_\lambda : [0,T]\times(\mathbb{R}^d)^k \to [0,+\infty)$, $k\in\mathbb{N}$, *such that the following conditions hold:*

(i) for every $\lambda>0$, *for* P^N*-almost every* $z^N\in\mathcal{P}(S)$,

$$\int_0^T \left\langle |b^N_t - b^N_{\lambda,t}|^2(\cdot, z^N_t), z^N_t \right\rangle dt \le \int_0^T \int_{((\mathbb{R}^d)^k)'} g_\lambda(t,x_1,\dots,x_k)\, d(z^N_t)^{\otimes k}\, dt; \tag{8.1}$$

(ii) for every $\lambda>0$ *and every* $\mu\in\mathcal{P}(S)$ *with* $R(\mu\|\mathbb{W})<\infty$ *and* $\mathbb{W}$*-almost every* W,

$$\int_0^T |b_t - b_{\lambda,t}|^2(W_t,\mu_t)\, dt \le \int_0^T \int_{(\mathbb{R}^d)^k} g_\lambda(t,W_t,y)\, d\mu_t^{\otimes k-1}(y)\, dt. \tag{8.2}$$

Suppose also that $(g_\lambda)_{\lambda>0}$ *satisfies for some* $K>0$,

$$\limsup_{\lambda\to 0} \mathbb{E}\left[e^{\beta\int_0^T g_\lambda(t,W^1_t,\dots,W^k_t)\,dt}\right] \le K, \qquad \forall\beta\in\mathbb{R}, \tag{8.3}$$

where $W^1, \dots, W^k$ are independent Brownian motions with common initial law ρ_0.

Then the family $\{Q^N_{b^N}\}$ of laws of $z^N_{\boldsymbol{X}}$ (for $\boldsymbol{X} = (X^{N,1}, \dots, X^{N,N})$ satisfying (7.1)*) has an LDP with rate function*

$$\mathcal{F}(\mu) = \begin{cases} R(\mu\|\mathbb{W}^\mu) & \text{if } R(\mu\|\mathbb{W}) < \infty, \\ +\infty, & \text{otherwise}, \end{cases} \tag{8.4}$$

where $\mathbb{W}^\mu$ is the law of the process X^μ_t satisfying the SDE

$$dX^\mu_t = b_t(X^\mu_t, \mu)\,dt + dW_t. \tag{8.5}$$

Remark 8.2. Note that the zeros of the rate function $\mathcal{F}$ are exactly the solution to the McKean–Vlasov SDE

$$\begin{cases} dX_t = b_t(X_t, \mathrm{Law}(X_t))\,dt + dW_t, \\ X_0 \text{ with law } \rho_0, \end{cases}$$

with $R(\mathrm{Law}(X_t)\|\mathbb{W}) < \infty$. In particular, since $\mathcal{F}$ has at least one zero, there exists at least one solution to the McKean–Vlasov SDE with finite relative entropy (with respect to $\mathbb{W}$). □

Note that under the assumptions of the above theorem, $Q^N_{b^N}$ is well-defined by Theorem 7.1.

Moreover, since (8.3) is quite general, an application of Khasminskii's lemma provides us with the following sufficient condition:

Lemma 8.3. *Let $(g_\lambda)_{\lambda>0}$ be a sequence of Borel functions $g_\lambda : [0,T] \times (\mathbb{R}^d)^k \to [0, +\infty)$, $k \in \mathbb{N}$, that satisfies*

$$\limsup_{\lambda\to 0} \sup_{x_1,\dots,x_k \in \mathbb{R}^d} \mathbb{E}^{x_1,\dots,x_k}\left[\int_0^T g_\lambda(t, W^{1,x_1}_t, \dots, W^{k,x_k}_t)\,dt\right] = 0, \tag{8.6}$$

where the expectation is over k independent Brownian motions $W^{1,x_1}, \dots, W^{k,x_k}$ starting at points $x_1, \dots, x_k \in \mathbb{R}^d$ respectively. Then (8.3) *is satisfied.*

Proof. By Khasminskii's lemma (cf. Lemma D.3), (8.6) implies that, for every $\beta \geq 0$ and $0 < \alpha < 1$, and for all $\lambda > 0$ sufficiently small

$$\sup_{(x_1,\dots x_k)\in\mathbb{R}^{kd}} \mathbb{E}\left[e^{\beta \int_0^T g_\lambda(t, W^{1,x_1}_t, \dots W^{k,x_k}_t)\,dt}\right] \leq \frac{1}{1-\alpha},$$

and so, averaging $(x_1, \dots x_k)$ over $\rho_0^{\otimes k}$ and using Jensen's inequality yields

$$\mathbb{E}\left[e^{\beta \int_0^T g_\lambda(t,W_t^1,\dots W_t^k)\,dt}\right] \leq \frac{1}{1-\alpha}, \qquad \text{for all } \lambda > 0 \text{ sufficiently small.}$$

Therefore we have, for every $\beta \geq 0$, using that α is arbitrary,

$$\limsup_{\lambda \to 0} \mathbb{E}\left[e^{\beta \int_0^T g_\lambda(t,W_t^1,\dots,W_t^k)\,dt}\right] = 1,$$

which gives (8.3) with $K = 1$. □

8.1 Proof of main result

Here we state the proof of the large deviation result Theorem 8.1.

Proof of Theorem 8.1. We would like to apply Theorem 5.1 to $\mathcal{E}^N_{b^N}$, $\mathcal{E}_b$, $\mathcal{E}^N_{b^N_\lambda}$, $\mathcal{E}_{b_\lambda}$, with

$$G^N_\lambda(x^1, \dots, x^k) = G_\lambda(x^1, \dots, x^k) = \int_0^T g_\lambda(t, x^1_t, \dots x^k_t)\,dt, \qquad x^i \in S,\ i = 1, \dots, N.$$

We claim that the conditions (3.4), (5.1a), (5.1b) and (5.2a), (5.2b) hold. Then Theorem 5.1 and Lemma 7.4 give an LDP for $Q^N_{b^N}$ with rate function $R(\mu \| \mathbb{W}) + \mathcal{E}_b(\mu)$.

We now prove the claims on the above conditions and the form (8.4) for the rate function. In particular, we prove that (8.1) and (8.2) imply (5.1a) and (5.1b) respectively. Finally, note that (8.3) directly implies (5.2a) and (5.2b), and (3.4) follows as in the proof of Lemma 7.4.

Preliminary uniform bounds: We call, for $\mu \in \mathcal{P}(S)$,

$$\begin{aligned}\overline{K}_{\lambda,\beta} &:= \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta N \langle \int_0^T |b^N_t - b^N_{\lambda,t}(\cdot, z^N_{\mathbf{W},t})|^2 dt, z^N \rangle}\right],\\ \overline{H}_{\lambda,\beta}(\mu) &:= \log \mathbb{E}\left[e^{\beta \int_0^T |b - b_\lambda(W,\mu)|^2 dt}\right].\end{aligned}$$

We will show that there exists $\lambda_0 > 0$ and $\beta_0 \gg 1$ arbitrarily large (more precisely, for every $\beta_0 > 0$ large, there exists $\lambda_0 > 0$), such that for all $\beta \leq \beta_0$:

$$\begin{aligned} &\overline{K}_{\lambda_0,\beta} \leq \overline{K}_{\lambda_0,\beta_0} < \infty,\\ \sup_{\mu, R(\mu\|\mathbb{W})<\infty} \overline{H}_{\lambda_0,\beta}(\mu) - (k-1)R(\mu\|\mathbb{W}) \leq &\sup_{\mu, R(\mu\|\mathbb{W})<\infty} \overline{H}_{\lambda_0,\beta_0}(\mu) - (k-1)R(\mu\|\mathbb{W}) < \infty.\end{aligned} \tag{8.7}$$

We start with the proof for $\overline{K}_{\lambda_0,\beta_0}$. Applying assumption (8.1), we get

$$\begin{aligned}\overline{K}_{\lambda_0,\beta_0} &:= \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta_0 N \langle \int_0^T |b_t^N - b_{\lambda_0,t}^N(\cdot, z_{W,t}^N)|^2 dt, z_W^N\rangle}\right] \\ &\le \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\beta_0 N \int_{(S^k)'} G_{\lambda_0}(x^1,\dots,x^k)\, d(z_W^N)^{\otimes k}(x^1,\dots x^k)}\right].\end{aligned}$$

Since (5.2a) holds, we can apply Remark 5.3: for any β_0, the above right-hand side is finite for $\lambda_0 > 0$ sufficiently small. For $\overline{H}_{\lambda_0,\beta_0}(\mu)$, we have

$$\begin{aligned}\overline{H}_{\lambda_0,\beta_0}(\mu) - (k-1)R(\mu\|\mathbb{W}) &:= \log \mathbb{E}\left[e^{\beta_0 \int_0^T |b - b_{\lambda_0}(W,\mu)|^2 dt}\right] - (k-1)R(\mu\|\mathbb{W}) \\ &\le \log \mathbb{E}\left[e^{\beta_0 \int_{S^{k-1}} G_{\lambda_0}(W,y) d\mu^{\otimes(k-1)}(y)}\right] - (k-1)R(\mu\|\mathbb{W}).\end{aligned}$$

Since (5.2b) holds, we can apply Remark 5.3: for any β_0, the above right-hand side is finite and bounded uniformly over μ for $\lambda_0 > 0$ sufficiently small.

As a consequence of (8.7), we have (for N large at least)

$$\mathbb{E}\left[e^{\frac{1}{2}\int_S V^1_{b^N}(x, z_W^N)\, dz_W^N(x)}\right] < \infty, \tag{8.8a}$$

$$\mathbb{E}\left[e^{\frac{1}{2}\int_S V^1_b(W,\mu)}\right] < \infty \qquad \forall \mu \text{ with } R(\mu\|\mathbb{W}) < \infty. \tag{8.8b}$$

In particular, $V^1_{b^N}(\cdot, z^N)$ and $V^{2,N}_{b^N}(\cdot, z^N)$ are in $L^1(z^N)$ for $\mathbb{P}^{\mathbb{N}}$-a.e. z^N and also, for every μ with $R(\mu\|\mathbb{W}) < \infty$, $V^1_b(\cdot, \mu)$ and $V^2_b(\cdot, \mu)$ are in $L^1(\mu)$, see Section 7.2 for details. Moreover, by Theorem 7.1, at least for N large, the system 7.1 admits a weak solution, whose unique (under the additional constraint) law has density given by (7.4). Finally, the inequality (8.8b) holds replacing $1/2$ with any $\beta > 0$ in the exponential, in particular, $\mathcal{E}_b$ is Borel by Lemma 7.2.

Verification of (5.1a) ***and*** (5.1b)***:*** We fix λ_0 and β_0 such that (8.7). In view of (5.1a), we show some easy uniform bounds. Using the inequality $|b^N|^2 \le 2|b^N - b^N_{\lambda_0}|^2 + 2|b^N_{\lambda_0}|^2$ and applying Hölder inequality, we get, for every $\ell \ge 0$,

$$\begin{aligned}\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\ell N \langle \int_0^T \langle |b_t^N(\cdot, z_{W,t}^N)|^2 dt, z_W^N\rangle}\right] &\le \limsup_{N\to\infty} \frac{1}{2N} \log \mathbb{E}\left[e^{4\ell N \langle \int_0^T |b_t^N - b_{\lambda_0,t}^N(\cdot, z_{W,t}^N)|^2 dt, z_W^N\rangle}\right] \\ &\quad + \limsup_{N\to\infty} \frac{1}{2N} \log \mathbb{E}\left[e^{4\ell N \langle \int_0^T |b_{\lambda_0,t}^N(\cdot, z_{W,t}^N)|^2 dt, z_W^N\rangle}\right] \\ &=: \frac{1}{2}\overline{K}_{\lambda_0,4\ell} + \frac{1}{2}\overline{K}'_{\lambda_0,4\ell} < \infty,\end{aligned} \tag{8.9}$$

where $\overline{K}'_{\lambda_0,4\ell}$ is finite because $\sup_{N\in\mathbb{N}} \|b^N_{\lambda_0}\|_\infty < \infty$. Using now the inequality $|b^N_\lambda|^2 \leq 2|b^N - b^N_\lambda|^2 + 2|b^N|^2$ and proceeding similarly, we also get, for every $\ell \geq 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{\ell N\langle \int_0^T |b^N_{\lambda,t}(\cdot,z^N_{W,t})|^2 dt, z^N_W\rangle}\right] \leq \frac{1}{2}\big(\overline{K}_{\lambda,4\ell} + \overline{K}_{\lambda_0,4\ell} + \overline{K}'_{\lambda_0,4\ell}\big). \tag{8.10}$$

To show (5.1a), we write, for $\beta \in \mathbb{R}$,

$$\mathbb{E}\left[e^{-\beta N(\mathcal{E}^N_{b^N} - \mathcal{E}^N_{b^N_\lambda})(z^N_W)}\right] \leq \mathbb{E}\left[e^{-\beta N\langle (V^1_{b^N} - V^1_{b^N_\lambda})(\cdot,z^N_W), z^N_W\rangle}\right]^{1/2} \mathbb{E}\left[e^{2\beta N\langle (V^2_{b^N} - V^2_{b^N_\lambda})(\cdot,z^N_W), z^N_W\rangle}\right]^{1/2} \tag{8.11}$$

and we control the differences $V^1_{b^N} - V^1_{b^N_\lambda}$ and $V^2_{b^N} - V^2_{b^N_\lambda}$ separately. Using the inequality

$$\beta(|b^N|^2 - |b^N_\lambda|^2) \leq |b^N|^2 + |b^N_\lambda|^2 + \frac{\beta^2}{2}|b^N - b^N_\lambda|^2,$$

and applying Hölder's inequality and the bounds (8.9) and (8.10) (with $\ell = 4$), we get

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}[e^{\beta N\langle (V^1_{b^N} - V^1_{b^N_\lambda})(\cdot,z^N_W), z^N_W\rangle}] \\
&\leq \frac{1}{4}\overline{K}_{\lambda,16} + \frac{1}{2}\big(\overline{K}_{\lambda_0,16} + \overline{K}'_{\lambda_0,16}\big) + \limsup_{N\to\infty} \frac{1}{2N} \log \mathbb{E}[e^{\beta^2 N\langle \int_0^T |(b^N_t - b^N_{\lambda,t})(\cdot,z^N_{W,t})|^2 dt, z^N_W\rangle}] \\
&= \frac{1}{4}\overline{K}_{\lambda,16} + \frac{1}{2}\big(\overline{K}_{\lambda_0,16} + \overline{K}'_{\lambda_0,16}\big) + \frac{1}{2}\overline{K}_{\lambda,\beta^2}.
\end{aligned} \tag{8.12}$$

For $V^2_{b^N} - V^2_{b^N_\lambda}$, we use Lemma D.2 to obtain

$$\mathbb{E}\left[e^{2\beta N\langle V^2_{b^N}(\cdot,z^N_W) - V^2_{b^N_\lambda}(\cdot,z^N_W), z^N_W\rangle}\right] \leq \mathbb{E}\left[e^{4\beta^2 \sum_{i=1}^N \int_0^T |b^N_t - b^N_{\lambda,t}|^2(W^i_t, z^N_{W,t})\, dt}\right]^{\frac{1}{2}},$$

and therefore

$$\limsup_{N\to\infty} \frac{1}{N} \log E^N\left[e^{4\beta N\langle V^2_{b^N}(\cdot,z^N_W) - V^2_{b^N_\lambda}(\cdot,z^N_W), z^N_W\rangle}\right] \leq \frac{1}{2}\overline{K}_{\lambda,4\beta}. \tag{8.13}$$

Putting together the inequalities (8.11), (8.12) and (8.13), we get, for some constant $c > 0$ (independent of β) and some $c_\beta > 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-\beta N(\mathcal{E}^N_{b^N} - \mathcal{E}^N_{b^N_\lambda})(z^N_W)}\right] \leq \bar{c} + c\overline{K}_{\lambda,c_\beta}, \tag{8.14}$$

with $\bar{c} = (\overline{K}_{\lambda_0,16} + \overline{K}'_{\lambda_0,16})/2 \geq 0$. The assumption (8.1) gives, for some new $c_\beta > 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[e^{-\beta N(\mathcal{E}^N_{b^N} - \mathcal{E}^N_{b^N_\lambda})(z^N_{\boldsymbol{W}})}\right]$$
$$\leq \bar{c} + c \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\exp\left(c_\beta N \int_0^T \int_{((\mathbb{R}^d)^k)'} g_\lambda(t, x_1, \dots x_k)\, d(z^N_{\boldsymbol{W},t})^{\otimes k}\, dt\right)\right]$$
$$\leq \bar{c} + c \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\exp\left(c_\beta N \int_{(S^k)'} \int_0^T g_\lambda(t, x^1_t, \dots x^k_t)\, dt\, d(z^N_{\boldsymbol{W}})^{\otimes k}\right)\right]$$

and (5.1a) follows.

As for (5.1b), we use Lemma B.1 to obtain, for every μ with $R(\mu\|\mathbb{W}) < \infty$,

$$\beta|\mathcal{E}_b(\mu) - \mathcal{E}_{b_\lambda}(\mu)| \leq R(\mu\|\mathbb{W}) + \log \mathbb{E}\left[e^{\beta|V_b(W,\mu) - V_{b_\lambda}(W,\mu)|}\right].$$

Using the same arguments as before, we get, for some $c' > 0$ (independent of β) and some $c_\beta > 0$,

$$\beta|\mathcal{E}_b(\mu) - \mathcal{E}_{b_\lambda}(\mu)| \leq \frac{k+1}{2} R(\mu\|\mu_0) + \bar{c} + c'\overline{H}_{\lambda,c_\beta}, \tag{8.15}$$

where $\bar{c} > 0$ is such that $\bar{c} \geq (\overline{H}_{\lambda_0,16}(\mu) - (k-1)R(\mu\|\mu_0) + \overline{H}'_{\lambda_0,16})/2$ for every μ with $R(\mu\|\mathbb{W}) < \infty$, and $\overline{H}'_{\lambda_0,\beta} = \log \mathbb{E}[e^{\beta \int_0^T |b_{\lambda_0}(W,\mu)|^2 dt}]$. The assumption (8.2) gives, for some new c_β,

$$\beta|\mathcal{E}_{b^N}(\mu) - \mathcal{E}_{b^N_\lambda}(\mu)| \leq \frac{k+1}{2} R(\mu\|\mu_0) + \bar{c} + c' \log \mathbb{E}\left[e^{c_\beta \int_{S^k} \int_0^T g_\lambda(t, W_t, x^2_t \dots x^k_t)\, dt\, d\mu^{\otimes(k-1)}}\right]$$

and (5.1b) follows.

Proof of (8.4)***:*** By (8.8b), b satisfies the assumption of Lemma 7.6, which then implies the representation formula (8.4) for the rate function. The proof is complete. □

The above proof shows that we can relax some of the assumptions, as we show below.

Proposition 8.4. *The results of Theorem 8.1 (namely the LDP for $Q^N_{b^N}$ with rate function $\mathcal{F}$) remain valid if any of the following statements hold:*

(a) *Instead of* $\mathcal{F}$Lip-*inducing drifts* $(b^N_\lambda, b_\lambda)_{\lambda>0}$, *we assume that for every* $\lambda > 0$ *the family* $Q^N_{b^N_\lambda}$ *has an LDP with rate function* $\mathcal{F}_{b_\lambda}$ *(defined similarly to* $\mathcal{F}$ *via* (8.4), (8.5)*), and for every* $\beta \in \mathbb{R}$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\exp\left(N\beta \int_0^T \left\langle |b^N_{\lambda,t}|^2(\cdot, z^N_{\boldsymbol{W},t}), z^N_{\boldsymbol{W},t}\right\rangle\right) dt\right] < \infty,$$
$$\sup_{\mu, R(\mu\|\mathbb{W})<\infty} \log \mathbb{E}\left[\exp\left(\beta \int_0^T |b_{\lambda,t}|^2(W_t, \mu_t)\, dt\right)\right] - R(\mu\|\mathbb{W}) < \infty. \tag{8.16}$$

(b) *Instead of the existence of* g_λ *and* (8.1), (8.2), *we assume there exists a constant* $K \in \mathbb{R}$ *such that for every* $\beta \in \mathbb{R}$,

$$\limsup_{\lambda\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\exp\left(N\beta \int_0^T \left\langle |b^N_t - b^N_{\lambda,t}|^2(\cdot, z^N_{\boldsymbol{W},t}), z^N_{\boldsymbol{W},t}\right\rangle\right) dt\right] \le K,$$
$$\limsup_{\lambda\to 0} \sup_{\mu, R(\mu\|\mathbb{W})<\infty} \log \mathbb{E}\left[\exp\left(\beta \int_0^T |b_t - b_{\lambda,t}|^2(W_t, \mu_t)\, dt\right)\right] - R(\mu\|\mathbb{W}) \le K. \tag{8.17}$$

Proof. The inequalities (8.16) imply via Theorem 7.1 and Lemmas D.2 and 7.6 both the representations of $Q^N_{b^N_\lambda}$ and $\mathcal{F}_{b_\lambda}$ in terms of $\mathcal{E}^N_{b^N_\lambda}, \mathcal{E}_{b_\lambda}$, and the estimates (3.4). Moreover, (8.17) combined with (8.14) and (8.15) imply conditions (3.5a) and (3.5b). Hence we can use directly Theorem 3.8 to deduce the LDP (the representation formula (8.4) follows again from Lemma 7.6). □

8.2 Applications to concrete examples

In this section, we consider common forms of drifts b^N that appear in applications.

8.2.1 Example: 2-point interaction

We start with the example of 2-point, translation-invariant interaction; while this is a particular case of the k-point interaction, we discuss this case separately, to highlight the essential ingredients of the result. In this example, we consider a simple class of drifts, that commonly appear in various fields of application, namely

$$b^N_t(x_i, z^N_{\boldsymbol{x}}) = \frac{1}{N} \sum_{j\neq i} \varphi_t(x_i - x_j), \qquad z^N_{\boldsymbol{x}} = \frac{1}{N} \sum_{j=1}^N \delta_{x_j} \tag{8.18}$$

for some Borel map $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$.

Remark 8.5. Strictly speaking, (8.18) is not a good definition, because the right-hand side depends on i and not just on x_i and $z^N_{\boldsymbol{x}}$. However, we can give a rigorous definition with a harmless change of (8.18). Precisely, we can define

$$b^N_t(x,\mu) = \int_{\mathbb{R}^d\setminus\{x\}} \varphi_t(x-y)\, d\mu_t(y). \tag{8.19}$$

Indeed, note that (recall $\boldsymbol{W} = (W^1,\dots,W^N)$ is a N-tuple of d-dimensional independent Brownian motions with law $\tilde{\boldsymbol{P}}^N = \mathbb{W}^{\otimes N}$)

$$\frac{1}{N}\sum_{j\neq i} \varphi_t(W^i_t - W^j_t) = \int_{\mathbb{R}^d\setminus\{x\}} \varphi_t(x-y)\, dz^N_{\boldsymbol{W},t}(y), \tag{8.20}$$

for a.e. $t\in[0,T]$ and $\tilde{P}^N$-a.e. $\boldsymbol{W}$, because the set $\{t\in[0,T] : W^i_t = W^j_t, j\neq i\}$ has Lebesgue measure zero for $\tilde{P}^N$-a.e. $\boldsymbol{W} = (W^1,\dots,W^N)$. Therefore, if Girsanov's theorem can be applied to the particle system (7.1), as it is the case in the next result, then (8.20) holds also replacing $\boldsymbol{W}$ and $\tilde{\boldsymbol{P}}^N$ respectively with $\boldsymbol{X}^N = (X^{N,1},\dots X^{N,N})$, the solution to the particle system (7.1) with drift (8.19), and its law $\tilde{Q}^N$. Hence such solution $\boldsymbol{X}^N$ solves also the original particle system with drift (8.18). □

We introduce also the drift of the associated McKean–Vlasov SDE, namely

$$b_t(x,\mu) = \int_{\mathbb{R}^d} \varphi_t(x-y)\, d\mu_t(y).$$

Proposition 8.6. *Assume the condition* (7.9) *on the initial law* ρ_0. *Suppose that*

$$\mathbb{E}\left[e^{\beta\int_0^T |\varphi_t|^2(W^1_t - W^2_t)\,dt}\right] < \infty, \qquad \forall\beta\in\mathbb{R}, \tag{8.21}$$

where W^1, W^2 *are independent Brownian motions with common initial law* ρ_0.

Then the family $\{Q^N_b\}$ *of laws of the empirical processes associated to the interacting system* (7.1) *with drifts of the form* (8.18) *satisfies an LDP with rate function* $\mathcal{F}$ *given in* (8.4).

In particular, (8.21) *holds whenever* φ *is in* $L^q_t(L^p_{x-y})$, *for* p, q *satisfying*

$$2\le p,q\le\infty, \quad \frac{d}{p}+\frac{2}{q} < 1. \tag{8.22}$$

Proposition 8.6 is a special case of Proposition 8.9 with $k=2$, $p_1=p, p_2=+\infty$, and hence we will postpone the proof. Here, $\varphi\in L^q_t(L^p_{x-y})$ is shorthand for $\varphi(x,y)=\varphi(x-y)$ and $\varphi\in L^q\big((0,T),L^p(\mathbb{R}^d)\big)$.

Remark 8.7. The condition $\varphi \in L^q_t(L^p_{x-y})$ with p, q satisfying (8.22) is well-known in the literature on the so-called regularization by noise phenomenon (where an ill-posed ODE or PDE gains well-posedness by addition of a suitable noise). Indeed, a d-dimensional SDE, with additive noise and drift in $L^q_t(L^p_x)$, with p, q satisfying (8.22), has the strong existence and uniqueness property, see for example [KR05, FF11] among many others; on the contrary, if p, q do not satisfy (8.22) not even with equality, there exist counterexamples to well-posedness for SDEs, even in the weak sense, see e.g. [BFGM19, Section 7].

For this reason, the exponents of (8.22) are likely optimal for irregular drifts φ in our example: we expect that, for $L^q_t(L^p_{x-y})$ drifts without condition (8.22), even the 2-particle system is not well-posed. However, there are likely drifts that satisfy (8.21) but are not in a $L^q_t(L^p_x)$ class with (8.22).

□

Remark 8.8. A relevant example of function φ verifying condition (8.22) is

$$\varphi(t, z) = |z|^\alpha g\left(\frac{z}{|z|}\right) 1_{|z|\leq R} + h(z) 1_{|z|>R},$$

with $g : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ and $h : \mathbb{R}^d \to \mathbb{R}^d$ both Borel bounded, $R > 0$, and exponent α satisfying

$$\alpha > -1 \text{ for } d \geq 2, \quad \text{and} \quad \alpha > -1/2 \text{ for } d = 1.$$

It is unclear whether the restriction $\alpha > -1/2$ in $d = 1$, which is due to the assumption $p \geq 2$, is really optimal or is a drawback of the method, nonetheless the assumption $p, q \geq 2$ appears in several works dealing with singular drifts, like [KR05, FF11].

When $d \geq 2$, the above condition includes the case

$$\varphi(x) = -\nabla\Phi(x), \qquad \Phi(t, x) = |x|^\alpha, \qquad \alpha > 0,$$

in some neighborhood around $x = 0$, and with φ bounded outside of this neighborhood. However, $\Phi(x) = \log|x|$ does not fall in this class. Moreover, as shown in [HH16], the exponential moment estimate of (8.21) actually blows up when β is large enough.

The latter is also related to the fact that for $d \geq 2$ and a logarithmic potential $\Phi(x) = \log|x|$, the law of SDE (7.1) might no longer be absolutely continuous with respect to the law of non-interacting Brownian motions—it is possible in the case of a sufficiently strong attractive force that the particles hit each other (see [FJ17]). Interestingly enough, as is done in [FJ17], it can still be shown that the particle system converges to the corresponding McKean–Vlasov equation. However, whether in this case a large deviation principle still exists is not known. □

8.2.2 Example: k-point interaction

We can also treat the case of a k-interaction drift, namely

$$b_t^N(x_i, z_{\boldsymbol{x}}^N) = \frac{1}{N^{k-1}} \sum_{j_1,\dots,j_{k-1} \neq i \text{ all distinct}} \varphi_t(x_i, x_{j_1}, \dots x_{j_{k-1}}) \tag{8.23}$$

for some Borel function $\varphi : [0,T] \times \mathbb{R}^{kd} \to \mathbb{R}^d$. As in the previous example, the rigorous definition of b^N can be given as in Remark 8.5. Similarly, we take the drift of the associated McKean–Vlasov SDE

$$b_t(x, \mu) = \int_{\mathbb{R}^{(k-1)d}} \varphi_t(x, y_1, \dots y_{k-1}) \, d\mu_t^{\otimes(k-1)}(y_1, \dots y_{k-1}). \tag{8.24}$$

Proposition 8.9. *Assume the condition* (7.9) *on the initial law* ρ_0. *Suppose that*

$$\mathbb{E}\left[e^{\beta \int_0^T |\varphi_t|^2(W_t^1,\dots,W_t^k)\, dt}\right] < \infty, \qquad \forall \beta \in \mathbb{R} \tag{8.25}$$

where $W^1, \dots, W^k$ *are independent Brownian motions with common initial law* ρ_0.

Then the family $\{Q_b^N\}$ *of laws of the empirical processes associated to the interacting system* (7.1) *with drifts of the form* (8.23) *satisfies an LDP with rate function* $\mathcal{F}$ *given in* (8.4).

In particular, (8.25) *holds whenever*

$$\varphi \in L_t^q(L_{x_1}^{p_1}(\dots(L_{x_k}^{p_k})\dots)), \tag{8.26}$$

with $p_1, \dots p_k, q$ *satisfying*

$$p_1, \dots p_k, q \in [2, \infty], \quad \frac{d}{p_1} + \dots + \frac{d}{p_k} + \frac{2}{q} < 1. \tag{8.27}$$

Remark 8.10. Similarly to Lemma D.5, (8.26) can be replaced by

$$\varphi \in L_t^q(L_{x_{\sigma(1)}}^{p_1}(\dots(L_{x_{\sigma(k)}}^{p_k})\dots))$$

for some permutation σ of $\{1, \dots k\}$. □

Remark 8.11. The space of φ satisfying condition (8.25) is a vector space (as easily checked). Hence, condition (8.25) also holds in the more general case of $\varphi = \sum_{j=1}^m \varphi_j$ for some m, where, for any $j = 1, \dots m$,

$$\varphi_j \in L_t^{q^{(j)}}(L_{x_1}^{p_1^{(j)}}(\dots(L_{x_k}^{p_k^{(j)}})\dots))$$

with $p_1^{(j)}, \dots p_k^{(j)}, q^{(j)}$ satisfying (8.27). In particular, we can allow φ to be a sum of a bounded function and a function satisfying (8.26) and (8.27) as in Theorem 6.1. □

Proof of Proposition 8.9. We will use three different approximations:

1. First, we consider b with φ Lipschitz bounded and show that (b^N, b) are $\mathcal{F}$Lip-inducing, by bounding away the self-interactions.

2. Then we consider the case b with φ Borel bounded. We apply Theorem 8.1 to get the desired LDP.

3. Finally, we extend this to φ satisfying (8.25) using truncation of φ and the approximation given in step (2).

The fact that (8.25) holds under conditions (8.26) and (8.27) follows from Lemma D.6.

(1) Assume that $\varphi : [0,T] \times \mathbb{R}^{kd} \to \mathbb{R}^d$ is Borel bounded and that the map $x \mapsto \varphi(t,x)$ is globally Lipschitz continuous for every $t \in [0,T]$, with Lipschitz constant $\mathrm{Lip}(\varphi)$ independent of t. We will show that (b^N, b) is $\mathcal{F}$Lip-inducing.

It is well known that $b \in \mathcal{F}$Lip: indeed, for any $x_1, x_2 \in \mathbb{R}^d$ and $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$, we have

$$\begin{aligned}
&|b_t^N(x_1,\mu_1) - b_t^N(x_2,\mu_2)| \\
&\quad \le \int_{\mathbb{R}^{(k-1)d}} \left|\varphi_t(x_1, y_1, \dots y_{k-1}) - \varphi_t(x_2, y_1', \dots y_{k-1}')\right| d\gamma(y_1,y_1') \cdot \ldots \cdot d\gamma(y_{k-1}, y_{k-1}') \\
&\quad \le \mathrm{Lip}(\varphi)\left(|x_1 - x_2| + \int_{\mathbb{R}^{(k-1)d}} \left(\sum_{i=1}^{k-1} |y_i - y_i'|\right) d\gamma(y_1,y_1') \cdot \ldots \cdot d\gamma(y_{k-1}, y_{k-1}')\right) \\
&\quad = \mathrm{Lip}(\varphi)\left(|x_1 - x_2| + (k-1)\int_{\mathbb{R}^{(k-1)d}} |y - y'|\, d\gamma(y,y')\right),
\end{aligned}$$

for any coupling γ between μ and ν; optimizing over all couplings then yields the required Lipschitz estimate involving the W_1-distance. Moreover, clearly $|b^N| \le |b|$ and hence $\sup_N \|b^N\|_\infty < \infty$. Finally, to show (7.8), note that, similarly to Lemma 4.3, we have for every N and P^N-almost every z^N

$$\begin{aligned}
&\int_0^T \left\langle |b_t - b_t^N|^2(\cdot, z_t^N), z_t^N \right\rangle dt \\
&\le \frac{2}{N}\int_0^T \sum_{i=1}^N \left(\frac{1}{N^{k-1}} \left| \left(\sum_{j_1,\dots,j_{k-1} \neq i \text{ all distinct}} - \sum_{j_1,\dots,j_{k-1}\neq i} \right) \varphi_t(W_t^i, W_t^{j_1}, \dots W_t^{j_{k-1}}) \right| \right)^2 dt \\
&\quad + \frac{2}{N}\int_0^T \sum_{i=1}^N \left(\frac{1}{N^{k-1}} \left| \left(\sum_{j_1,\dots,j_{k-1} \neq i} - \sum_{j_1,\dots,j_{k-1}} \right) \varphi_t(W_t^i, W_t^{j_1}, \dots W_t^{j_{k-1}}) \right| \right)^2 dt \\
&\le 2T\|\varphi\|_\infty^2 \frac{1}{N^{2(k-1)}} \left(\left|\frac{(k-1)(k-2)}{2} N^{k-2}\right|^2 + \left|(k-1)N^{k-2}\right|^2 \right) \le \frac{C}{N^2},
\end{aligned}$$

for some generic constant C. Hence (b^N, b) is $\mathcal{F}$Lip-inducing.

(2) Next, suppose that $\varphi : [0,T] \times \mathbb{R}^{kd} \to \mathbb{R}^d$ is Borel bounded. We can find a sequence of Lipschitz approximations $(\varphi_\lambda)_{\lambda>0}$, with φ_λ as in **(1)** for each $\lambda > 0$, such that $\varphi_\lambda \to \varphi$ Lebesgue-a.e. as $\lambda \to 0$ and $\|\varphi_\lambda\|_\infty \le \|\varphi\|_\infty$ (for example, we can take φ_λ as convolutions of φ with approximations of identity). We take b^N_λ, b_λ the corresponding drifts for the particle system and the McKean–Vlasov SDE (as respectively in (8.23), (8.24) with φ_λ in place of φ). In order to apply Theorem 8.1, with (b^N_λ, b_λ) as sequence of $\mathcal{F}$Lip-inducing drifts, we verify now the assumptions (8.1), (8.2) and (8.3). The estimates (8.1) and (8.2) are easily obtained with

$$g_\lambda(t, x_1, \dots, x_k) = \left|\varphi(t, x_1, \dots, x_k) - \varphi_\lambda(t, x_1, \dots, x_k)\right|^2.$$

Concerning (8.3), we start noting that

$$g_\lambda(t, W^1, \dots W^k_t) \to 0 \quad \text{and} \quad g_\lambda(t, W^1, \dots W^k_t) \le 4\|\varphi\|^2_\infty, \quad \text{for } \mathbb{P} \otimes dt\text{-a.e. } (\omega, t).$$

Hence, for each $\beta > 0$, we apply dominated convergence theorem twice, i.e. to the time integral and then to the expectation, thereby obtaining

$$\mathbb{E}\left[e^{\beta \int_0^T g_\lambda(t, W^1_t, \dots, W^k_t)\, dt}\right] \to 1 \quad \text{as} \quad \lambda \to 0, \tag{8.28}$$

that is (8.3). Hence we can apply Theorem 8.1 and obtain the desired LDP.

(3) Finally, assume that φ satisfies (8.25). We take

$$\tilde\varphi_\lambda = (\varphi \wedge 1/\lambda) \vee (-1/\lambda).$$

We take also an increasing sequence $(\beta_\lambda)_\lambda$ with $\beta_\lambda \to \infty$. For each $\lambda > 0$ fixed, applying (8.28) to $\tilde\varphi_\lambda$ in place of φ, we get the existence of a Lipschitz function φ_λ as in **(1)**, such that

$$\mathbb{E}\left[e^{\beta_\lambda \int_0^T |\tilde\varphi_\lambda - \varphi_\lambda|^2 (t, W^1_t, \dots, W^k_t)\, dt}\right] < 1 + \lambda. \tag{8.29}$$

Now we take b^N_λ, b_λ the drifts for the particle system and the McKean–Vlasov SDE, as respectively in (8.23), (8.24) with φ_λ in place of φ. As before, in order to apply Theorem 8.1, with (b^N_λ, b_λ) as sequence of $\mathcal{F}$Lip-inducing drifts, we verify the assumptions (8.1), (8.2) and (8.3). As before, the estimates (8.1) and (8.2) are easily obtained with

$$g_\lambda(t, x_1, \dots, x_k) = \left|\varphi(t, x_1, \dots, x_k) - \varphi_\lambda(t, x_1, \dots, x_k)\right|^2.$$

Concerning (8.3), we split $g_\lambda \le 2|\varphi - \tilde\varphi_\lambda|^2 + 2|\tilde\varphi_\lambda - \varphi_\lambda|^2$ and study the two terms separately. For the first term, for every $\beta > 0$, we have by dominated convergence theorem

$$\limsup_{\lambda\to 0} \mathbb{E}\left[e^{\beta \int_0^T |\varphi - \tilde\varphi_\lambda|^2 (t, W^1_t, \dots, W^k_t)\, dt}\right] = 1.$$

For the second term, the bound (8.29) implies, for every $\beta > 0$,

$$\limsup_{\lambda\to 0}\mathbb{E}\left[e^{\beta\int_0^T|\tilde{\varphi}_\lambda-\varphi_\lambda|^2(t,W_t^1,\dots,W_t^k)\,dt}\right]\le\limsup_{\lambda\to 0}\mathbb{E}\left[e^{\beta_\lambda\int_0^T|\tilde{\varphi}_\lambda-\varphi_\lambda|^2(t,W_t^1,\dots,W_t^k)\,dt}\right]=1.$$

The two bounds above give (8.3). Hence we can apply Theorem 8.1 and obtain the desired LDP. The proof is complete. □

8.2.3 Example: Measure-dependent drift

As an example of a more general interaction, we consider drifts of the form (cf. [GNP21])

$$b_t^N(x_i,z_{\boldsymbol{x}}^N)=\frac{1}{N}\sum_{j\neq i}\Psi\left(x_i,x_j,\frac{1}{N}\sum_{\ell\neq i}\varphi_t(x_i,x_\ell),\frac{1}{N}\sum_{\ell\neq j}\varphi_t(x_j,x_\ell),(z_{\boldsymbol{x}}^N)_t\right),\tag{8.30}$$

where $\Psi:\mathbb{R}^{4d}\times\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}^d$ and $\varphi:[0,T]\times\mathbb{R}^{2d}\to\mathbb{R}^d$ are Borel maps. As in the previous examples, the rigorous definition of b^N can be given as in Remark 8.5. Similarly, we take the drift of the associated McKean–Vlasov SDE

$$b_t(x,\mu)=\int_{\mathbb{R}^d}\Psi\left(x,y,\int_{\mathbb{R}^d}\varphi_t(x,z)\,d\mu_t(z),\int_{\mathbb{R}^d}\varphi_t(y,z)\,d\mu_t(z),\mu_t\right)d\mu_t(y).$$

Throughout we will use both the 1-Wasserstein metric, W_1 and the bounded Lipschitz metric, d_{BL} on $\mathcal{P}(\mathbb{R}^d)$, with the latter given by

$$d_{BL}(\mu,\nu):=\sup_{\phi:\,\|\phi\|_\infty\le 1,\mathrm{Lip}(\phi)\le 1}\left\{\int_{\mathbb{R}^d}\phi(x)d\mu(x)-\int_{\mathbb{R}^d}\phi(x)d\nu(x)\right\}.\tag{8.31}$$

Note that $d_{BL}\le W_1$.

Proposition 8.12. *Assume that the initial law ρ_0 satisfies condition* (7.9). *Furthermore, let $\Psi\in\mathrm{Lip}(\mathbb{R}^{4d}\times(\mathcal{P}(\mathbb{R}^d),d_{BL}))$, and suppose that there exists a constant L such that*

$$\Psi(x,y,a,b,\mu)\le L(1+|a|+|b|),\qquad x,y,a,b\in\mathbb{R}^d,\mu\in\mathcal{P}(\mathbb{R}^d)\tag{8.32}$$

and

$$\mathbb{E}\left[e^{\beta\int_0^T|\varphi_t|^2(W_t^1,W_t^2)\,dt}\right]<\infty,\qquad \textit{for all }\beta\in\mathbb{R},\tag{8.33}$$

where W^1,W^2 are independent Brownian motions with common initial law ρ_0.

Then the family Q_b^N of laws of the empirical processes associated to the interacting system (7.1) *induced by drifts of the form* (8.30) *satisfies an LDP with rate function $\mathcal{F}$ given in* (8.4).

In particular, (8.33) *holds whenever φ is translation invariant with $\varphi \in L_t^q(L_{x-y}^p)$ for $p, q \in [2, \infty]$ satisfying*

$$\frac{d}{p} + \frac{2}{q} < 1. \tag{8.34}$$

Remark 8.13. A couple of comments:

1. Similar to Remark 8.11, Proposition 8.12 holds for any $\varphi = \sum_{j=1}^m \varphi_j$ with each φ_j satisfying (8.34) for suitable exponents $p^{(j)}, q^{(j)}$.
2. As in Remark 4.6, when φ is bounded, we can include self-interactions in the summations in (8.18), (8.23) and (8.30) without changing the results above.

□

Remark 8.14. The example of Theorem 6.1 is a particular case of (8.30), which can be seen by setting

$$\Psi(x, x_2, y, x_3, \mu) := \psi(x, \mu, y).$$

The LDP for Q_b^N follows from Proposition 8.12. Moreover, the convergence to the McKean–Vlasov SDE will be shown in the next chapter, and follows directly from Proposition 9.6. □

Proof of Proposition 8.12. The proof is an adaptation of the proof for Proposition 8.9. As in step **(1)** of the proof of Lemma 8.9, we first consider suitable bounded Lipschitz functions φ with corresponding drifts b. Also in this case, it is not difficult to see that (b^N, b) is $\mathcal{F}$Lip-inducing. First, to show that $b \in \mathcal{F}$Lip, note that b is bounded by (8.32) and the boundedness of φ yields

$$|b_t(x_1, \mu_1) - b_t(x_2, \mu_2)| \leq \mathrm{Lip}(\Psi) \iint_{\mathbb{R}^d \times \mathbb{R}^d} \Big(|x_1 - x_2| + |y_1 - y_2| + (\mathrm{I}) + (\mathrm{II}) \Big) \, d\gamma(y_1, y_2) + (\mathrm{III})$$

where (III) $= \mathrm{Lip}(\Psi)\, d_{BL}(\mu, \nu)$ and

$$(\mathrm{I}) \leq \mathrm{Lip}(\varphi) \iint_{\mathbb{R} \times \mathbb{R}^d} \Big(|x_1 - x_2| + |w_1 - w_2| \Big) \, d\gamma(w_1, w_2),$$
$$(\mathrm{II}) \leq \mathrm{Lip}(\varphi) \iint_{\mathbb{R} \times \mathbb{R}^d} \Big(|y_1 - y_2| + |w_1 - w_2| \Big) \, d\gamma(w_1, w_2).$$

Noting $d_{BL} \leq W_1$, inserting estimates (I) and (II) into the previous inequality and optimizing over all couplings γ between μ_1 and μ_2 yields

$$|b_t(x_1, \mu_1) - b_t(x_2, \mu_2)| \leq \mathrm{Lip}(\Psi)(2 + 3\mathrm{Lip}(\varphi))\big(|x_1 - x_2| + W_1(\mu_1, \mu_2)\big),$$

i.e., the Lipschitz estimates holds. Next, note that we have for P^N-almost every z^N

$$\int_0^T \left\langle |b_t - b_t^N|^2(\cdot, z_t^N), z_t^N \right\rangle dt \leq \left(\mathrm{Lip}(\Psi)\|\varphi\|_\infty \frac{2}{N} + \frac{1}{N} L(1 + 2\|\varphi\|_\infty) \right)^2,$$

which vanishes as $N \to \infty$.

Finally, since Ψ is Lipschitz, we find that for any φ, b and approximating sequence $\varphi^\lambda, b^\lambda$ the estimates (8.1) and (8.2) may be obtained with $g_\lambda(t, x, y) = \mathrm{Lip}(\Psi)^2 |\varphi_t - \varphi_t^\lambda|^2$, and the rest of the proof follows similar to Proposition 8.9. □

Background There are several papers dealing with large deviations for McKean–Vlasov SDEs. One of the first papers is [DG87], which proves an LDP for the paths of empirical measure, assuming continuity on the drift b among other hypotheses. The papers [DMZ03] and [BDF12] prove an LDP (for the empirical measures on the path space, as here), again assuming also continuity of the drift. The work [Lac18] proves propagation of chaos and a large deviation upper bound, assuming b bounded and continuous in the measure argument, but with respect to a stronger topology (the τ topology). Outside the context of bounded (or linear growth) drift, we are aware only of the result in [Fon04], which shows an LDP for a system of one-dimensional particles with a repulsive two-body interaction of order $1/x$, that is the framework of Section 8.2.1 but with $\varphi(x) = 1/x$, which is outside the class we can deal with here.

We also recall the recent papers [Orr20] and [dRST19] on large deviations for interacting diffusions/McKean–Vlasov SDEs when also the noise intensity tends to 0, and [GNP21] for the large deviations of the Brownian one-dimensional hard-rod system.

Chapter 9

Uniqueness for the McKean–Vlasov SDE

In this chapter, we consider the McKean–Vlasov SDE associated with the particle system (7.1), namely

$$\begin{cases} dX_t = b_t(X_t, \mathrm{Law}(X_t))\,dt + dW_t, \\ X_0 \text{ with law } \rho_0, \end{cases} \tag{9.1}$$

for a given Borel drift $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and a given initial law ρ_0.

We assume conditions on b which include the examples in the previous chapters. We show, under these conditions, uniqueness in law for the McKean–Vlasov SDE (9.1). As a consequence, in all examples in the previous chapter, the empirical measures associated with the particle system (7.1) converge, as $N \to \infty$, to the law of the solution of the McKean–Vlasov SDE (9.1).

We keep the notation of the previous chapters, with $S = C([0,T];\mathbb{R}^d)$, $\mathbb{W}$ the d-dimensional Wiener measure with initial law ρ_0 and W the Brownian motion with initial law ρ_0 (similarly W^i are independent Brownian motion starting from ρ_0). We fix $m > 2$ and, for $\eta > 0$, we set

$$M_\eta := \left\{ \mu \in \mathcal{P}(S) \;\middle|\; \left\| \frac{d\mu}{d\mathbb{W}} \right\|_{L^m(S,\mathbb{W})} \le \eta \right\}.$$

We introduce some notation and give some remarks on the restriction from paths on $[0,T]$ to path on a subinterval $[0,T_2]$. We call $\pi_{[0,T_2]} : S = C([0,T];\mathbb{R}^d) \to S_{[0,T_2]} := C([0,T_2];\mathbb{R}^d)$ the restriction map; for $\gamma \in S$, we call $\gamma_{[0,T_2]} = \pi_{[0,T_2]}(\gamma)$ its restriction to $[0,T_2]$. For $\rho \in \mathcal{P}(S)$, we call $\rho_{[0,T_2]}$ the restriction of ρ to the interval $[0,T_2]$, that is $\rho_{[0,T_2]} = (\pi_{[0,T_2]})_\# \rho$, which is a probability measure on $S_{[0,T_2]}$; similarly we use $d\rho_{[0,T_2]}/d\mathbb{W}$ for the density of $\rho_{[0,T_2]}$ with respect to $\mathbb{W}_{[0,T_2]}$. Note that $d\rho_{[0,T_2]}/d\mathbb{W}$ is a version of the conditional expectation of $d\rho/d\mathbb{W}$ given $\pi_{[0,T_2]}$; in particular, if $d\rho/d\mathbb{W} \in L^\gamma(S,\mathbb{W})$ for some $\gamma \geq 1$, then $d\rho_{[0,T_2]}/d\mathbb{W}$ is in $L^\gamma(S_{[0,T_2]}, \mathbb{W}_{[0,T_2]})$ and we have (by Jensen's inequality for conditional expectation)

$$\|d\rho_{[0,T_2]}/d\mathbb{W}\|_{L^\gamma} \leq \|d\rho/d\mathbb{W}\|_{L^\gamma}.$$

We keep a similar notation for $\rho^{\otimes k}_{[0,T_2]}$; note that $d\rho^{\otimes k}_{[0,T_2]}/d\mathbb{W}^{\otimes k}$ (the density of $\rho^{\otimes k}_{[0,T_2]}$ with respect to $\mathbb{W}^{\otimes k}_{[0,T_2]}$) coincides with $(d\rho_{[0,T_2]}/d\mathbb{W})^{\otimes k}$ (the density tensorised k times). Finally, we call $V_{b,[0,T_2]}$ the map V_b (as defined in Section 7.1) with final time T_2 instead of T.

Theorem 9.1. *Fix the initial law $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$. Fix $m > 2$. Assume that, for every $\eta > 0$, for every $\beta > 0$,*

$$\sup_{\mu\in\mathcal{P}(S),\, R(\mu\|\mathbb{W})\leq\eta} \mathbb{E}\left[\exp\left(\beta\int_0^T |b_t(W_t,\mu_t)|^2 dt\right)\right] < \infty. \tag{9.2}$$

Assume also that there exists a non-negative Borel function $g : [0,T]\times(\mathbb{R}^d)^k \to \mathbb{R}$ and, for every $\eta > 0$, a constant $C_\eta \geq 0$, such that, for every $0 \leq T_1 < T_2 \leq T$,

$$\begin{aligned}\int_{T_1}^{T_2} &|b_t(x_t,\mu_t) - b_t(x_t,\nu_t)|^2 dt\\ &\leq C_\eta \int_{S^{k-1}_{[0,T_2]}} \left(\int_{T_1}^{T_2} g_t(x_t,y_t)\,dt\right) \left|u^{\otimes(k-1)}_{[0,T_2]}(y) - v^{\otimes(k-1)}_{[0,T_2]}(y)\right|^2 d\mathbb{W}^{\otimes(k-1)}_{[0,T_2]}(y)\\ &\qquad \text{for } \mathbb{W}\text{-a.e. } x \text{ and every } \mu,\nu \in M_\eta,\end{aligned} \tag{9.3}$$

where $u_{[0,T_2]} = d\mu_{[0,T_2]}/d\mathbb{W}$ and $v_{[0,T_2]} = d\nu_{[0,T_2]}/d\mathbb{W}$, and that, for some $\tilde{m} > m$ with $\tilde{m} \geq (m/2)' = m/(m-2)$,

$$\mathbb{E}\left[\left(\int_0^T g_t(W^1_t,\dots W^k_t)\,dt\right)^{\tilde{m}}\right] < \infty. \tag{9.4}$$

Then uniqueness in law holds for the McKean–Vlasov equation (9.1) *among laws with finite relative entropy with respect to $\mathbb{W}$, that is: if X and Y are two solutions with laws $\bar{\mu}$ and $\bar{\nu}$ respectively, and if $R(\bar{\mu}\|\mathbb{W}) < \infty$ and $R(\bar{\nu}\|\mathbb{W}) < \infty$, then $\bar{\mu} = \bar{\nu}$.*

9.1 Proof of main result

Proof. For any $\mu \in \mathcal{P}(S)$ with $R(\mu\|\mathbb{W}) < \infty$, assumption (9.2) gives via Girsanov's theorem D.1 the existence of a weak solution X^μ to the SDE

$$dX^\mu_t = b_t(X^\mu_t,\mu_t)\,dt + dW_t, \tag{9.5}$$

with law $F(\mu) := \mathbb{W}^\mu = \mathrm{Law}(X^\mu)$ given by

$$\frac{dF(\mu)}{d\mathbb{W}}(W) = \exp\big(V_b(W,\mu)\big),$$

with V_b as defined in Section 7.1. Note that $F(\mu)$ is the unique law solving (9.5) and having finite entropy with respect to $\mathbb{W}$. Indeed, if ν is the law of another solution Y^μ to (9.5) with $R(\nu\|\mathbb{W}) < \infty$, then, by Lemma B.1,

$$\int_S \int_0^T |b_t(x_t,\mu_t)|^2 dt\, d\nu(x) \leq R(\nu\|\mathbb{W}) + \log \mathbb{E}\left[e^{\int_0^T |b_t(W_t,\mu_t)|^2 dt}\right] < \infty.$$

Therefore the uniqueness condition (D.3) is met under ν and so $\nu = F(\mu)$.

Moreover, the density of $F(\mu)$ with respect to $\mathbb{W}$ is in $L^\gamma(S,\mathbb{W})$ for every finite $\gamma > 1$: Indeed, following standard computations (similarly to the proof of Theorem 8.1) and using assumption (9.2), we have for some c_γ,

$$\mathbb{E}\left[\left|\frac{dF(\mu)}{d\mathbb{W}}(W)\right|^\gamma\right] \leq \mathbb{E}\left[\exp\left(c_\gamma \int_0^T |b_t(W_t,\mu_t)|^2 dt\right)\right] < \infty. \tag{9.6}$$

Finally, note that a process X, with $R(\mathrm{Law}(X)\|\mathbb{W}) < \infty$, is a solution to the McKean–Vlasov SDE if and only if its law is a fixed point for F. In particular, if $\bar\mu$ is the law of a solution to the McKean–Vlasov SDE, with $R(\bar\mu\|\mathbb{W}) < \infty$, then $\bar\mu \in L^\gamma(S,\mathbb{W})$ for every finite γ and satisfies (9.6) (with $\bar\mu$ in place of μ and $F(\mu)$).

Let $\bar\mu$ and $\bar\nu$ be the laws of two solutions to the McKean–Vlasov SDE and $\mathcal{T}$ be the subset of $[0,T]$ consisting of all times T_1 such that $\bar\mu$ and $\bar\nu$ coincide when restricted to $[0,T_1]$ (that is, $\bar\mu_{[0,T_1]} = \bar\nu_{[0,T_1]}$). Clearly, $\mathcal{T}$ is an interval containing 0. We will show that $\mathcal{T}$ is both open and closed in $[0,T]$, thus implying that $\mathcal{T}$ must coincide with $[0,T]$, and hence $\bar\mu = \bar\nu$.

The fact that $\mathcal{T}$ is closed follows easily from the fact that $\bar\mu$ and $\bar\nu$ are measures on continuous paths: indeed, if $(T_n)_n$ is an increasing sequence in $[0,T]$ converging to T' and $\bar\mu$ and $\bar\nu$ coincide up to T_n for every n, then $\bar\mu$ and $\bar\nu$ coincide up to T' as well.

We will now show that $\mathcal{T}$ is open, that is, if $\bar\mu$ and $\bar\nu$ coincide up to T_1, then there exists $T_2 > T_1$ such that $\bar\mu$ and $\bar\nu$ coincide up to T_2. Fix $\eta > 0$ such that $\bar\mu$ and $\bar\nu$ belong to M_η (such η exists due to (9.6)). Fix also $T_2 > T_1$ to be specified later. Unless differently specified, all constants in the proof are independent of T_2.

Using the elementary inequality $|e^a - e^b| \leq \frac{1}{2}(e^a + e^b)|a-b|$, we get

$$\begin{aligned}\left|\frac{d\bar\mu_{[0,T_2]}}{d\mathbb{W}} - \frac{d\bar\nu_{[0,T_2]}}{d\mathbb{W}}\right| &= \left|\frac{dF(\bar\mu)_{[0,T_2]}}{d\mathbb{W}} - \frac{dF(\bar\nu)_{[0,T_2]}}{d\mathbb{W}}\right| \\ &\leq \frac{1}{2}\left(\frac{dF(\bar\mu)_{[0,T_2]}}{d\mathbb{W}} + \frac{dF(\bar\nu)_{[0,T_2]}}{d\mathbb{W}}\right)|V_{b,[0,T_2]}(\cdot,\bar\mu) - V_{b,[0,T_2]}(\cdot,\bar\nu)|.\end{aligned}$$

So by Hölder's inequality with $1/\gamma+1/\tilde{m} = 1/m$ (with shorthand $L^\gamma = L^\gamma(S_{[0,T_2]}, \mathbb{W}_{[0,T_2]})$, $\gamma > 1$),

$$\left\| \frac{dF(\bar{\mu})_{[0,T_2]}}{d\mathbb{W}} - \frac{dF(\bar{\nu})_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m} \leq \frac{1}{2} \left(\left\| \frac{dF(\bar{\mu})}{d\mathbb{W}} \right\|_{L^\gamma} + \left\| \frac{dF(\bar{\nu})}{d\mathbb{W}} \right\|_{L^\gamma} \right) \|V_b(\cdot, \bar{\mu}) - V_b(\cdot, \bar{\nu})\|_{L^{\tilde{m}}}.$$

By (9.6), (9.2) and the fact that $R(\rho\|\mathbb{W}) \leq \|d\rho/d\mathbb{W}\|^m_{L^m(S,\mathbb{W})}$ for any $\rho \in \mathcal{P}(S)$, we have, for some $C_\eta > 0$,

$$\begin{aligned} \left\| \frac{dF(\bar{\mu})}{d\mathbb{W}} \right\|_{L^\gamma} + \left\| \frac{dF(\bar{\nu})}{d\mathbb{W}} \right\|_{L^\gamma} &\leq \left\| \frac{dF(\bar{\mu})}{d\mathbb{W}} \right\|_{L^\gamma(S,\mathbb{W})} + \left\| \frac{dF(\bar{\nu})}{d\mathbb{W}} \right\|_{L^\gamma(S,\mathbb{W})} \\ &\leq 2 \sup_{\rho \in M_\eta} \mathbb{E} \left[\exp\left(c_\gamma \int_0^T |b_t(W_t, \rho_t)|^2 dt \right) \right] \\ &\leq 2 \sup_{\rho \in \mathcal{P}(S), R(\rho\|\mathbb{W}) \leq \eta^m} \mathbb{E} \left[\exp\left(c_\gamma \int_0^T |b_t(W_t, \rho_t)|^2 dt \right) \right] \leq C_\eta. \end{aligned}$$

By definition of $V_{b,[0,T_2]}$ (in Section 7.1 with final time T_2), we obtain

$$\begin{aligned} \|V_{b,[0,T_2]}(\cdot, \bar{\mu}) - V_{b,[0,T_2]}(\cdot, \bar{\nu})\|_{L^{\tilde{m}}} &\leq \left\| \int_0^{T_2} (b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)) \cdot dW_t \right\|_{L^{\tilde{m}}} \\ &\quad + \frac{1}{2} \left\| \int_0^{T_2} \Big| |b_t(W_t, \bar{\mu}_t)|^2 - |b_t(W_t, \bar{\nu}_t)|^2 \Big| dt \right\|_{L^{\tilde{m}}} = \text{(I)} + \text{(II)}. \end{aligned}$$

For the first term, we obtain by the Burkholder–Davis–Gundy inequality

$$\text{(I)} \leq \tilde{c}_1 \left\| \int_0^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right\|^{\frac{1}{2}}_{L^{\tilde{m}/2}} \leq c_1 \left\| \int_0^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right\|^{\frac{1}{2}}_{L^{\tilde{m}}}$$

for some constants $\tilde{c}_1, c_1 > 0$. As for the second term, we estimate as follows

$$\begin{aligned} \text{(II)} &\leq \tilde{c}_2 \left\| \left(\int_0^{T_2} \Big(|b_t(W_t, \bar{\mu}_t)|^2 + |b_t(W_t, \bar{\nu}_t)|^2 \Big) dt \right)^{\frac{1}{2}} \cdot \left(\int_0^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right)^{\frac{1}{2}} \right\|_{L^{\tilde{m}}} \\ &\leq c_2 \left\| \int_0^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right\|^{\frac{1}{2}}_{L^{\tilde{m}}} \\ &\quad \cdot \left(\left\| \int_0^{T_2} |b_t(W_t, \bar{\mu}_t)|^2 dt \right\|^{\frac{1}{2}}_{L^{\tilde{m}}} + \left\| \int_0^{T_2} |b_t(W_t, \bar{\nu}_t)|^2 dt \right\|^{\frac{1}{2}}_{L^{\tilde{m}}} \right) \end{aligned}$$

for some constants $\tilde{c}_2, c_2 > 0$. Using the inequality $a^{\tilde{m}} \leq e^{\tilde{m}a}$, we obtain the estimate

$$\left\| \int_0^{T_2} |b_t(W_t, \bar{\mu}_t)|^2 dt \right\|_{L^{\tilde{m}}}^{\frac{1}{2}} \leq \mathbb{E}\left[\exp\left(\tilde{m}\int_0^{T_2} |b_t(W_t, \bar{\mu}_t)|^2 dt\right)\right]^{\frac{1}{2\tilde{m}}}$$
$$\leq \sup_{\rho\in M_\eta} \mathbb{E}\left[\exp\left(\tilde{m}\int_0^{T} |b_t(W_t, \rho_t)|^2 dt\right)\right]^{\frac{1}{2\tilde{m}}} \leq C_\eta,$$

where we have used again (9.2) and the fact that $R(\rho\|\mathbb{W}) \leq \|d\rho/d\mathbb{W}\|_{L^m(S,\mathbb{W})}^m$. The same argument holds for the term with $\int_0^{T_2} |b_t(W_t, \bar{\nu}_t)|^2 dt$.

Putting the terms (I) and (II) together, and using that $\bar{\mu}_{[0,T_1]} = \bar{\nu}_{[0,T_1]}$, we then obtain

$$\|V_{b,[0,T_2]}(\cdot, \bar{\mu}) - V_{b,[0,T_2]}(\cdot, \bar{\nu})\|_{L^{\tilde{m}}} \leq c_3 \left\| \int_0^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right\|_{L^{\tilde{m}}}^{\frac{1}{2}}$$
$$= c_3 \left\| \int_{T_1}^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right\|_{L^{\tilde{m}}}^{\frac{1}{2}}$$

for some constant $c_3 > 0$. Using assumption (9.3) (with the notation $\boldsymbol{W} = (W^1, \dots, W^k)$ and $\boldsymbol{W}^{,1} = (W^2, \dots, W^k)$, where W^i are independent d-dimensional Brownian motions with initial law ρ_0), we further estimate the right-hand side to obtain

$$\left\| \int_{T_1}^{T_2} |b_t(W_t, \bar{\mu}_t) - b_t(W_t, \bar{\nu}_t)|^2 dt \right\|_{L^{\tilde{m}}}^{1/2}$$
$$\leq c_4 \mathbb{E}^{W^1}\left[\left(\mathbb{E}^{\boldsymbol{W}^{,1}}\left[\int_{T_1}^{T_2} g_t(\boldsymbol{W}_t) dt \left| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}}^{\otimes(k-1)} - \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}}^{\otimes(k-1)} \right|^2 ({\boldsymbol{W}^{,1}}_{[0,T_2]})\right]\right)^{\tilde{m}}\right]^{1/(2\tilde{m})}$$
$$\leq c_5 \mathbb{E}^{W^1}\Bigg[\Bigg(\mathbb{E}^{\boldsymbol{W}^{,1}}\Bigg[\int_{T_1}^{T_2} g_t(\boldsymbol{W}_t)\, dt \cdot$$
$$\cdot \sum_{j=2}^{k} \left| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} - \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right|^2 (W^j_{[0,T_2]}) \prod_{2\leq \ell\leq k, \ell\neq j} \left| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} \vee \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right|^2 (W^\ell_{[0,T_2]})\Bigg]\Bigg)^{\tilde{m}}\Bigg]^{1/(2\tilde{m})}$$
$$\leq c_6 \left(\mathbb{E}\left[\int_{T_1}^{T_2} g_t(\boldsymbol{W}_t)\, dt\right]^{\tilde{m}}\right)^{1/(2\tilde{m})} \left\| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} \vee \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m}^{k-2} \left\| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} - \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m}$$
$$\leq c_7 \left(\mathbb{E}\left[\int_{T_1}^{T_2} g_t(\boldsymbol{W}_t)\, dt\right]^{\tilde{m}}\right)^{1/(2\tilde{m})} \eta^{k-2} \left\| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} - \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m}$$

for appropriate constants $c_i > 0$, $i = 1, \dots, 7$. We conclude that, for some $C_\eta > 0$ (independent of T_1 and T_2),

$$\left\| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} - \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m} = \left\| \frac{dF(\bar{\mu})_{[0,T_2]}}{d\mathbb{W}} - \frac{dF(\bar{\nu})_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m}$$
$$\leq C_\eta \left(\mathbb{E} \left(\int_{T_1}^{T_2} g_t(\boldsymbol{W}_t)\, dt \right)^{\tilde{m}} \right)^{1/(2\tilde{m})} \left\| \frac{d\bar{\mu}_{[0,T_2]}}{d\mathbb{W}} - \frac{d\bar{\nu}_{[0,T_2]}}{d\mathbb{W}} \right\|_{L^m}.$$

By assumption (9.4), we can find $T_2 > T_1$ close enough to T_1 such that

$$C_\eta \left(\mathbb{E} \left[\int_{T_1}^{T_2} g_t(\boldsymbol{W}_t)\, dt \right]^{\tilde{m}} \right)^{1/(2\tilde{m})} < 1.$$

Hence, for such T_2, we get that $d\bar{\mu}_{[0,T_2]}/d\mathbb{W} = d\bar{\nu}_{[0,T_2]}/d\mathbb{W}$, that is $\bar{\mu}$ and $\bar{\nu}$ coincide up to T_2. The proof is complete. □

Remark 9.2. As the proof shows, assumption (9.3) may be replaced by the following weaker assumption: For $\mathbb{W}$-a.e. x and every $\mu, \nu \in M_\eta$,

$$\int_{T_1}^{T_2} |b_t(x_t, \mu_t) - b_t(x_t, \nu_t)|^2 dt \leq C_\eta \int_{S^{k-1}} \left(\int_{T_1}^{T_2} g(t, x_t, y_t)\, dt \right) \cdot \tag{9.7}$$
$$\cdot \sum_{j=2}^{k} \left| u_{[0,T_2]}(y^j) - v_{[0,T_2]}(y^j) \right|^2 \prod_{\substack{2 \leq \ell \leq k \\ \ell \neq j}} (u_{[0,T_2]} \vee v_{[0,T_2]})(y^\ell)^2 d\mathbb{W}^{\otimes(k-1)}(y),$$

where $u_{[0,T_2]} = d\mu_{[0,T_2]}/d\mathbb{W}$ and $v_{[0,T_2]} = d\nu_{[0,T_2]}/d\mathbb{W}$.

We will use this generalization for the example of Section 8.2.3. □

Remark 9.3. The proof of Theorem 9.1 shows, morally, a contraction bound for F in M_η, provided that T_2 is sufficiently close to T_1. From this bound it should be possible to get local existence for the McKean–Vlasov SDE. Unfortunately, the iteration in time scheme does not work easily, because the constant η bounding the L^m norm of the solution and the time step $T_2 - T_1$ can change from one time step to the next. Nonetheless, in all the examples where our LPD result Theorem 8.1 applies, existence for the McKean–Vlasov SDE holds due to existence of a minimizer of the rate function $\mathcal{F}$. □

Corollary 9.4. *Under the assumptions of Theorems 8.1 and 9.1 (possibly modified as in Remark 9.2), as $N \to \infty$, the family of empirical measures $z^N_{\boldsymbol{X}}$ associated with the particle system* (7.1) *converges almost surely to the (unique) law $\bar{\mu}$ of the McKean–Vlasov SDE* (9.1) *(with finite relative entropy with respect to $\mathbb{W}$).*

Proof. The zeros of the rate function $\mathcal{F}$ in Theorem 8.1 are exactly the laws of the solutions to the McKean–Vlasov SDE (9.1) with finite relative entropy with respect to $\mathbb{W}$. In particular, since $\mathcal{F}$ is a good rate function, there exists at least one law solution to the McKean–Vlasov SDE. By Theorem 9.1, such a law is unique (among the probability measures with finite entropy with respect to $\mathbb{W}$), and it is then the unique zero of the rate function $\mathcal{F}$ of the LDP for $(z^N_{\boldsymbol{X}})_N$ in Theorem 8.1. Hence Lemma 3.4 applies. □

Remark 9.5. As noted by Sznitman in [Szn98, Szn91], convergence in law of the empirical measures z^N to the constant variable $\bar{\mu}$ implies that the sequence $\tilde{Q}^N_{b^N} \in \mathcal{P}(S^N)$ is $\bar{\mu}$-chaotic, in the sense that for every $k \in \mathbb{N}$,

$$\text{Law}\left(X^{N,1}, \dots, X^{N,k}\right) \to \bar{\mu}^{\otimes k},$$

weakly as $N \to \infty$ on $S = C([0,T];\mathbb{R}^d)$. In particular, we have a form of propagation of chaos, namely that for all $k \in \mathbb{N}$,

$$\text{Law}\left(X^{N,1}_t, \dots, X^{N,k}_t\right) \to \bar{\mu}^{\otimes k}_t, \qquad \forall t \in [0,T].$$

Moreover, adapting the arguments of Remark 4.8 and the proof of Proposition 4.9, we would expect to obtain entropic chaoticity, i.e.

$$\lim_{N\to\infty} \frac{1}{N} R(Q^N | P^N_{\bar{\mu}}) = 0, \tag{9.8}$$

which, after a subadditivity and a projection argument, implies that we would have *entropic propagation of chaos*, i.e. that for all $k \in \mathbb{N}$,

$$R\left(\text{Law}\left(X^{N,1}_t, \dots, X^{N,k}_t\right) \middle| \bar{\mu}^{\otimes k}_t\right) \to 0, \qquad \forall t \in [0,T].$$

We will establish a similar phenomenon in Part II.B.

One thing stopping us from concluding (9.8) is the fact that even for $b \in \mathcal{F}\text{Lip}$, we do not a priori know if

$$\lim_{N\to\infty} \int_{\mathcal{X}} \mathcal{E}^N_b(\mu) dQ^N_b = \mathcal{E}_b(\bar{\mu}).$$

It is possible that rough-path arguments as in [DFMS18] can alleviate this problem. □

9.2 Application to concrete examples

Now we come back to the examples in Sections 8.2.1, 8.2.2, 8.2.3, recalling the drift for the corresponding McKean–Vlasov SDEs, namely:

- for the example in Section 8.2.1,

$$b_t(x,\mu) = \int_{\mathbb{R}^d} \varphi_t(x,y)\,d\mu_t(y);$$

- for the example in Section 8.2.2,

$$b_t(x,\mu) = \int_{\mathbb{R}^{(k-1)d}} \varphi_t(x,y_2,\dots y_k)\,d\mu_t(y_2,\dots y_k);$$

- for the example in Section 8.2.3,

$$b_t(x,\mu) = \int_{\mathbb{R}^d} \Psi\left(x, y, \int_{\mathbb{R}^d} \varphi_t(x,z)\,d\mu_t(z), \int_{\mathbb{R}^d} \varphi_t(y,z)\,d\mu_t(z), \mu\right) d\mu_t(y).$$

Proposition 9.6. *For the examples in Sections 8.2.1, 8.2.2, 8.2.3, under the assumptions of Propositions 8.6, 8.9 and 8.12 respectively, the corresponding McKean–Vlasov SDEs admit a unique solution X with $R(\mathrm{Law}(X)\|\mathbb{W}) < \infty$, the family of empirical measures $z^N_{\boldsymbol{X}}$ converges almost surely to the law of X.*

Proof. We have seen in the previous section that the examples above satisfy the assumptions of Theorem 8.1, so it is enough to verify the assumptions of Theorem 9.1, possibly modified as in Remark 9.2. Assumption (9.2) is a consequence of (8.7) in the proof of Theorem 8.1. Assumptions (9.3) and (9.4) are satisfied, with any given m and $\tilde{m}$ as in Theorem 9.1:

- for the example in Section 8.2.1, taking $g(t,x,y) = |\varphi(t,x,y)|^2$;
- for the example in Section 8.2.2, taking $g(t,x_1,\dots x_k) = |\varphi(t,x_1,\dots x_k)|^2$.

We will not show the proof of this fact, which is similar to, and easier than, the next proof for the example in Section 8.2.3.

For the example in Section 8.2.3, we will show that assumptions (9.7) and (9.4) are satisfied, with L the linear growth constant for Ψ as in (8.32), taking

$$g(t,x,y,z) = \left(\mathrm{Lip}(\Psi)^2 + L^2\right)\left(1 + |\varphi_t(x,z)|^2 + |\varphi_t(y,z)|^2\right),$$

We start showing (9.7). We denote, for $0 \leq t \leq T_2$,

$$\Psi_t(x, y, \mu^1_{[0,T_2]}, \mu^2_{[0,T_2]}, \mu^3_{[0,T_2]}) = \Psi\left(x_t, y_t, \int_{S_{[0,T_2]}} \varphi_t(x_t, z_t) d\mu^1_{[0,T_2]}(z), \int_{S_{[0,T_2]}} \varphi_t(y_t, z_t) d\mu^2_{[0,T_2]}(z), \mu^3_t \right),$$

and also $u_{[0,T_2]} = d\mu_{[0,T_2]}/d\mathbb{W}$ and $v_{[0,T_2]} = d\nu_{[0,T_2]}/d\mathbb{W}$. For simplicity of notation, in the following proof of (9.7), we often omit the subscript $[0, T_2]$ from μ, ν, u, v, $\mathbb{W}$, S (up to the end of the proof of (9.7), every time we write μ, ν, ... we mean $\mu_{[0,T_2]}$, $\nu_{[0,T_2]}$). We have the following estimates, with a generic constant $C > 0$ (independent of T_1 and T_2): for any T_1, T_2 with $0 \leq T_1 < T_2 \leq T$,

$$\begin{aligned}
&\int_{T_1}^{T_2} |b_t(x_t, \mu_t) - b_t(x_t, \nu_t)|^2 dt \\
&= \int_{T_1}^{T_2} \left| \int_S \Psi_t(x, y, \mu, \mu, \mu)\, d\mu(y) - \int_S \Psi_t(x, y, \nu, \nu, \nu) d\nu(y) \right|^2 dt \\
&\leq C \int_{T_1}^{T_2} \left| \int_S \Psi_t(x, y, \mu, \mu, \mu) d(\mu - \nu)(y) \right|^2 dt \\
&\quad + C \int_{T_1}^{T_2} \left| \int_S [\Psi_t(x, y, \mu, \mu, \mu) - \Psi_t(x, y, \nu, \mu, \mu)] d\nu(y) \right|^2 dt \\
&\quad + C \int_{T_1}^{T_2} \left| \int_S [\Psi_t(x, y, \nu, \mu, \mu) - \Psi_t(x, y, \nu, \nu, \mu)] d\nu(y) \right|^2 dt \\
&\quad + C \int_{T_1}^{T_2} \left| \int_S [\Psi_t(x, y, \nu, \nu, \mu) - \Psi_t(x, y, \nu, \nu, \nu)] d\nu(y) \right|^2 dt \\
&=: \text{(I)} + \text{(II)} + \text{(III)} + \text{(IV)}.
\end{aligned}$$

For the term (I), we have

$$\begin{aligned}
\text{(I)} &= C \int_{T_1}^{T_2} \left| \int_S \Psi_t(x, y, \mu, \mu, \mu) d(\mu - \nu)(y) \right|^2 dt \\
&\leq CL^2 \int_{T_1}^{T_2} \left| \int_S \left(1 + \left| \int_S \varphi_t(x_t, z_t) d\mu(z) \right| + \left| \int_S \varphi_t(y_t, z_t) d\mu(z) \right| \right) |u(y) - v(y)| d\mathbb{W}(y) \right|^2 dt \\
&\leq CL^2 \int_{T_1}^{T_2} \left| \int_S \int_S \left(1 + |\varphi_t(x_t, z_t)| + |\varphi_t(y_t, z_t)| \right) |u(y) - v(y)| d\mathbb{W}(y) d\mu(z) \right|^2 dt \\
&= CL^2 \int_{T_1}^{T_2} \left| \int_{S^2} \left(1 + |\varphi_t(x_t, z_t)| + |\varphi_t(y_t, z_t)| \right) |u(y) - v(y)| u(z) d\mathbb{W}^{\otimes 2}(y, z) \right|^2 dt \\
&\leq 3C \int_{S^2} \left(\int_{T_1}^{T_2} L^2 (1 + |\varphi(t, x_t, z_t)|^2 + |\varphi(t, y_t, z_t)|^2) dt \right) |u(y) - v(y)|^2 u(z)^2 d\mathbb{W}^{\otimes 2}(y, z).
\end{aligned}$$

For the term (II), we have

$$
\begin{aligned}
\text{(II)} &\leq C \int_{T_1}^{T_2} \int_S \operatorname{Lip}(\Psi)^2 \left| \int_S \varphi(t, x_t, z_t) d(\mu - \nu)(z) \right|^2 d\nu(y) dt \\
&= C \int_{T_1}^{T_2} \operatorname{Lip}(\Psi)^2 \left| \int_{S^2} \varphi(t, x_t, z_t)(u(z) - v(z)) u(y) d\mathbb{W}^{\otimes 2}(y, z) \right|^2 dt \\
&\leq C \int_{S^2} \left(\int_{T_1}^{T_2} \operatorname{Lip}(\Psi)^2 |\varphi(t, x_t, z_t)|^2 dt \right) |u(z) - v(z)|^2 u(y)^2 d\mathbb{W}^{\otimes 2}(y, z),
\end{aligned}
$$

where $u(y)$ has been added artificially to satisfy (9.7) with $k = 3$. As for (III),

$$
\begin{aligned}
\text{(III)} &\leq C \int_{T_1}^{T_2} \int_S \operatorname{Lip}(\Psi)^2 \left| \int_S \varphi(t, y_t, z_t) d(\mu - \nu)(z) \right|^2 d\nu(y) dt \\
&= C \int_{T_1}^{T_2} \int_S \operatorname{Lip}(\Psi)^2 \left| \int_S \varphi(t, y_t, z_t)(u(z) - v(z)) d\mathbb{W}(z) \right|^2 v(y) d\mathbb{W}(y) dt \\
&\leq C \int_{S^2} \left(\int_{T_1}^{T_2} \operatorname{Lip}(\Psi)^2 |\varphi(t, y_t, z_t)|^2 dt \right) |u(z) - v(z)|^2 v(y)^2 d\mathbb{W}^{\otimes 2}(y, z).
\end{aligned}
$$

In view of the term (IV), we recall that, for any $t \in [0, T_2]$,

$$
d_{BL}(\mu_t, \nu_t) \leq d_{BL}(\mu_{[0,T_2]}, \nu_{[0,T_2]}),
$$

which follows from the fact that the map $e_t : S_{[0,T_2]} \to \mathbb{R}^d$, $e_t(x) := x_t$, is 1-Lipschitz. Hence,

$$
\begin{aligned}
\text{(IV)} &= C \int_{T_1}^{T_2} \left| \int_S [\Psi_t(x, y, \nu, \nu, \mu) - \Psi_t(x, y, \nu, \nu, \nu)] d\nu(y) \right|^2 dt \\
&\leq C \int_{T_1}^{T_2} \operatorname{Lip}(\Psi)^2 \, d_{BL}(\mu_t, \nu_t)^2 dt \\
&\leq C(T_2 - T_1) \operatorname{Lip}(\Psi)^2 \, d_{BL}(\mu, \nu)^2.
\end{aligned}
$$

We also recall that the bounded Lipschitz metric d_{BL} is bounded by the total variation distance d_{TV}. Therefore we derive, adding again an artificial term $u(z)$,

$$
\begin{aligned}
\text{(IV)} &\leq C(T_2 - T_1) \operatorname{Lip}(\Psi)^2 d_{TV}(\mu, \nu)^2 = C(T_2 - T_1) \operatorname{Lip}(\Psi)^2 \|u - v\|_{L^1}^2 \\
&= C \left(\int_{T_1}^{T_2} \operatorname{Lip}(\Psi)^2 \, dt \right) \left(\int_{S^2} |u(y) - v(y)| u(z) d\mathbb{W}^{\otimes 2}(y, z) \right)^2 \\
&\leq C \int_{S^2} \left(\int_{T_1}^{T_2} \operatorname{Lip}(\Psi)^2 \, dt \right) |u(y) - v(y)|^2 u(z)^2 d\mathbb{W}^{\otimes 2}(y, z).
\end{aligned}
$$

Putting together (I), (II), (III), (IV), we get (9.7). Finally, we verify assumption (9.4) for g (with fixed m and $\tilde{m}$ as in Theorem 9.1). Note that it is enough to prove

$$\mathbb{E}\left[\left(\int_0^T |\varphi_t(W_t^1, W_t^2)|^2 \, dt\right)^{\tilde{m}}\right] < \infty.$$

But this easily follows from the assumption (8.33) on φ,

$$\mathbb{E}\left[e^{\beta \int_0^T |\varphi_t|^2(W_t^1, W_t^2)\, dt}\right] < \infty, \qquad \forall \beta \in \mathbb{R},$$

which implies the finiteness of all moments of the variable $\int_0^T |\varphi_t(W_t^1, W_t^2)|^2 dt$. □

Background The convergence of the particle system to the McKean–Vlasov SDE, in the sense of Corollary 9.4, is classical in the case of Lipschitz bounded drift, see e.g. [Szn91]. The case of non-Lipschitz drift has also been treated in various works and we mention only some of them. In the context of the example in Section 8.2.1, the paper [JW18] proves the convergence, with quantitative bounds, for φ in $W^{-1,\infty}$ (which includes our example) such that $\mathrm{div}(\varphi)$ is in $W^{-1,\infty}$. The work [GQ15] covers the case $\varphi = -\nabla\Phi$ with $\Phi(x) = |x|^\alpha$ with α in $(0,1)$, which is a relevant example of Section 8.2.1 for $d \geq 2$ (see Remark 8.8). Both in [JW18] and [GQ15] the initial conditions are assumed to be diffuse in a suitable sense. The paper [Tom20] (which appeared after [HHMT20] came on arXiv), exploiting the technique in [JTT18], shows well-posedness and convergence of the particle system for φ as in our example in Section 8.2.1 (namely φ in $L_t^q(L_x^p)$ with p, q satisfying (8.22)), also possibly non-translation invariant, with a condition of continuity of φ outside a set of singular points. As examples of convergence in critical cases, that are not covered by our results, we recall [FJ17], for $\varphi = -\nabla\Phi$ with $\Phi(x) = \log|x|$, and [FHM14], for the 2D Navier-Stokes equations and the associated vortex system, that is $\varphi(x) = x^\perp/|x|^2$.

Outside the context of Section 8.2.1, we already mentioned [Lac18], proving convergence in the τ topology when b is bounded and satisfies a suitable continuity assumption in the measure argument. The paper [Jab19] proves convergence for a general measure-dependent drift, assuming a quite weak condition but depending on the solution to the McKean–Vlasov itself; the result is then applied to the case of bounded drifts. However we are not aware of a convergence result that covers our Corollary 9.4 and Proposition 9.6, although the recent work [HRZ22] (which has also appeared after [HHMT20] came on arXiv) proves convergence for a class of singular drifts similar to that in Proposition 9.6.

Concerning (weak or strong) uniqueness for McKean–Vlasov SDEs with irregular drifts, we mention [MS19, HvS21, CdR20, RZ21b, MV20, FPZ19]. It is also worth mentioning [Del19] on a regularization by noise phenomenon, via an infinite-dimensional noise, for a related mean-field game problem.

Appendix A

Limits of convex functions

Let V be a vector space. We consider a sequence of convex functions $\phi_n : V \to \overline{\mathbb{R}}$ (with domains D_n), and two convex functions ϕ_L and ϕ_U with $\phi_L \leq \phi_U$ on V (with respective domains D_L, D_U), which will act as asymptotic lower and upper bounds for ϕ_n in a sense specified below.

Moreover, we consider pairs of sequences and points $(\{y_n\}_{n\in\mathbb{N}}, y) \in V^{\mathbb{N}} \times V$, which for simplicity will be referred to as the pairs (y_n, y), and denote $(y_n, y) \in D$ for the statement $y \in D_U$ and $y_n \in D_n$ for every n. We then have the following approximation result.

Theorem A.1. *Let $\gamma > 1$, and let the family of pairs $\{(y_{\lambda,n}, y_\lambda)\}_\lambda$ be such that for every $\lambda > 0$,*

$$\limsup_{n\to\infty} \phi_n(\gamma y_{\lambda,n}) < \infty, \tag{A.1a}$$

$$\phi_U(\gamma y_\lambda) < \infty, \tag{A.1b}$$

$$\phi_L(y_\lambda) \leq \liminf_{n\to\infty} \phi_n\left(y_{\lambda,n}\right) \leq \limsup_{n\to\infty} \phi_n\left(y_{\lambda,n}\right) \leq \phi_U(y_\lambda). \tag{A.2}$$

Moreover, let the couple (x_n, x) be such that there exists a $K \in \mathbb{R}$ such that for all $\beta \in \mathbb{R}$,

$$\limsup_{\lambda\to 0} \limsup_{n\to\infty} \phi_n\left(\beta(y_{\lambda,n} - x_n)\right) \leq K, \tag{A.3a}$$

$$\limsup_{\lambda\to 0} \phi_U\left(\beta(y_\lambda - x)\right) \leq K. \tag{A.3b}$$

Then for any $0 < \gamma' < \gamma$,

$$\limsup_{n\to\infty} \phi_n(\gamma' x_n) < \infty, \tag{A.4a}$$

$$\phi_U(\gamma' x) < \infty, \tag{A.4b}$$

and

$$\phi_L(x) \le \liminf_{n\to\infty} \phi_n\big(x_n\big) \le \limsup_{n\to\infty} \phi_n\big(x_n\big) \le \phi_U(x). \tag{A.5}$$

Remark A.2. Some comments on these assumptions:

(i) Recall that the domain of a convex function ϕ is given by $D(\phi) := \{x \in V : \phi(x) < \infty\}$, and note that the convex functions ϕ_L, ϕ_N, ϕ are all allowed to be *improper*, i.e. allowed to be equal to $-\infty$ on their domain, or to have empty domain.

(ii) The constant K in (A.3) is independent of β, which is crucial in proving the convergence. As an example, note that when ϕ_U is even with $\phi_U(0) = 0$ the assumptions imply that (A.3b) holds for $K = 0$ as well.

□

First, we will establish some technical properties for limits of convex functions: a generalization of the classical statement on continuity of convex function on the interior of their domains to certain pointwise limits of convex functions, and a result that shows how under (A.3a) the limits in n, λ of $\phi_n(x), \phi_n(y_{\lambda,n})$ in effect 'commute'.

Lemma A.3. *Let $g_n : \mathbb{R} \to \overline{\mathbb{R}}$ be a sequence of convex functions such that for some $a, b \in \mathbb{R}$*

$$\limsup_{n\to\infty} g_n(a) < \infty, \qquad \limsup_{n\to\infty} g_n(b) < \infty.$$

Then the functions

$$\bar{g}(z) := \limsup_{n\to\infty} g_n(z), \qquad g(z) := \liminf_{n\to\infty} g_n(z)$$

are both either equal to $-\infty$ on (a, b), or finite and continuous on (a, b).

Proof. First, note that there exists a large enough N such that for all $n \ge N$ both $g_n(a)$ and $g_n(b)$ are bounded from above and, in particular, $[a, b] \subset D(g_n)$. Moreover, it is easy to verify that $\bar{g} = \limsup_{n\to\infty} g_n(z)$ is convex as well, $[a, b] \subset D(\bar{g})$, and with $\bar{g}$ bounded from above on $[a, b]$. We will show that either $g \equiv -\infty$ on (a, b) or that $\limsup_{n\to\infty} \|g_n\|_\infty$ is finite — in which case we will derive that g is Lipschitz continuous at any proper sub-interval $[c, d] \subsetneq [a, b]$ with $a < c < d < b$. The case for $\bar{g}$ follows from a similar argument and the fact that $\bar{g}$ is convex itself.

Now, suppose that $\limsup_{n\to\infty} \|g_n\|_\infty = \infty$. Since $\bar{g}(z)$ is bounded from above on $[a, b]$, it follows that there exists a subsequence $n' \in \mathbb{N}$ and a sequence $z_{n'} \in [a, b]$ such that

$$\lim_{n'\to\infty} g_{n'}(z_{n'}) = -\infty.$$

By compactness in $[a, b]$ we can choose a converging sequence $z_{n'} \to z^*$. We will treat the case $z^* \in (a, b)$ and $z^* = a, b$ separately. First, assume the former and fix $s \in (a, b)$. Then for small enough $\epsilon > 0$ and large enough n' such that $g_{n'}(z_{n'}) < \infty$ and $|z_{n'} - z^*| < \epsilon$ with $a + \epsilon < z^* < b - \epsilon$, we have by convexity of $g_{n'}$

$$\begin{aligned} g_{n'}(s) &\leq \frac{z_{n'} - s}{z_{n'} - a} g_{n'}(a) + \frac{s - a}{z_{n'} - a} g_{n'}(z_{n'}) \\ &\leq g_{n'}(a)^+ + \frac{s - a}{z^* - a - \epsilon} g_{n'}(z_{n'}), \end{aligned}$$

whenever $s \in (a, z_{n'}]$, and where $g_{n'}(a)^+ = \min(0, g_{n'}(a))$. Repeating the argument for $s \in [z_{n'}, b)$ we derive

$$\begin{aligned} \limsup_{n' \to \infty} g_{n'}(s) \leq \limsup_{n' \to \infty} & \max \left(g_{n'}^+(a), g_{n'}^+(b) \right) \\ & + \max \left(\frac{s - a}{z^* - a - \epsilon}, \frac{b - s}{b - z^* - \epsilon} \right) \limsup_{n' \to \infty} g_{n'}(z_{n'}) \\ = -\infty, & \end{aligned}$$

In particular, we conclude that $g(s) := \liminf_{n \to \infty} g_n(s) = -\infty$. The case for $z^* = b$ (or $z^* = a$) is similar, using the fact that $s \in (a, z_n')$ for large enough n'.

Next, suppose otherwise, i.e. $\limsup_{n \to \infty} \|g_n\|_\infty < \infty$. Recall from classical convex analysis on $\mathbb{R}^d$ (e.g. similar to [EG15, Theorem 6.7]) that bounded convex functions on convex sets O are uniformly Lipschitz on certain subsets $A \subsetneq O$ with $d(A, O^c) > 0$. In particular, when $f : [a, b] \to \mathbb{R}$ is convex and bounded, $a + \delta \leq c < d \leq b - \delta$, for some $\delta > 0$, the Lipschitz constant of f on the interval $[c, d]$ is bounded by

$$K := 2\delta^{-1} \|f\|_\infty.$$

Since pointwise limits or pointwise minima of K-Lipschitz functions are also K-Lipschitz, and

$$g(z) = \lim_{n \to \infty} \left(\lim_{m \to \infty} \min_{n \leq l \leq m} g_l(z) \right),$$

it follows that g is K-Lipschitz as well. Since c, d are arbitrary, we conclude that g is continuous on (a, b). □

Lemma A.4. *Let $(\phi_n)_{n \in \mathbb{N}}$ be a sequence of convex functions on V and $\{(y_{\lambda,n})\}_\lambda$ a family of sequences and (x_n) a sequence in V. Suppose there exists a $\gamma > 1$ such that*

$$\limsup_{n \to \infty} \phi_n\big(\gamma y_{\lambda,n}\big) < \infty, \qquad \textit{for all } \lambda > 0, \tag{A.6}$$

and suppose there exists some constant $K \in \mathbb{R}$ such that

$$\limsup_{\lambda \to 0} \limsup_{n \to \infty} \phi_n\big(\beta(y_{\lambda,n} - x_n)\big) \leq K \qquad \textit{for all } \beta \in \mathbb{R}. \tag{A.7}$$

Then for any $0 \le \gamma' < \gamma$,

$$\limsup_{n\to\infty} \phi_n(\gamma' x_n) < \infty, \tag{A.8}$$

and

$$\lim_{\lambda\to 0} \limsup_{n\to\infty} \phi_n\big(y_{\lambda,n}\big) = \limsup_{n\to\infty} \phi_n\big(x_n\big), \tag{A.9a}$$

$$\lim_{\lambda\to 0} \liminf_{n\to\infty} \phi_n\big(y_{\lambda,n}\big) = \liminf_{n\to\infty} \phi_n\big(x_n\big). \tag{A.9b}$$

Proof. First, note that if ϕ is convex, we have for any $x, y \in V$ and $\alpha \in (0,1)$ for which both $\alpha^{-1}(y-x), (1-\alpha)^{-1}x \in D(\phi)$ that $x + y \in D(\phi)$

$$\phi(x+y) \le \alpha\phi\big(\alpha^{-1}x\big) + (1-\alpha)\phi\big((1-\alpha)^{-1}y\big),$$

and therefore,

$$\begin{aligned}\phi(y) &= \phi(y - x + x)\\ &\le \alpha\phi\left(\alpha^{-1}(y-x)\right) + (1-\alpha)\phi\left((1-\alpha)^{-1}x\right).\end{aligned} \tag{A.10}$$

Now, we begin by establishing (A.8). By (A.6) and (A.7) there exists a large enough N^* such that for all $n \ge N^*$, λ, $\beta \in R$, and all $\beta \in \mathbb{R}$ both $\gamma y_{\lambda,n} \in D_n$ and $\beta(x - y_\lambda) \in D_n$. Therefore, for any $\gamma' \in [0,\gamma)$, by (A.10) we have

$$\phi_n(\gamma' x_n) \le (1-\alpha)\phi_n\left((1-\alpha)^{-1}\gamma'(x_n - y_{\lambda,n})\right) + \alpha\phi_n(\gamma y_{\lambda,n}).$$

with $\alpha = \gamma'/\gamma$. Passing to the limes superior in n, we obtain

$$\limsup_{n\to\infty} \phi_n(\gamma' x_n) \le (1-\alpha)\limsup_{n\to\infty} \phi_n\left((1-\alpha)^{-1}\gamma'(x_n - y_{\lambda,n})\right) + \alpha\limsup_{n\to\infty} \phi_n(\gamma y_{\lambda,n}).$$

Note that the left-hand side is independent of $\lambda > 0$, while the second term on the right-hand side is finite for every $\lambda > 0$ by (A.6). Moreover, by (A.7) it follows that the first term on the right-hand side is finite for sufficiently small $\lambda > 0$. Hence, the left-hand side is finite as well and, in particular, we have shown (A.8).

Next, from (A.10) we find for any $\alpha \in (0,1)$

$$\begin{aligned}\phi_n(\alpha x_n) &\le (1-\alpha)\phi_n\left(\frac{\alpha}{1-\alpha}(x_n - y_{\lambda,n})\right) + \alpha\phi_n(y_{\lambda,n}),\\ \phi_n(y_{\lambda,n}) &\le (1-\alpha)\phi_n\left((1-\alpha)^{-1}(y_{\lambda,n} - x_n)\right) + \alpha\phi_n\left(\alpha^{-1}x_n\right).\end{aligned}$$

Taking limits in λ and n, we obtain

$$\limsup_{n\to\infty} \phi_n(\alpha x_n) \le (1-\alpha)K + \alpha \liminf_{\lambda\to 0} \limsup_{n\to\infty} \phi_n(y_{\lambda,n}), \tag{A.11a}$$

$$\limsup_{\lambda\to 0} \limsup_{n\to\infty} \phi_n(y_{\lambda,n}) \le (1-\alpha)K + \alpha \limsup_{n\to\infty} \phi_n\left(\alpha^{-1}x_n\right), \tag{A.11b}$$

where we made use of (A.7). Note that $\bar{g}(z) := \limsup_{n\to\infty} \phi_n(zx_n)$ is bounded from above around $z = 1$, and by Lemma A.3 is either continuous in $(0, \gamma')$ or $\bar{g}(z) \equiv -\infty$ in a neighborhood around $z = 1$. In both cases, passing to the limit $\alpha \to 1$ in (A.11a) and (A.11a), we recover (A.9a).

Similarly, to establish (A.9b), we first note that after repeating the argument of (A.11), we find

$$\liminf_{n\to\infty} \phi_n(\alpha x_n) \le (1-\alpha)K + \alpha \liminf_{\lambda\to 0} \liminf_{n\to\infty} \phi_n(y_{\lambda,n}),$$
$$\limsup_{\lambda\to 0} \liminf_{n\to\infty} \phi_n(y_{\lambda,n}) \le (1-\alpha)K + \alpha \liminf_{n\to\infty} \phi_n\left(\alpha^{-1} x_n\right).$$

Set $g_n(z) := \phi_n\left(zx_n\right)$, $g(z) := \liminf_{n\to\infty} g_n(z)$, and recall that $\bar{g}(z)$ is bounded from above on $[0, \gamma']$. Hence, again applying Lemma A.4 and passing to the limit $z \to 1$ in $g(z)$, and $\alpha \to 1$ above, we conclude the proof. □

Now, as a special case, for when $\phi_n = \phi$, $x_n = x$ and $y_{\lambda,n} = y_\lambda$, we have the following statement.

Corollary A.5. *Let ϕ be a convex function on V. Suppose there is a $\gamma > 1$ such that $\phi(\gamma y_\lambda) < \infty$ for all $\lambda > 0$, and that there exists a constant $K \in \mathbb{R}$ such that*

$$\limsup_{\lambda\to 0} \phi\big(\beta(x - y_\lambda)\big) \le K \qquad \text{for all } \beta \in \mathbb{R}.$$

Then for any $0 < \gamma' < \gamma$, $\phi(\gamma' x) < \infty$, and $\phi(x) = \lim_{\lambda\to 0} \phi(y_\lambda)$.

Together, these results show how to connect $\phi_L(x)$ to $\phi_L(y_\lambda)$, $\phi_n(x_n)$ to $\phi_n(y_{\lambda,n})$, and $\phi_U(x)$ to $\phi_U(y_\lambda)$. Since by assumption $\phi_n(y_{\lambda,n})$ is connected to $\phi_L(y_\lambda)$ and $\phi_U(y_\lambda)$, we derive corresponding statements for (x_n, x). Indeed, we now show how the above results imply Theorem A.1.

Proof of Theorem A.1. Applying Corollary A.5 to both ϕ_L and ϕ_U separately, we obtain from (A.1b) and (A.3b) the convergences

$$\phi_L(x) = \lim_{\lambda\to 0} \phi_L(y_\lambda), \qquad \phi_U(x) = \lim_{\lambda\to 0} \phi_U(y_\lambda).$$

Similarly, from (A.1a) and (A.3a) it follows from Lemma A.4 that

$$\lim_{\lambda\to 0} \limsup_{n\to\infty} \phi_n\big(y_{\lambda,n}\big) = \limsup_{n\to\infty} \phi_n\big(x_n\big),$$
$$\lim_{\lambda\to 0} \liminf_{n\to\infty} \phi_n\big(y_{\lambda,n}\big) = \liminf_{n\to\infty} \phi_n\big(x_n\big).$$

Finally, we use the above and the relationship between $y_{\lambda,n}$ and y_λ assumed in (A.2), to obtain

$$\begin{aligned}\phi_L(x) = \lim_{\lambda\to 0} \phi_L(y_\lambda) &\le \lim_{\lambda\to 0} \liminf_{n\to\infty} \phi_n\big(y_{\lambda,n}\big) \\ &= \liminf_{n\to\infty} \phi_n\big(x_n\big) \le \limsup_{n\to\infty} \phi_n\big(x_n\big) \\ &= \lim_{\lambda\to 0} \limsup_{n\to\infty} \phi_n\big(y_{\lambda,n}\big) \le \lim_{\lambda\to 0} \phi_U\big(y_\lambda\big) = \phi_U\big(x\big).\end{aligned}$$

The boundedness conditions (A.4) follow similarly. □

Appendix B

Variational representation of exponential integrals

Below we will give an extension of the classic variational formulation of exponential integrals. This is exploited in various arguments of Parts I.A and I.B.

Lemma B.1. *Let* $(X, \mathcal{A})$ *be a measurable space and* $V : X \to [0, \infty]$ *be a non-negative* $\mathcal{A}$*-measurable function. Then*

$$\log \int_X e^V \, d\mu = \sup_{\nu \in D} \left\{ \int_X V \, d\nu - R(\nu \| \mu) \right\}, \tag{B.1}$$

where

$$D = \left\{ \nu \in \mathcal{P}(X) \,\middle|\, R(\nu \| \mu) < \infty \right\}.$$

In the case of Polish spaces X, the result also follows directly from [DE97, Proposition 4.5.1], which deals with potentials V that are either bounded from below or from above. Nevertheless, for completeness, we give here a direct and elementary proof.

Proof. We will first prove that for every $\nu \in D$,

$$\int_X V \, d\nu - R(\nu \| \mu) \le \log \int_X e^V \, d\mu, \tag{B.2}$$

and then we will show by approximation of bounded functions that

$$\log \int_X e^V \, d\mu \le \sup_{\nu \in D} \left\{ \int_X V \, d\nu - R(\nu \| \mu) \right\}, \tag{B.3}$$

which together will imply (B.1).

For the first inequality, we take any $\nu \in D$ and assume, without loss of generality, that e^V is integrable with respect to μ. In this case, Young's inequality (in the form $ab \leq e^a - b + b\log b$) and the finiteness of $R(\nu\|\mu) < +\infty$ imply the finiteness of $\int_X V\,d\nu$. Let $\mu_V \in \mathcal{P}(S)$ be defined by

$$\mu_V = \frac{1}{Z_V} e^V \mu \qquad \text{with} \qquad Z_V = \int_X e^V\,d\mu.$$

Now let us rewrite the expression on the left-hand side of (B.2) as follows,

$$\begin{aligned}
\int_X V\,d\nu - R(\nu\|\mu) &= \int_X V\,d\nu - \int_X \log\frac{d\nu}{d\mu}\,d\nu \\
&= \log\int_X e^V\,d\mu + \int_X \log e^V\,d\nu - \log\int_X e^V\,d\mu - \int_X \log\frac{d\nu}{d\mu}\,d\nu \\
&= \log\int_X e^V\,d\mu - \int_X \log\left(\frac{d\nu}{d\mu}\left(\frac{d\mu_V}{d\mu}\right)^{-1}\right)d\nu \\
&= \log\int_X e^V\,d\mu - R(\nu\|\mu_V)
\end{aligned} \tag{B.4}$$

By non-negativity of the entropy $R(\cdot\|\mu_V)$, (B.2) holds for any $\nu \in D$.

To show (B.3), consider the sequence $V_n = \min\{V, n\}$ for $n \in \mathbb{N}$. Note that V_n is a non-decreasing sequence of non-negative bounded functions converging pointwise to V. Since V_n is bounded, we have that $R(\mu_{V_n}\|\mu) < +\infty$ and that $\int_X e^{V_n} d\mu < \infty$, and we define again μ_{V_n} as above. Hence, repeating the argument of (B.4) for V_n it follows that

$$\int_X V_n\,d\nu - R(\nu\|\mu) = \log\int_X e^{V_n}\,d\mu - R(\nu\|\mu_{V_n}).$$

In particular, taking $\nu = \mu_{V_n}$, we get

$$\log\int_X e^{V_n}\,d\mu = \int_X V_n\,d\mu_{V_n} - R(\mu_{V_n}\|\mu) \leq \int_X V\,d\mu_{V_n} - R(\mu_{V_n}\|\mu)$$

Maximizing over n, we conclude (B.3). □

Remark B.2. In the proof for bounded V the equality in B.1 follows by checking that $\nu = \mu_V$ is the unique maximizer (B.1). However, this is not possible for a general unbounded V, even when e^V is integrable with respect to μ. A simple example over $X = [0, 1/2]$ follows from taking $V(x) = -\log x + \alpha\log\log x^{-1} + 1$ with any $1 < \alpha < 2$: for this example, μ_V is well-defined and clearly $R(\mu_V\|\mu_V) = 0$, but both $R(\mu_V\|\mu)$ and $\int V\,d\mu_V$ are infinite. In particular, μ_V does not belong to D and is not a maximizer of (B.1), even though μ_{V_n} is a maximizing sequence.

In contrast, if V satisfies a slightly stronger exponential integrability condition, then $R(\nu\|\mu_V) < \infty$ does imply $R(\nu\|\mu) < \infty$ and $\int V d\nu < \infty$, as we show below. □

Corollary B.3. *Let $V : X \to \overline{\mathbb{R}}$ be a measurable function and $\mu_V = Z_V^{-1} e^V \mu$ with normalization constant Z_V. Further, suppose there exists $\gamma > 1$ such that*

$$\int_X e^{\gamma|V|} d\mu < \infty.$$

Then

$$R(\nu\|\mu_V) < \infty \iff R(\nu\|\mu) < \infty.$$

Moreover, the following equality holds:

$$R(\nu\|\mu_V) = R(\nu\|\mu) - \int_X V d\nu + \log \int_X e^V d\mu. \tag{B.5}$$

Proof. Repeating carefully the proof of (B.4), one gets formula (B.5) if V is in $L^1(\nu)$ and one of the two conditions $R(\nu\|\mu) < \infty$ and $R(\nu\|\mu_V) < \infty$ holds. Hence it remains to show that each of the conditions $R(\nu\|\mu) < \infty$ and $R(\nu\|\mu_V) < \infty$ implies that V is in $L^1(\nu)$.

In the case $R(\nu\|\mu) < \infty$, by (B.2) we have

$$\int_X |V| \, d\nu \leq R(\nu\|\mu) + \log \int_X e^{|V|} d\mu,$$

which is finite by assumption.

In the case $R(\nu\|\mu_V) < \infty$, we apply again (B.2) but with base measure μ_V instead of μ and $(\gamma - 1)V$ instead of V, getting

$$\begin{aligned}\int_X (\gamma - 1)|V| \, d\nu &\leq R(\nu\|\mu_V) + \log \int_X e^{(\gamma-1)|V|} d\mu_V \\ &\leq R(\nu\|\mu_V) + \log \int_X e^{\gamma|V|} d\mu - \log \int_X e^V d\mu,\end{aligned}$$

which is finite by assumption. The proof is complete. □

Lemma B.4. *Let $F : X^k \to [0, \infty)$, $k \in \mathbb{N}$ be a nonnegative measurable function satisfying*

$$\int_{X^k} \exp(F(x_1, \dots, x_k)) \, d\mu^{\otimes k} < \infty,$$

and $\nu \in \mathcal{P}(S)$ is such that $R(\nu\|\mu) < \infty$. Then

$$\log \int_X \exp\left(\int_{X^{k-1}} F(x, y) \, d\nu^{\otimes k-1}(y) \right) d\mu(x) \leq (k-1) R(\nu\|\mu) + \log \int_{X^k} e^F d\mu^{\otimes k}.$$

Proof. A simple application of Lemma B.1 yields

$$\begin{aligned}
&\log \int_X \exp\left(\int_{X^{k-1}} F(x,y)\, d\nu^{\otimes k-1}(y)\right) d\mu(x) \\
&\qquad = \sup_\rho \left\{ \left\langle \rho, \int_{X^{k-1}} F(x,y)\, d\nu^{\otimes k-1}(y) \right\rangle - R(\rho \| \mu) \right\} \\
&\qquad = \sup_\rho \left\{ \langle \rho \otimes \nu^{\otimes k-1}, F \rangle - R(\rho \otimes \nu^{\otimes k-1} \| \mu \otimes \mu^{\otimes k-1}) \right\} + (k-1) R(\nu \| \mu) \\
&\qquad \le \sup_\sigma \left\{ \langle \sigma, F \rangle - R(\sigma \| \mu^{\otimes k}) \right\} + (k-1) R(\nu \| \mu) \\
&\qquad = \log \int_{X^k} e^F \, d\mu^{\otimes k} + (k-1) R(\nu \| \mu),
\end{aligned}$$

which is the desired estimate. □

Appendix C

On measurability and exponential approximations

C.1 Measurability of integral maps

In this section, we give a measurability result for integral maps. This in particular implies measurability of $\mathcal{E}_V$ and $\mathcal{E}_V^N$ in Chapter 4. In the following, X and Y are Polish spaces, we recall that $\mathcal{P}(Y)$ endowed with the weak topology is also Polish; X, Y and $\mathcal{P}(Y)$ are endowed with their Borel σ-algebras.

Theorem C.1. *Given any Borel function*

$$X \times Y \times \mathcal{P}(Y) \ni (x, y, \mu) \mapsto f(x, y, \mu) \in [-\infty, +\infty],$$

then the set $\{\mu \in \mathcal{P}(Y) \mid f \in L^1(\mu)\}$ *is Borel and the mapping*

$$F_f : X \times \mathcal{P}(Y) \ni (x, \mu) \mapsto 1_{f \in L^1(\mu)} \int_Y f(x, y, \mu)\, d\mu(y)$$

is also Borel.

The case when $f(x, y, \mu) = W(y)$ for some measurable W is a classical question, and treated in great generalization in for example [Bog07, Chapter 8]. However, for simplicity, in our setting we stick to the case of metric spaces, and adapt an argument of [AB06] (Theorem 15.13). First, we provide a generalization of Lemma 7.3.12 of [DZ10].

Lemma C.2. *Suppose that f is in $C_b(X \times Y \times \mathcal{P}(Y))$. Then F_f is in $C_b(X \times \mathcal{P}(Y))$.*

Proof. In the following, we denote $Z = X \times Y \times \mathcal{P}(Y)$. The idea of the proof is similar to the fact that for Polish spaces X and Y, the set of functions $\{f(x)g(y) \mid f \in$

$C_b(X), g \in C_b(Y)\}$ is convergence determining for $\mathcal{P}(X \times Y)$ (see for example Theorem 3.4.5b of [EK05]), which is used in Lemma 7.3.12 of [DZ10]. Namely, first note for any $g(x, y, \mu) := g_1(x)g_2(y)g_3(\mu)$, with $g_1 \in C_b(X)$, $g_2 \in C_b(Y)$ and $g_3 \in C_b(\mathcal{P}(Y))$, the boundedness and continuity of F_g is trivial.

Now, consider a sequence $(x_n, \mu_n)_n$ converging to (x^*, μ^*). In particular, the subsets $L := \{x_n\} \cup \{x^*\} \subset X$ and $M := \{\mu_n\}_{n\geq 1} \cup \{\mu^*\} \subset \mathcal{P}(Y)$ are compact in X and $\mathcal{P}(Y)$ respectively. In particular, the set $M \subset \mathcal{P}(Y)$ is tight, i.e., for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset Y$ such that $\mu^*(K^c) < \epsilon$ and $\mu_n(K_\epsilon^c) < \epsilon$ for all $n \geq 1$.

Therefore, fix any $\epsilon > 0$. By applying Stone-Weierstrass on the compact set $B_\epsilon = L \times K_\epsilon \times M$, we find a sequence $(g^{\epsilon,l})_{l\geq 1} \in C_b(B_\epsilon)$ with $g^{\epsilon,l}(x, y, \mu) := g_1^{\epsilon,l}(x)g_2^{\epsilon,l}(y)g_3^{\epsilon,l}(\mu)$, where $g_1^{\epsilon,l} \in C_b(L)$, $g_2^{\epsilon,l} \in C_b(K_\epsilon)$ and $g_3^{\epsilon,l} \in C_b(M)$, such that

$$\lim_{l\to\infty} \|f - g^{\epsilon,l}\|_{C_b(B_\epsilon)} = 0.$$

Now also fix $l \in \mathbb{N}$. By Tietze's extension theorem we find a $\tilde{g}_2^{\epsilon,l} \in C_b(Y)$ such that

$$\tilde{g}_2^{\epsilon,l}(y) = g_2^{\epsilon,l}(y) \text{ on } K_\epsilon, \qquad \|\tilde{g}_2^{\epsilon,l}\|_{C_b(Y)} \leq \|g_2^{\epsilon,l}\|_{C_b(K_\epsilon)}.$$

Hence, define on $\tilde{B} = L\times Y\times M$ the continuous function $\tilde{g}^{\epsilon,l}(x, y, \mu) := g_1^{\epsilon,l}(x)\tilde{g}_2^{\epsilon,l}(y)g_3^{\epsilon,l}(\mu)$, and note that by construction $\|\tilde{g}^{\epsilon,l}\|_{C_b(\tilde{B})} \leq \|g^{\epsilon,l}\|_{C_b(B_\epsilon)}$.

We now compute for any $l, n \geq 1$,

$$\begin{aligned}
|F_f(x^*, \mu^*) - F_f(x_n, \mu_n)| &= \left|\int_Y f(x^*, y, \mu^*)\, d\mu^*(y) - \int_Y f(x_n, y, \mu_n)\, d\mu_n(y)\right| \\
&\leq \left|\int_{K_\epsilon} f(x^*, y, \mu^*)\, d\mu^*(y) - \int_{K_\epsilon} f(x_n, y, \mu_n)\, d\mu_n(y)\right| + 2\epsilon\|f\|_{C_b(Z)} \\
&\leq \left|\int_{K_\epsilon} g^{\epsilon,l}(x^*, y, \mu^*)\, d\mu^*(y) - \int_{K_\epsilon} g^{\epsilon,l}(x_n, y, \mu_n)\, d\mu_n(y)\right| \\
&\qquad + 2\|f - g^{\epsilon,l}\|_{C_b(B_\epsilon)} + 2\epsilon\|f\|_{C_b(Z)} \\
&\leq \left|\int_Y \tilde{g}^{\epsilon,l}(x^*, y, \mu^*)\, d\mu^*(y) - \int_Y \tilde{g}^{\epsilon,l}(x_n, y, \mu_n)\, d\mu_n(y)\right| \\
&\qquad + 2\epsilon\|g^{\epsilon,l}\|_{C_b(B_\epsilon)} + 2\|f - g^{\epsilon,l}\|_{C_b(B_\epsilon)} + 2\epsilon\|f\|_{C_b(Z)}.
\end{aligned}$$

Thus, by the continuity of $F_{\tilde{g}^{\epsilon,l}}$ on $\tilde{B}$,

$$\limsup_{n\to\infty} |F_f(x^*, \mu^*) - F_f(x_n, \mu_n)| \leq 2\epsilon\|g^{\epsilon,l}\|_{C_b(B_\epsilon)} + 2\|f - g^{\epsilon,l}\|_{C_b(B_\epsilon)} + 2\epsilon\|f\|_{C_b(Z)}.$$

Taking subsequent limits, first in $l \to \infty$ and then in $\epsilon \to 0$ we conclude the proof. □

Next, we paraphrase Theorem 4.33 of [AB06].

Theorem C.3 (Monotone class theorem)**.** *Let X be a metrizable space, and let $\mathcal{F}$ be a vector subspace of $B_b(X)$ including $C_b(X)$. Then $\mathcal{F} = B_b(X)$ if and only if $\mathcal{F}$ is closed under monotone sequential pointwise limits in $B_b(X)$.*

Proof of theorem C.1. Let $\mathcal{F}$ be the set of bounded Borel-measurable functions given by

$$\mathcal{F} = \left\{ f \in B_b(X \times Y \times \mathcal{P}(Y)) \;\middle|\; F_f : X \times \mathcal{P}(Y) \to \mathbb{R} \text{ is Borel-measurable} \right\}.$$

It is clear that $\mathcal{F}$ is a vector subspace of $B_b(X \times Y \times \mathcal{P}(Y))$, and by Lemma C.2 it contains $C_b(X \times Y \times \mathcal{P}(Y))$. To show stability under monotone pointwise limit, consider any sequence f_n in $\mathcal{F}$ with $f_n \uparrow f$ and f in $B_b(X \times Y \times \mathcal{P}(Y))$. For any measure $\mu \in \mathcal{P}(Y)$, it follows from monotone convergence that

$$\lim_{n\to\infty} F_{f_n}(x,\mu) = \lim_{n\to\infty} \int_Y f_n(x,y,\mu)\, d\mu(y) = \int_Y f(x,y,\mu)\, d\mu(y) = F_f(x,\mu).$$

Hence by the monotone class theorem, we get that $\mathcal{F} = B_b(X \times Y \times \mathcal{P}(Y))$, that is F_f is Borel for any Borel bounded function f.

For a Borel non-negative function f, it follows from approximation with bounded functions and monotone convergence that

$$(x,\mu) \mapsto \int_Y f(x,y,\mu)\, d\mu(y)$$

is Borel and, in particular, the set $\{\mu \mid f \in L^1(\mu)\}$ is Borel. The case of a general Borel function f follows splitting f into f^+ and f^-. The proof is complete. □

Corollary C.4. *Let S be a Polish space and let $V : S^k \to \overline{\mathbb{R}}$ be Borel-measurable. Then both $\mathcal{E}_V, \mathcal{E}_V^N : \mathcal{P}(S) \to \overline{\mathbb{R}}$ are also Borel-measurable.*

C.2 Exponential approximation theorem in Polish spaces

Theorem C.5. *Suppose X is Polish, μ is a finite Borel measure, and $V : X \to \overline{\mathbb{R}}$ is Borel measurable such that*

$$\int_X e^{\beta|V|}\, d\mu < +\infty \qquad \text{for any } \beta \geq 0. \tag{C.1}$$

Then there exists continuous bounded functions $V_\lambda : X \to \mathbb{R}$ such that

$$\lim_{\lambda\to 0} \log \int_X e^{\beta|V - V_\lambda|}\, d\mu = 0 \qquad \text{for any } \beta \geq 0.$$

The argument is similar to the denseness of $C_b(X)$ in $L^p(X)$ spaces, and rely on Tietze's extension theorem and Lusin's Theorem, see for example [AB06, Theorems 2.47 and 12.8] (note that any Polish space X is normal, and so these theorems apply).

Proof of Theorem C.5. First, suppose that V is bounded. Then by Lusin's theorem, for any $\epsilon > 0$, there exists a compact set K_ϵ such that $\mu(X \setminus K_\epsilon) < \epsilon$.

By Tietze's extension theorem, we obtain a continuous function V_ϵ such that $V_\epsilon = V$ on the compact set K_ϵ, and satisfying $\sup_{x\in X} |V_\epsilon(x)| \leq \sup_{x\in X} |V(x)|$. Therefore, for any $\beta \geq 0$,

$$\begin{aligned}\int_X e^{\beta|V-V_\epsilon|}\,d\mu &= \int_{K_\epsilon} e^{\beta|V-V_\epsilon|}\,d\mu + \int_{X\setminus K_\epsilon} e^{\beta|V-V_\epsilon|}\,d\mu \\ &= \int_{K_\epsilon} e^0\,d\mu + \int_{X\setminus K_\epsilon} e^{\beta|V-V_\epsilon|}\,d\mu \leq 1 + \epsilon \cdot e^{2\beta \sup_{x\in X}|V|(x)}.\end{aligned}$$

Note that the choices $\epsilon, K_\epsilon, V_\epsilon$ are independent of β, and thus, by continuity of the logarithm,

$$\lim_{\epsilon\to 0} \log \int_X e^{\beta|V-V_\epsilon|}\,d\mu = 0.$$

Next, for the case of unbounded V, suppose (C.1) holds. Let V_n be the bounded truncation of V to the interval $[-n, n]$, i.e.

$$V_n := \min\{n, \max\{-n, V\}\}, \qquad n \in \mathbb{N}.$$

It is clear that V_n is bounded and $\lim_{n\to\infty} |V - V_n|(x) \to 0$ pointwise in $x \in X$. Moreover, $|V - V_n| \leq |V|$ for all $n \in \mathbb{N}$. Hence, by the assumption on V and the dominated convergence,

$$\lim_{n\to\infty} \int_X e^{\beta|V-V_n|}\,d\mu = 1 \qquad \text{for any } \beta \geq 0,$$

which implies that

$$\lim_{n\to\infty} \log \int_X e^{\beta|V-V_n|}\,d\mu = 0 \qquad \text{for any } \beta \geq 0.$$

From the previous argument on bounded functions, we obtain, for each $n \in \mathbb{N}$, a sequence of bounded continuous functions $(V_{n,\epsilon})_{\epsilon>0}$ such that

$$\lim_{\epsilon\to 0} \log \int_X e^{\beta|V_n-V_{n,\epsilon}|}\,d\mu = 0 \qquad \text{for any } \beta \geq 0.$$

Finally, since by convexity,

$$\log \int_X e^{\beta |V - V_{n,\epsilon}|} \, d\mu \leq \frac{1}{2} \log \int_X e^{2\beta |V - V_n|} \, d\mu + \frac{1}{2} \log \int_X e^{2\beta |V_n - V_{n,\epsilon}|} \, d\mu,$$

we can find an appropriate sequence $V_\lambda := V_{n(\lambda),\epsilon(\lambda)}$ such that

$$\lim_{\lambda \to 0} \log \int_X e^{\beta |V - V_\lambda|} \, d\mu = 0 \qquad \text{for any } \beta \geq 0,$$

thereby concluding the proof. □

Appendix D

Stochastic estimates

D.1 Basic facts on the Girsanov theorem and the Novikov condition

Theorem D.1 (Girsanov)**.** *Let W be an m-dimensional Brownian motion on a filtered probability space (satisfying the standard assumption), with general initial law* Law(W_0)*, let $b : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ be a Borel function. Consider the SDE*

$$\begin{aligned} dX_t &= b(t, X_t)\, dt + dW_t, \qquad t \in [0,T], \\ \mathrm{Law}(X_0) &= \mathrm{Law}(W_0). \end{aligned} \tag{D.1}$$

If

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |b(t,W_t)|^2\, dt} \right] < \infty, \tag{D.2}$$

then there exists a weak solution to (D.1)*. Its law $\tilde{P}^X$ is equivalent to the Wiener measure $\mathbb{W}$ (starting with the same initial law of X_0) and satisfies,*

$$\frac{d\tilde{P}^X}{d\mathbb{W}}(W) := \exp\left(-\frac{1}{2}\int_0^T |b(t,W_t)|^2\, dt + \int_0^T b(t,W_t) \cdot dW_t \right).$$

Moreover, if Y is another weak solution to (D.1) *(defined possibly on another probability space satisfying the standard assumption), with law $\tilde{P}^Y$, such that*

$$\int_0^T |b(t,Y_t)|^2 dt < \infty \quad \tilde{P}^Y\text{-a.s.}, \tag{D.3}$$

then $\tilde{P}^Y$ coincides with $\tilde{P}^X$.

This result is classical, here we recall uniqueness, in the line of [Fed09], Section 3.

Proof. Existence and the representation formula are a classical consequence of Girsanov's theorem, which can be applied thanks to Novikov's condition (D.2). When the initial distribution ρ_0 is a Dirac delta, uniqueness follows from [LS01, Theorem 7.7]. Uniqueness in the general case follows from conditioning X_0 to be a single point. □

Lemma D.2. *Let W be an m-dimensional Brownian motion and let $b : \mathbb{R}^m \to \mathbb{R}$ be a Borel function such that*

$$\mathbb{E}\left[e^{2\int_0^T |b(t,W_t)|^2 dt}\right] < \infty.$$

Then

$$\mathbb{E}\left[e^{\int_0^T b(t,W_t)\cdot dW_t}\right] \leq \mathbb{E}\left[e^{2\int_0^T |b(t,W_t)|^2 dt}\right]^{\frac{1}{2}}$$

Proof. The proof is classical, we sketch the idea: By Novikov's criterion, the exponential local martingale

$$\exp\left(-\frac{1}{2}\int_0^T |2b(t,W_t)|^2 dt + \int_0^T (2b)(t,W_t)\cdot dW_t\right)$$

is a true martingale, in particular it has expectation 1. Then it is enough to apply Hölder inequality to get the required estimate. □

D.2 $L_t^q(L_x^p)$-estimates

Khasminskii's Lemma is classical, see for example [Has59], [Szn98, Chapter 1, Lemma 2.1], [FF11, Lemma 13].

Lemma D.3 (Khasminskii's Lemma)**.** *Let W be a d-dimensional Brownian motion starting from 0, let $f : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ be a non-negative Borel function and assume that*

$$\alpha_f := \sup_{x\in\mathbb{R}^d} \mathbb{E}\int_0^T f(t,x+W_t)\,dt < 1.$$

Then it holds

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}\left[\exp\left[\int_0^T f(t,x+W_t)\,dt\right]\right] \leq \frac{1}{1-\alpha_f}.$$

Lemma D.4 ($L_t^q(L_x^p)$ estimates)**.** *Let W be a d-dimensional Brownian motion starting from 0. Take $1 \leq p,q \leq \infty$ satisfying*

$$\frac{d}{p} + \frac{2}{q} < 2. \tag{D.4}$$

Then there exists a constant C (depending on p, q, d and T) such that, for every $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ non-negative Borel function,

$$\sup_{x \in \mathbb{R}^d} \int_0^T \mathbb{E}[f(t, x + W_t)]\, dt \le C \|f\|_{L_t^q(L_x^p)}. \tag{D.5}$$

This bound is classical (see e.g. [FF11, Lemma 11]), with an elementary proof that we recall here.

Proof. We use Hölder's inequality applied at t and x fixed for the convolution with the Gaussian density p_t:

$$\mathbb{E}[f(t, x + W_t)] = f_t \star p_t(x) \le \|f_t\|_{L_x^p} \|p_t\|_{L_x^{p'}}$$

We recall that $\|p_t\|_{L_x^{p'}} \le c t^{-d/2p}$ for some constant c depending on p and d (as one can see via the change of variable $x' = t^{-1/2}x$). Therefore, for every x, we get by Hölder's inequality for any $q < \infty$,

$$\int_0^T f(t, x + W_t)\, dt \le c \int_0^T \|f_t\|_{L_x^p} t^{-d/2p}\, dt \le c \|f\|_{L_t^q(L_x^p)} \left(\int_0^T t^{-dq'/2p}\, dt \right)^{1/q'}$$

(note that (D.4) implies $q > 1$, so $q' < \infty$). As for $q = \infty$, we estimate similarly,

$$\int_0^T f(t, x + W_t)\, dt \le c \|f_t\|_{L_t^\infty(L_x^p)} \int_0^T t^{-d/2p}\, dt.$$

Now the assumption on p and q is equivalent to $dq'/2p < 1$ for $q < \infty$ and to $d < 2p$ for $q = \infty$. Hence the time integral of $t^{-dq'/2p}$ is finite. Hence the bound (D.5) holds with $C = c \left(\int_0^T t^{-dq'/2p} dt \right)^{1/q'}$. The proof is complete. □

The previous bound can be easily generalized to the case of k independent Brownian motions, as in the following:

Lemma D.5. *Let $W^1, \dots W^k$ be k independent d-dimensional Brownian motions starting from 0. Take $1 \le p_1, \dots p_k, q \le \infty$ satisfying*

$$\frac{d}{p_1} + \dots \frac{d}{p_k} + \frac{2}{q} < 2.$$

Then there exists a constant C (depending on $p_1, \dots p_k, q, d$ and T) such that, for every $f : [0,T] \times \mathbb{R}^{kd} \to \mathbb{R}$ non-negative Borel function,

$$\sup_{x_1, \dots x_k \in \mathbb{R}^d} \int_0^T \mathbb{E}[f(t, x_1 + W_t^1, \dots x_k + W_t^k)] dt \le C \|f\|_{L_t^q(L_{x_1}^{p_1}(\dots(L_{x_k}^{p_k})\dots))}. \tag{D.6}$$

More generally, one can replace the above right-hand side by $\|f\|_{L_t^q(L_{x_{\sigma(1)}}^{p_1}(\dots(L_{x_{\sigma(k)}}^{p_k})\dots))}$ for any permutation σ of $\{1, \dots k\}$.

Proof. The proof is similar to the previous one. We write

$$\mathbb{E}[f(t, x_1 + W_t^1, \dots x_k + W_t^k)] = f_t \star p_t^{\otimes k}(x_1 \dots x_k)$$

and use Hölder inequality in the x_k variable, to get

$$\mathbb{E}[f(t, x_1 + W_t^1, \dots x_k + W_t^k)] \le c\|f_t(x_1, \dots x_{k-1}, \cdot)\|_{L^{p_k}_{x_k}} t^{-d/2p_k} p_t^{\otimes k}(x_1, \dots x_{k-1}).$$

Then we proceed similarly with the other variables and get

$$\mathbb{E}[f(t, x_1 + W_t^1, \dots x_k + W_t^k)] \le c\|f_t\|_{L^q_t(L^{p_1}_{x_1}(\dots(L^{p_k}_{x_k})\dots))} t^{-d/2p_1} \cdot \dots \cdot t^{-d/2p_k}.$$

We then conclude on (D.6) as in the previous proof, taking

$$C = C_T = c\left(\int_0^T t^{-dq'(1/2p_1 + \dots 1/2p_k)} dt\right)^{1/q'} = c_1 T^{1-1/q-d/(2p_1)-\dots-d/(2p_k)}, \tag{D.7}$$

for some constant $c_1 > 0$ independent of T. The bound for a general permutation σ follows from (D.6) applied to $f(x_{\sigma^{-1}(1)}, \dots x_{\sigma^{-1}(k)})$. □

Finally, we put together the previous bounds to obtain an exponential estimate for $L^q(L^p)$ functions (see [FF11, Corollary 14] for a similar statement).

Lemma D.6. *Let $W^1, \dots W^k$ be k independent d-dimensional Brownian motions starting from* 0*. Take $1 \le p_1, \dots p_k, q \le \infty$ satisfying*

$$\frac{d}{p_1} + \dots \frac{d}{p_k} + \frac{2}{q} < 2. \tag{D.8}$$

Then there exists a constant $c > 0$ (depending on $p_1, \dots p_k, q, T$) such that, for every $f : [0, T] \times \mathbb{R}^{kd} \to \mathbb{R}$ non-negative Borel function with $f \in L^q_t(L^{p_1}_{x_1} \dots (L^{p_k}_{x_k}) \dots)$,

$$\sup_{x_1, \dots x_k \in \mathbb{R}^d} \mathbb{E}\left[\exp\left[\int_0^T f(t, x_1 + W_t^1, \dots x_k + W_t^k)\, dt\right]\right] \le \exp\left[c\left(1 + \|f\|^{1/(1-\alpha)}_{L^q_t(L^{p_1}_{x_1}\dots(L^{p_k}_{x_k})\dots)}\right)\right],$$

with $\alpha = 1 - 1/q - d/(2p_1) - \dots - d/(2p_k)$.

Proof. We take

$$h = \left(2c_1\|f\|_{L^q_t(L^{p_1}_{x_1}\dots(L^{p_k}_{x_k})\dots)}\right)^{-1/(1-\alpha)} \wedge T,$$

and let $t_j = hj \wedge T$, and m the first positive integer with $t_m = T$, in particular

$$m = \left\lceil \frac{T}{h} \right\rceil \le T\left(2c_1\|f\|_{L^q_t(L^{p_1}_{x_1}\dots(L^{p_k}_{x_k})\dots)}\right)^{1/(1-\alpha)} + 1.$$

With this choice of h, we have

$$C_h \sup_{j=0,\dots m-1} \left(\int_{t_j}^{t_{j+1}} \|f_t\|^q_{L^{p_1}_{x_1}(\dots(L^{p_k}_{x_k})\dots)} dt \right)^{1/q} \le C_h \|f\|_{L^q_t(L^{p_1}_{x_1}\dots(L^{p_k}_{x_k})\dots)} \le \frac{1}{2},$$

C_h being the constant in (D.7). As a consequence of Lemma D.5, we have

$$\begin{aligned}
\sup_{j=0,\dots,m-1} \sup_{x_1,\dots x_k \in \mathbb{R}^d} \int_{t_j}^{t_{j+1}} \mathbb{E}[f(t, x_1 + W^1_t - W^1_{t_j}, \dots x_k + W^k_t - W^k_{t_j})]\, dt \\
\le C_h \sup_{j=0,\dots,m-1} \left(\int_{t_j}^{t_{j+1}} \|f_t\|^q_{L^{p_1}_{x_1}(\dots(L^{p_k}_{x_k})\dots)} dt \right)^{1/q} \le \frac{1}{2}.
\end{aligned}$$

Hence, we can apply Lemma 8.3 and get

$$\sup_{j=0,\dots,m-1} \sup_{x_1,\dots x_k \in \mathbb{R}^d} \mathbb{E}\left[\exp\left[\int_{t_j}^{t_{j+1}} f(t, x_1 + W^1_t - W^1_{t_j}, \dots, x_k + W^k_t - W^k_{t_j})\, dt \right]\right] \le 2. \tag{D.9}$$

Now we come back to the bound on the whole time interval $[0, T]$. We split the time integral over $[0, T]$ into the integrals over $[t_j, t_{j+1}]$ and use conditional expectation with respect to $\mathcal{F}_{t_{m-1}}$: we have, for every $x_1, \dots, x_k \in \mathbb{R}^d$,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left[\int_0^T f(t, x_1 + W^1_t, \dots x_k + W^k_t)\, dt \right]\right] \\
&= \mathbb{E}\left[\prod_{j=0}^{m-1} \exp\left[\int_{t_j}^{t_{j+1}} f(t, x_1 + W^1_t, \dots x_k + W^k_t)\, dt \right]\right] \\
&= \mathbb{E}\left[\prod_{j=0}^{m-2} \exp\left[\int_{t_j}^{t_{j+1}} f(t, x_1 + W^1_t, \dots x_k + W^k_t)\, dt \right] \cdot \right. \\
&\qquad \left. \cdot\, \mathbb{E}\left[\exp\left[\int_{t_{m-1}}^{t_m} f(t, x_1 + W^1_t, \dots x_k + W^k_t)\, dt \right] \middle| \mathcal{F}_{t_{m-1}} \right]\right]
\end{aligned}$$

(all exponentials are ≥ 1 and so the above products make sense and we can use the rule $\mathbb{E}[XY] = \mathbb{E}[X\mathbb{E}[Y \mid \mathcal{F}_s]]$ for X $\mathcal{F}_s$-measurable). Now we apply the Markov property and the bound (D.9) and get

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left[\int_0^T f(t, x_1 + W_t^1, \dots x_k + W_t^k)\, dt\right]\right] \\
&= \mathbb{E}\left[\prod_{j=0}^{m-2} \exp\left[\int_{t_j}^{t_{j+1}} f(t, x_1 + W_t^1, \dots x_k + W_t^k)\, dt\right] \cdot \right.\\
&\quad \left. \cdot \mathbb{E}\left[\exp\left[\int_{t_{m-1}}^{t_m} f(t, y_1 + W_t^1 - W_{t_{m-1}}^1, \dots y_k + W_t^k - W_{t_{m-1}}^k)\, dt\right]\right] \Big|_{y_1 = x_1 + W_{t_{m-1}}^1, \dots, y_k = x_k + W_{t_{m-1}}^k} \right] \\
&\leq 2\mathbb{E}\left[\prod_{j=0}^{m-2} \exp\left[\int_{t_j}^{t_{j+1}} f(t, x_1 + W_t^1, \dots x_k + W_t^k)\, dt\right]\right].
\end{aligned}
$$

Iterating this argument on j, we get finally, for every $x_1, \dots x_k \in \mathbb{R}^d$,

$$
\mathbb{E}\left[\exp\left[\int_0^T f(t, x_1 + W_t^1, \dots x_k + W_t^k) dt\right]\right] \leq 2^m \leq 2^{T(2c_1\|f\|_{L_t^q(L_{x_1}^{p_1} \dots (L_{x_k}^{p_k}) \dots)})^{1/(1-\alpha)} + 1},
$$

which concludes the proof. □

Part II.A

Variational and dissipation structures for jump processes

Chapter 10

Introduction

In the previous part, we discussed large deviation principles on path space for interacting diffusions. We now turn to the realm of jump processes, or to be more precise, to the evolution equation that describes the law of a jump process, which is called the forward Kolmogorov equation (FKE).

We will introduce two variational formulations involving one-way fluxes, where one formulation arises formally from large deviations of many non-interacting copies of the corresponding jump process, and the other depends on a given reference measure π and, in particular, describes the dissipation of the relative entropy with respect to π.

The latter is our main object of study and is a generalization of existing so-called gradient or force structures in three main ways. For one, we consider one-way fluxes instead of net fluxes, which greatly simplifies the framework and incorporates more information. Secondly, we work on arbitrary or locally compact Polish spaces and with finite but possibly unbounded jump kernels. And finally, we let π be an arbitrary measure instead of an invariant measure of the underlying system.

Therefore, we do not consider gradient or so-called GENERIC structures, even in a generalized sense, since the entropy is no longer a Lyapunov function. However many of the same characteristics play a role and we will adopt much of the same language. In particular, we will still talk about *energy-dissipation principles* (EDPs), consisting of the entropy as a driving energy functional, and a dissipation functional relating the entropy and the forward Kolmogorov equation.

It should be said that removing the assumption that π is an invariant measure is slightly unconventional since it is precisely the reason why EDPs are studied in the first place, and why it is connected to various questions in statistical physics.

However, this is all to set the stage for Part II.B, where we will consider large-population limits of interacting particle systems involving jump, birth, and death. Here our main use of an entropy function is to control its convergence, not its decay.

With those applications in mind, we consider several assumptions on the jump kernel: bounded kernels; unbounded kernels but bounded fluxes; and unbounded fluxes, where we only consider the case of detailed balance.

In all cases, we will show that the corresponding *EDP-functional* is equal to the expected rate functional from large deviation theory, and that both have the unique solution to the FKE as null-minimizers.

10.1 FKE and fluxes

Let $\mathcal{X}$ be a Polish space and $\kappa(x, \mathrm{d}y)$ a finite and measurable jump kernel, i.e. a Borel family $\{\kappa(x,\cdot)\}_{x\in\mathcal{X}} \subset \mathcal{M}^+(\mathcal{X})$, with $\mathcal{M}^+(\mathcal{X})$ the space of finite nonnegative measures over $\mathcal{X}$. The forward Kolmogorov equation then reads as follows

$$\partial_t \rho_t = Q^* \rho_t, \qquad \rho_t \in \mathcal{P}(\mathcal{X}), \tag{FKE}$$

where $\mathcal{P}(\mathcal{X})$ is the space of probability measures over $\mathcal{X}$, and Q^* is the dual of the operator Q, which under suitable assumptions is the generator of the associated process $(X_t)_{t\in[0,T]} \subset \mathcal{X}$, with

$$(Qf)(x) = \int_{\mathcal{X}} (f(y) - f(x))\kappa(x, \mathrm{d}y), \qquad f \in \mathbb{F},$$

for a family of measurable functions $\mathbb{F}$ that will be specified below and depends on further assumptions on κ. Alternatively, if all the terms involved are bounded, we have

$$(Q^*\rho)(\mathrm{d}x) = -\left(\int_{y\in\mathcal{X}} \kappa(x, \mathrm{d}y)\right)\rho(\mathrm{d}x) + \int_{y\in\mathcal{X}} \rho(\mathrm{d}y)\kappa(y, \mathrm{d}x), \qquad \rho \in \mathcal{P}(\mathcal{X}).$$

In general, we will not consider the corresponding stochastic process or its existence, but take (FKE) as a starting point, defined in a suitable weak form. However, we will discuss existence at several points and still will refer to the simple case of $\mathcal{X}$ finite for illustratory purposes.

To discuss variational structures associated with (FKE) we fix a time window $[0, T]$, set the *edge space* $E := \mathcal{X}^2$, and introduce pairs of curves of laws $(\rho_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{X})$ and so-called unidirectional or one-way fluxes $(j_t)_{t\in[0,T]} \subset \mathcal{M}^+_{\mathrm{loc}}(E)$ satisfying the *continuity equation*

$$\partial_t \rho_t + \overline{\nabla} \cdot j_t = 0, \tag{CE}$$

in a suitable weak sense, denoted as $(\rho, j) \in \mathsf{CE}$. Here $\mathcal{M}_{\mathrm{loc}}$ is the space of Radon measures, and for any $f \in B(\mathcal{X})$, $j \in \mathcal{M}^+_{\mathrm{loc}}(E)$ the discrete gradient $\overline{\nabla} f$, and divergence $\overline{\nabla} \cdot j$ are given under suitable integrability constraints by:

$$(\overline{\nabla} f)(x, y) := f(y) - f(x), \qquad \overline{\nabla} \cdot j(\mathrm{d}x) = \int_{y \in \mathcal{X}} j(\mathrm{d}x, \mathrm{d}y) - \int_{y \in \mathcal{X}} j(\mathrm{d}y, \mathrm{d}x).$$

Note that $\overline{\nabla}$ and $-\overline{\nabla}\cdot$ are dual operators, and $\overline{\nabla} \cdot j(\mathrm{d}x) = \int_{y \in \mathcal{X}} j^{\mathrm{net}}(\mathrm{d}x, \mathrm{d}y)$, with $\mathcal{M}(E) \ni j^{\mathrm{net}} := j - \mathsf{s}_{\#} j$ as the *net flux* and the flip function $\mathsf{s} : E \to E$ as $\mathsf{s}(x, y) = (y, x)$.

Finally, while the precise formulation of the continuity equation and solutions to (FKE) will change, in all cases ρ_t will be a solution if and only if $(\rho, \nu_\rho) \in \mathsf{CE}$, with the ρ-dependent measure $\nu_\rho \in \mathcal{M}^+_{\mathrm{loc}}(E)$ given by

$$\nu_\rho(\mathrm{d}x, \mathrm{d}y) := \rho(\mathrm{d}x)\kappa(x, \mathrm{d}y).$$

10.2 Variational formulation and LDPs

A straightforward variational formulation for (FKE) is now obtained via the following functional over curves $(\rho, j) \in \mathsf{CE}$

$$\mathcal{I}(\rho, j) := \int_0^T \mathcal{L}(\rho_t, j_t)\, \mathrm{d}t, \tag{10.1}$$

with Lagrangian $\mathcal{L} : \mathcal{P}(\mathcal{X}) \times \mathcal{M}^+_{\mathrm{loc}}(E) \to [0, +\infty]$ defined as

$$\mathcal{L}(\rho, j) = \mathcal{E}\mathrm{nt}(j | \nu_\rho).$$

Here $\mathcal{E}\mathrm{nt}(\mu|\nu)$ is the relative entropy of μ with respect to ν, given by

$$\mathcal{E}\mathrm{nt}(\mu|\nu) := \begin{cases} \displaystyle\int_Y \phi\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right) \mathrm{d}\nu, & \text{if } \mu \ll \nu, \\ +\infty, & \text{otherwise.} \end{cases} \tag{10.2}$$

where $\phi : [0, +\infty) \to [0, +\infty)$ is the entropy function ϕ, with

$$\phi(z) := z \log z - z + 1.$$

It is clear that $\mathcal{I}$ is only zero if and only if $j_t = \nu_{\rho_t}$ for a.e. $t \in [0, T]$, i.e. when ρ_t is a solution to (FKE).

When $\mathcal{X}$ is a finite set, the functional $\mathcal{I}$ arises from large deviations of the empirical measure of many identical copies of the same jump process. Namely, let L^n_t be the rescaled empirical measure

$$L^n_t(x) := \sum_{i=1}^N \delta_{X^i_t}(x),$$

and W_t^n the *integrated fluxes*:

$$W_t^n(x,y) := \frac{1}{n}\#\Big\{\text{jumps from } x \text{ to } y \text{ in the time-window } [0,t)\Big\}.$$

Moreover, assume that the particles are initially i.d.d. distributed at time $t = 0$ with common measure $\bar{\rho}$. Then one can derive under suitable assumptions, see for example [Ren18, Kra21], that the triple (L_t^n, W_t^n) is a well-defined Markov process and satisfies a large-deviation principle as $n \to \infty$ with rate function $\mathcal{I}(\rho, j)$ in the sense that asymptotically (as $n \to \infty$)

$$\operatorname{Prob}\left(L_t^n \approx \rho_t,\, W_t^n \approx \int_0^t j_s\,\mathrm{d}s,\, \forall t \in [0,T]\right) \asymp e^{-n\left(\mathcal{I}^0(\rho_0)+\mathcal{I}(\rho,j)\right)}$$

where $\mathcal{I}^0(\rho) := \mathcal{E}\mathrm{nt}(\rho|\bar{\rho})$.

Now, even though we do not consider LDPs in our more general setting of $\mathcal{X}$ Polish, we will still use it as our underlying intuition for $\mathcal{I}$, and refer to $\mathcal{I}$ as the *rate functional*, the fluxes $(j_t)_{t\in[0,T]}$ as *observed fluxes* and ν_ρ as *expected flux* (i.e. the expected flux when the particles are distributed according to ρ).

10.3 Energy-dissipation structure

We will now introduce the corresponding EDP-functional associated to a fixed measure $\pi \in \mathcal{M}^+(\mathcal{X})$, which we call the reference measure for the system. While we will discuss the necessary technical assumptions later, let us state our main underlying assumption:

Assumption 10.1. *For π-a.e. x we have $\kappa(x,\cdot) \ll \pi$. Moreover, $\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) \in \mathcal{M}^+(E)$ is finite and there exists a finite jump kernel $\kappa^\dagger$ with*

$$\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) = \pi(\mathrm{d}y)\kappa^\dagger(y,\mathrm{d}x). \tag{10.3}$$

Note that for π-a.e. x we have

$$\kappa^\dagger(x,\mathrm{d}y) = \pi(\mathrm{d}y)\frac{\mathrm{d}\kappa(y,\mathrm{d}x)}{\mathrm{d}\pi(\mathrm{d}x)},$$

and hence depends π-a.e. explicitly on both κ and π. We will therefore refer to $\kappa^\dagger$ as the *π-backward kernel*, or backward kernel for short.

The jump kernel is said to satisfy the *detailed balance condition* with respect to an invariant measure $\pi \in \mathcal{M}^+(\mathcal{X})$ if

$$\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) = \pi(\mathrm{d}y)\kappa(y,\mathrm{d}x). \tag{10.4}$$

Note that the detailed balance condition holds if and only if $\kappa^\dagger = \kappa$.

Remark 10.2. The definition of $\kappa^\dagger$ is motivated by the fact that if π is an invariant measure, i.e. $Q^*\pi = 0$, then under suitable assumptions on $\mathcal{X}$ the time-reversal of the corresponding stationary process, the so-called adjoint or backward process, is itself a Markov jump process with jump kernel $\kappa^\dagger$, also referred to as backward kernel [KJZ19, CL22]. In that case, the generator of the backward process is the $L^2(\pi)$-adjoint of Q. □

Remark 10.3. As will be shown in Section 10.5.5, the Dirichlet form associated to the generator Q and the space $L^2(\pi)$ is a semi-Dirichlet form. If π is an invariant measure it becomes an asymmetric Dirichlet form, in which case two processes are associated to the Dirichlet form if it is regular: the so-called forward process stemming from κ, and the backward process related to the backward kernel $\kappa^\dagger$. If the detailed balance condition holds the Dirichlet form will be symmetric. □

We will now modify the functional (10.22) and its ingredients using the backward kernel $\kappa^\dagger$. Recall $\nu_\rho \in \mathcal{M}^+_{\text{loc}}(E)$, let $\nu^\dagger_\rho \in \mathcal{M}^+_{\text{loc}}(E)$ be defined as

$$\nu_\rho(\mathrm{d}x, \mathrm{d}y) := \rho(\mathrm{d}x)\kappa(x, \mathrm{d}y), \qquad \nu^\dagger_\rho(\mathrm{d}x, \mathrm{d}y) := \rho(\mathrm{d}y)\kappa^\dagger(y, \mathrm{d}x), \tag{10.5}$$

and let $\theta_\rho \in \mathcal{M}^+_{\text{loc}}(E)$ be the geometric average of $\nu_\rho, \nu^\dagger_\rho$ in the sense of (11.1), i.e.

$$\mathrm{d}\theta_\rho := \sqrt{\frac{\mathrm{d}\nu_\rho}{\mathrm{d}\sigma}\frac{\mathrm{d}\nu^\dagger_\rho}{\mathrm{d}\sigma}}\mathrm{d}\sigma, \tag{10.6}$$

for an arbitrary dominating measure σ. Note that we have the symmetry

$$\nu_\pi = \nu^\dagger_\pi. \tag{10.7}$$

We define the dissipation potential $\mathcal{R} : \mathcal{P}(\mathcal{X}) \times \mathcal{M}^+_{\text{loc}}(E) \to [0, +\infty]$,

$$\mathcal{R}(\rho, j) := \mathcal{E}\mathrm{nt}(j|\theta_\rho)$$

and the Fisher information $\mathcal{D} : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$,

$$\mathcal{D}(\rho) := H^2(\nu_\rho, \nu^\dagger_\rho) + \frac{1}{2}\int_E \mathrm{d}(\nu_\rho - \nu^\dagger_\rho),$$

with squared Hellinger distance H^2, see (11.9), and defined only when the second term is finite. Finally, we set the free energy $\mathcal{F} : \mathcal{P}(\mathcal{X}) \to [0, +\infty]$ as one-half of the relative entropy of ρ with respect to π:

$$\mathcal{F}(\rho) := \frac{1}{2}\mathcal{E}\mathrm{nt}(\rho|\pi).$$

The EDP-functional $\mathcal{J}$ for any curve $(\rho, j) \in \mathsf{CE}$ is now given as

$$\mathcal{J}(\rho, j) := \int_0^T (\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t))\,\mathrm{d}t + \mathcal{F}(\rho_T) - \mathcal{F}(\rho_0), \tag{10.8}$$

with the term $\int \mathcal{R}\mathrm{d}t$ denoted as the *action* of a curve.

We will consider solutions satisfying an EDP, called EDP-solutions for short, as those solutions to (FKE) that satisfy the *energy-dissipation balance*

$$\mathcal{E}\mathrm{nt}(\rho_t|\pi) - \mathcal{E}\mathrm{nt}(\rho_0|\pi) = \int_0^t 2\left(\mathcal{R}(\rho_s, \nu_{\rho_s}) + \mathcal{D}(\rho_s)\right)\mathrm{d}s, \text{ for all } t \in [0, T]. \tag{10.9}$$

In particular EDP-solutions are curves such that $\mathcal{J}(\rho, \nu_\rho) = 0$, by taking $t = T$, and they are null-minimizers of $\mathcal{J}$ under the assumption of a suitable chain rule inequality for the entropy. We will show that the reverse implication holds as well, i.e. null-minimizers of $\mathcal{J}$ are EDP-solutions, and that for any ρ the total dissipation is equivalent to

$$2\left(\mathcal{R}(\rho, \nu_\rho) + \mathcal{D}(\rho)\right) = \mathcal{K}(\nu_\rho, \nu_\rho^\dagger)$$

where

$$\mathcal{K}(\nu, \mu) := \mathcal{E}\mathrm{nt}(\nu|\mu) + \int_E \mathrm{d}(\nu - \mu). \tag{10.10}$$

Remark 10.4. One can bound the increase in relative entropy along an EDP-solution as

$$\partial_t \mathcal{E}\mathrm{nt}(\rho_t|\pi) \le \nu_{\rho_t}^\dagger(E) - \nu_{\rho_t}(E),$$

if the latter terms are bounded. Moreover, it is easy to verify that if π is an invariant measure then the two terms cancel, and we recover the fact that in that case $\mathcal{E}(\cdot|\pi)$ is a Lyapunov functional along the solution. □

Remark 10.5. There exist various notions of Fisher information for jump processes, that either relate to the total entropy production or are derived from the associated Dirichlet form, see for example [HPST20]. In this work, we use a natural generalization of the definition introduced in [PRST22], adapted to our setting, and is the unique term such that $\mathcal{R} + \mathcal{D}$ is the same as above and

$$\inf_j \mathcal{R}(\rho, j) = 0.$$

However, as shown in Section 10.5.5 and is clear from the comment below, if $\mathrm{d}\rho = u\mathrm{d}\pi$ and under suitable integrability constraints we have in addition

$$\mathcal{D}(\rho) = \mathcal{E}\left(\sqrt{u}, \sqrt{u}\right)$$

where $\mathcal{E}$ is the semi-Dirichlet form on $L^2(\pi)$ corresponding to the generator Q. □

Remark 10.6. It is straightforward to check that if $\mathrm{d}\rho = u\mathrm{d}\pi$ then θ_ρ and $\mathcal{D}$ reduce to

$$\begin{aligned}\theta_\rho(\mathrm{d}x, \mathrm{d}y) &= \sqrt{u(x)u(y)}\,\nu_\pi(\mathrm{d}x, \mathrm{d}y),\\ \mathcal{D}(\rho) &= \frac{1}{2}\int_E \left(\sqrt{u(y)} - \sqrt{u(x)}\right)^2 \nu_\pi(\mathrm{d}x, \mathrm{d}y) + \frac{1}{2}\int_{\mathcal{X}} u(x)(\overline{\nabla}\cdot\nu_\pi)(\mathrm{d}x)\\ &= \frac{1}{2}\int_E \left(\sqrt{u(y)} - \sqrt{u(x)}\right)^2 \nu_\pi(\mathrm{d}x, \mathrm{d}y) + \frac{1}{2}\int_E (u(x) - u(y))\nu_\pi(\mathrm{d}x, \mathrm{d}y)\\ &= \int_E \left(u(x) - \sqrt{u(x)u(y)}\right)\nu_\pi(\mathrm{d}x, \mathrm{d}y),\end{aligned}$$

if all the terms involved are bounded. In particular, if $\rho \ll \pi$ and π is an invariant measure it follows that

$$\mathcal{D}(\rho) = \frac{1}{2}\int_E \left(\sqrt{u(y)} - \sqrt{u(x)}\right)^2 \nu_\pi(\mathrm{d}x, \mathrm{d}y) \geq 0. \tag{10.11}$$

□

Finally, for technical purposes, let us consider the duals of the dissipation potential $\mathcal{R}$ and Lagrangian $\mathcal{L}$, given by

$$\mathcal{R}^*(\rho, w) := \int_E (e^{w(x,y)} - 1)\theta_\rho(\mathrm{d}x, \mathrm{d}y), \quad \mathcal{H}^*(\rho, w) := \int_E (e^{w(x,y)} - 1)\nu_\rho(\mathrm{d}x, \mathrm{d}y),$$

defined over suitable $\rho \in \mathcal{P}(\mathcal{X})$ and $w \in B(E)$ such that the integrals are bounded from below.

10.4 Main results

So far we have not specified additional assumptions on the kernels κ and $\kappa^\dagger$, or on the specific curves (ρ, j) we consider. In light of the applications in Part II.B, we consider three cases: the kernel κ is bounded (Chapter 12); κ is unbounded but all the fluxes involved are finite (Chapter 13); the fluxes are merely locally finite, and κ is unbounded but satisfies the detailed balance condition (Chapter 14). The three families of assumptions are as follows:

Assumption 10.7 (Bounded kernel)**.** *There exists a constant $c_\kappa < \infty$ such that*

$$\sup_{x\in\mathcal{X}} \int_{\mathcal{X}} \kappa(x, \mathrm{d}y),\ \sup_{x\in\mathcal{X}} \int_{\mathcal{X}} \kappa^\dagger(x, \mathrm{d}y) \leq c_\kappa. \tag{10.12}$$

The curves (ρ, j) are in CE_b, see Definition 12.1, with the continuity equation defined over bounded fluxes in duality with the space of bounded measurable functions $B_b(\mathcal{X})$.

Assumption 10.8 (Unbounded kernel, bounded fluxes)**.** *The curves (ρ, j) are in CE_b^*, see Definition 13.1, i.e. curves $(\rho, j) \in \mathsf{CE}_b$ such that*

$$\sup_{t\in[0,T]} \int_E \rho_t(\mathrm{d}x)\kappa(x,\mathrm{d}y),\ \sup_{t\in[0,T]} \int_E \rho_t(\mathrm{d}x)\kappa^\dagger(x,\mathrm{d}y) < \infty. \tag{10.13}$$

Assumption 10.9 (Unbounded kernel and fluxes, but detailed balance)**.** *The space $\mathcal{X}$ is locally compact, and κ satisfies detailed balance with respect to the invariant measure π. Moreover,*

(B1) there exists a measurable function $a : \mathcal{X} \to (0, 1/2]$, bounded from below on compact sets, with the property that

$$\max\left(\sup_{x\in\mathcal{X}} a(x) \int_{\mathcal{X}} \kappa(x,\mathrm{d}y), \sup_{x\in\mathcal{X}} \int_{\mathcal{X}} a(y)\kappa(x,\mathrm{d}y)\right) =: \tfrac{1}{2} c_{\kappa,a} < \infty, \tag{10.14}$$

(B2) there exists a sequence of compactly supported functions $\{\xi_m\}_{m\in\mathbb{N}}$, such that $0 \le \xi_m \le 1$ for all $m \in \mathbb{N}$, $\xi_m(x) \to 1$ as $m \to +\infty$ for all $x \in \mathcal{X}$.

(B3) The curves (ρ, j) are in CE_a^*, see Definition 14.1, with the continuity equation defined in duality with the space of compactly supported measurable functions $B_c(\mathcal{X})$, and such that

$$\sup_t \int_{\mathcal{X}^2} |\overline{\nabla}\xi_m| \rho_t(\mathrm{d}x)\kappa(x,\mathrm{d}y) < \infty, \qquad \text{for all } x \in \mathcal{X}, m \ge 0. \tag{10.15}$$

We refer to Chapters 12, 13, 14 for more discussion on the relevant assumptions and, in particular, the relation of (10.15) to the existence of adapted Heine–Borel metrics. Still, it should be noted that in Part II.B the mean-field dynamics are an adaptation of the case of a bounded kernel, enhanced with birth and death; the case of bounded fluxes is related to quadratic moments in terms of the mass of the measure-valued process; and the case of detailed balance is related to merely finite entropy and first moment, which in that case are satisfied for all curves with finite $\mathcal{J}$.

With the above assumptions in hand, we can state our main result. It directly follows from Theorems 12.3, 13.2 and 14.3, and is threefold: it establishes the equivalence of the EDP-functional $\mathcal{J}$ and the rate functional $\mathcal{I}$, it characterizes solutions to (FKE) as null minimizers of $\mathcal{J}$, and states that solutions satisfy the energy-dissipation balance.

Theorem 10.10. *Suppose that Assumption 10.1 holds, and that $\kappa, \kappa^\dagger$ and (ρ, j) are such that $\mathcal{F}(\rho_0) < \infty$ and either Assumptions 10.7, 10.8, or 10.9 holds. Then,*

$$\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j). \tag{10.16}$$

In particular, $\mathcal{J}(\rho, j) \geq 0$ *and*

$$\mathcal{J}(\rho, j) = 0 \iff \begin{cases} \rho_t \textit{ is the unique solution to } (\mathsf{FKE}), \\ j_t = \nu_{\rho_t} \textit{ for a.e. } t \in [0, T]. \end{cases} \tag{10.17}$$

Finally, for the solution ρ_t *to* (FKE) *the following energy-dissipation balance holds: for all* $t \in [0, T]$

$$\mathcal{E}\mathrm{nt}(\rho_0|\pi) - \mathcal{E}\mathrm{nt}(\rho_t|\pi) = \int_0^t 2\left(\mathcal{R}(\rho_s, \nu_{\rho_s}) + \mathcal{D}(\rho_s)\right)\,\mathrm{d}s = \int_0^t \mathcal{K}(\nu_{\rho_s}, \nu^\dagger_{\rho_s})\,\mathrm{d}s. \tag{10.18}$$

The precise formulation of solutions to (FKE) and necessary constraints on the initial data will depend on the individual setting and on the existence of appropriate moment functions, but in all three cases uniqueness and existence will hold in a suitable class of solutions.

Remark 10.11. The equivalence $\mathcal{J} = \mathcal{I}$ can be easily verified under the assumption that $t \mapsto \mathcal{F}(\rho_t)$ is differentiable and the curves are sufficiently regular in the sense that $\mathrm{d}\rho_t = u_t \mathrm{d}\pi$ and u_t is uniformly bounded from above and below, and all the terms below are finite. Namely using $\mathrm{d}j_t = g_t \mathrm{d}\nu_\pi$ and the shorthand $u_t = u_t(x)$, $v_t = u_t(y)$, one can write

$$\mathcal{R}(\rho_t, j_t) = \int_E \left(g_t \log \frac{g_t}{\sqrt{u_t v_t}} - g_t + \sqrt{u_t v_t} \right) \mathrm{d}\nu_\pi, \quad \mathcal{D}(\rho_t) = \int_E (u_t - \sqrt{u_t v_t})\,\mathrm{d}\nu_\pi,$$

and

$$\partial_t \mathcal{F}(\rho_t) = \frac{1}{2}\int_E (\log v_t - \log u_t) g_t\,\mathrm{d}\nu_\pi, \quad \mathcal{L}(\rho_t, j_t) = \int_E g_t \log \frac{g_t}{u_t} - g_t + u_t.$$

Moreover, for these curves, the equivalence can also be written in dual form,

$$\begin{aligned} \mathcal{R}^*(\rho, w) &= \mathcal{H}(\rho, w + \overline{\nabla} D\mathcal{F}) - \mathcal{H}(\rho, \overline{\nabla} D\mathcal{F}), & \mathcal{D}(\rho) &= -\mathcal{H}(\rho, \overline{\nabla} D\mathcal{F}), \\ \mathcal{H}(\rho, w) &= \mathcal{R}^*(\rho, w - \overline{\nabla} D\mathcal{F}) - \mathcal{R}^*(\rho, -\overline{\nabla} D\mathcal{F}), & \mathcal{D}(\rho) &= \mathcal{R}^*(\rho, -\overline{\nabla} D\mathcal{F}), \end{aligned}$$

where $D\mathcal{F}(\rho) := \frac{1}{2}\log u$ is the variational derivative of $\mathcal{F}$ with respect to ρ. See also [MPR14, PRS21] for a discussion on such relations for dissipation potentials depending on ρ and either $\dot\rho$ or the net flux.

The technical issue is of course showing that the equivalence indeed holds for *all* curves (ρ, j) satisfying a suitable continuity equation, which we achieve by using both of the above approaches and a sequence of regularized entropies. □

Remark 10.12. Since the solution to (FKE) is characterized as the minimizer of $\mathcal{J}$, we would expect from a minimization approach over small time-windows that

$$\begin{aligned} \partial_t \rho_t + \overline{\nabla} \cdot v_{\rho_t} &= 0 \\ v_{\rho_t} &= (\partial_2 \mathcal{R}^*) \left(\rho_t, -\tfrac{1}{2} \overline{\nabla} D\mathcal{F}(\rho_t) \right). \end{aligned} \tag{10.19}$$

As it turns out, see Proposition 12.14 and Remark 12.18, these relations do indeed hold for the solution for a.e. t. Namely, for such t we have $v_{\rho_t}(\{u(x), u(y) > 0\}^c) = 0$ and hence

$$(\partial_2 \mathcal{R}^*) \left(\rho_t, -\tfrac{1}{2} \overline{\nabla} DS(\rho_t) \right) = 1_{u_t(x), u_t(y) > 0} e^{\frac{1}{2}(\log u_t(x) - \log u_t(y))} \sqrt{u_t(x) u_t(y)} v_\pi = v_{\rho_t}.$$

Apparently, both (10.19) and the balance equation (10.18) are equivalent ways to describe the solution to (FKE). However, similar as for the setting of gradient structures in the case of detailed balance, see for example [LMPR17] for a discussion and historical comments, one can characterize it in additional variational ways, either involving $\mathcal{R}$ or $\mathcal{R}^*$. □

10.5 Discussion and bibliographical notes

In the case where the jump kernel satisfies the detailed balance condition (10.4), various gradient structures and corresponding variational formulations for the net evolution or net flux exist with the entropy $\mathcal{E}\mathrm{nt}(\rho|\pi)$ as a driving functional. This was already observed for discrete Markov chains in [CHLZ12, Maa11, Mie11, Mie13] for a quadratic structure, and subsequently extended to for example Lévy-kernels [Erb14], mean-field interaction [EFLS16] and the Boltzmann equation [Erb16], and limits of reacting particle systems [MM20].

Meanwhile, in trying to duplicate some of the same correspondence between gradient structures and large deviations as for diffusions, and in studying the stability of variational structures, nonlinear or so-called generalized gradient structures for the net evolution were investigated in [Mie16, LMPR17]. Moreover, it was stated that already in the 1930s it was noted that linear force-flux relationships were insufficient for studying systems that were not close to thermal equilibrium, and instead appropriate nonlinear generalizations should be used.

We also refer to [PS22] for a grand overview of how generalized gradient structures involving the hyperbolic cosine can arise, and to [PR19] for how some of them arise as contractions of the functional $\mathcal{I}$. In addition, with particular interest we note the work [PRST22], which considered net fluxes and bounded kernels over Polish spaces.

In the above works on net fluxes for gradient and generalized gradient structures, the fundamental object of study is the EDP-functional $\mathcal{J}_{net}$ over curves (ρ, j^{net}) satisfying a modified continuity equation, with

$$\mathcal{J}_{\mathrm{net}}(\rho, j^{\mathrm{net}}) := \int_0^T \left(\mathcal{R}_{\mathrm{net}}(\rho_t, j_t^{\mathrm{net}}) + \mathcal{D}(\rho_t)\right) + \mathcal{F}(\rho_T) - \mathcal{F}(\rho_0). \tag{10.20}$$

Here $\mathcal{F}$ is the free energy, defined as above, i.e. the relative entropy with respect to π

$$\mathcal{F}(\rho) := \mathcal{E}\mathrm{nt}(\rho|\pi),$$

$\mathcal{R}_{\mathrm{net}}(\rho, j^{\mathrm{net}})$ is the dissipation functional, even in its second argument and with corresponding dual $\mathcal{R}^*_{\mathrm{net}}(\rho, w)$, and the Fisher information $\mathcal{D}$ is the term such that (for suitable curves)

$$\mathcal{D}(\rho) = \mathcal{R}^*_{\mathrm{net}}(\rho, \tfrac{1}{2}\overline{\nabla}\partial_\rho \mathcal{F}). \tag{10.21}$$

Here the additional factor 2 arises due to the definition of net flux used in this thesis. The EDP-functional characterizes the solution $\hat{\rho}_t$ to (FKE), since under appropriate assumptions the latter is the unique null-minimizer of $\mathcal{J}_{\mathrm{net}}$, and

$$\begin{aligned}\partial_t \hat{\rho}_t &= -\tfrac{1}{2}\overline{\nabla} \cdot \hat{j}_t^{\mathrm{net}} \\ \hat{j}_t^{\mathrm{net}} &= \partial_2 \mathcal{R}^*_{\mathrm{net}}(\hat{\rho}_t, \tfrac{1}{2}\overline{\nabla}\partial_\rho \mathcal{F}).\end{aligned}$$

Similar functionals exist when the restriction of detailed balance is lifted, see for example [Ren18, KJZ18], leaving gradient structures for so-called force structures, and are rooted in observations from Macroscopic Fluctuation Theory, see [BDSG+15]. This was generalized to reacting particle systems in [RZ21a], and to (anti)symmetric decompositions with respect to arbitrary quasi-potentials in [PRS21].

The main point of Part II.A is that similar decompositions exist for the one-way flux, and for the case where π is no longer an invariant measure. Already in the author's work [HT22] on variational structures for one-way fluxes under the assumption of detailed balance we employ the functional

$$\int_0^T \left(\mathcal{E}\mathrm{nt}\left(j_t \middle| \sqrt{\nu_{\rho_t}(\mathsf{s}_\# \nu_{\rho_t})}\right) + H^2(\nu_{\rho_t}, \mathsf{s}_\# \nu_{\rho_t})\right) \mathrm{d}t + \frac{1}{2}(\mathcal{E}\mathrm{nt}(\rho_T|\pi) - \mathcal{E}\mathrm{nt}(\rho_0|\pi)), \tag{10.22}$$

with the squared Hellinger distance H^2, and note the equivalence to $\mathcal{I}$ for suitably nice curves. In this thesis, the detailed balance is lifted, and the decomposition is generalized to suitable arbitrary π and κ.

Now, while Theorem 10.10 proves the equivalence of the EDP-functional $\mathcal{J}$ and the rate functional $\mathcal{I}$, and Remark 10.11 already sheds some light on this equivalence, let us provide some additional intuition and relate it to existing structures and inequalities. Moreover, we discuss the generalization to time-dependent π.

10.5.1 Symmetric/antisymmetric splitting of Lagrangian

The equivalence $\mathcal{J} = \mathcal{I}$ can be represented in various ways as shown in Remark 10.11: one is using $\mathcal{R}^*$ and Hamiltonians, or in a direct way as in the proof of Proposition 12.14. The latter can also be seen via a symmetrization argument involving time-reversal and taking adjoints, and is related to forms of detailed balance. Namely, recall that $\mathsf{s}_\# \nu^\dagger_\rho(\mathrm{d}x, \mathrm{d}y) = \rho(\mathrm{d}x)\kappa^\dagger(x, \mathrm{d}y)$, and consider the Lagrangian induced by the backward kernel $\kappa^\dagger$:

$$\mathcal{L}^\dagger(\rho, j) := \mathcal{E}\mathrm{nt}(j|\mathsf{s}_\#\nu^\dagger_\rho).$$

Suppose $(\rho, j) \in \mathsf{CE}$ is such that $u_t := \mathrm{d}\rho_t/\mathrm{d}\pi$ is uniformly bounded from above and below. One can then show that for a.e. $t \in [0, T]$

$$\begin{aligned}
\tfrac{1}{2}\left(\mathcal{L}(\rho_t, j_t) + \mathcal{L}^\dagger(\rho_t, \mathsf{s}_\# j_t)\right) &= \mathcal{E}\mathrm{nt}(j|\theta_{\rho_t}) + H^2(\nu_{\rho_t}, \nu^\dagger_{\rho_t}).\\
\tfrac{1}{2}\left(\mathcal{L}(\rho_t, j_t) - \mathcal{L}^\dagger(\rho_t, \mathsf{s}_\# j_t)\right) &= \tfrac{1}{2}\partial_t \mathcal{E}\mathrm{nt}(\rho_t|\pi) + \tfrac{1}{2}\int_E (\mathrm{d}\nu_{\rho_t} - \mathrm{d}\nu^\dagger_{\rho_t}).
\end{aligned}$$

In particular, if π is an invariant measure and hence κ satisfies so-called semidetailed or complex balance, the final term of the second equation drops out, and

$$\tfrac{1}{2}\left(\mathcal{L}(\rho_t, j_t) - \mathcal{L}^\dagger(\rho_t, \mathsf{s}_\# j_t)\right) = \tfrac{1}{2}\partial_t \mathcal{E}\mathrm{nt}(\rho_t|\pi).$$

This identity is well-known, see for example [Ren18, KJZ18] for a discussion and historical comments on the relation to macroscopic fluctuation theory (see [BDSG+15]), and can be related to time-reversal in the following way. Namely, it implies that

$$\mathcal{E}\mathrm{nt}(\rho_0|\pi) + \int_0^T \mathcal{L}(\rho, j)\,\mathrm{d}t = \mathcal{E}\mathrm{nt}(\rho_T|\pi) + \int_0^T \mathcal{L}^\dagger(\rho_{T-t}, \mathsf{s}_\# j_{T-t})\,\mathrm{d}t.$$

where the final term can be interpreted as following the path (ρ, j) *backwards* and swapping κ by $\kappa^\dagger$, noting that $(\rho_{T-t}, \mathsf{s}_\# j_{T-t}) \in \mathsf{CE}$. See also [CL22] for related equalities involving Schrodinger problems for stationary forward and backward processes.

In the case that ρ_t is a loop, i.e. with $\rho_0 = \rho_T$, we have the equality

$$\int_0^T \mathcal{L}(\rho, j)\,\mathrm{d}t = \int_0^T \mathcal{L}^\dagger(\rho_{T-t}, \mathsf{s}_\# j_{T-t})\,\mathrm{d}t. \tag{10.23}$$

Suppose now for simplicity that in addition $\mathcal{X}$ is finite and π is positive. Since from a large deviation perspective exponentials over integrals over Lagrangians correspond to products of probabilities, the above statement is reminiscent of the fact that semidetailed balance implies that for any collection of states $x_1, \dots, x_n$ with $x_1 = x_n$ we have the equality

$$\kappa(x_1, x_2) \cdots \kappa(x_{n-1}, x_n) = \kappa^\dagger(x_n, x_{n-1}) \cdots \kappa^\dagger(x_2, x_1),$$

since

$$\kappa^\dagger(x, y) = \frac{\pi(y)}{\pi(x)} \kappa(y, x).$$

10.5.2 Contraction to net flux

The EDP-functional $\mathcal{J}$ is given in terms of the one-way fluxes, but this can be turned into a variational formulation in terms of net fluxes via a minimization approach. Namely, using the fact that $\langle w, j \rangle = \frac{1}{2}\langle w, j^{\mathrm{net}} \rangle =$ for antisymmetric $w \in B_b(E)$, i.e w such that $w(y, x) = w(x, y)$, one can show that the dual of

$$\mathcal{R}_{\mathrm{net}}(\rho, j^{\mathrm{net}}) := \inf_{j \in \mathcal{M}^+(E)} \left\{ \mathcal{R}(\rho, j) \ : \ j^{\mathrm{net}} = j - \mathsf{s}_\# j \right\}$$

is given by

$$\begin{aligned} \mathcal{R}^*_{\mathrm{net}}(\rho, w) &= \int_E (e^{2w} - 1) \mathrm{d}\nu_\pi \\ &= \frac{1}{2} \int_E \sqrt{u(x)u(y)} \left((e^{2w(x,y)} - 1)\kappa(x, \mathrm{d}y) + (e^{-2w(x,y)} - 1)\kappa^\dagger(x, \mathrm{d}y) \right) \pi(\mathrm{d}x) \end{aligned}$$

for $\rho \in \mathcal{P}(\mathcal{X})$ with $\mathrm{d}\rho = u \mathrm{d}\pi$.

In the case of detailed balance where $\kappa^\dagger = \kappa$ this can be simplified to

$$\mathcal{R}^*_{\mathrm{net}}(\rho, w) = \frac{1}{2} \int_E \Psi(2w) \, \mathrm{d}\theta_\rho, \quad \mathcal{R}_{\mathrm{net}}(\rho, j^{\mathrm{net}}) = \frac{1}{2} \int_E \Psi \left(\frac{\mathrm{d}j^{\mathrm{net}}}{\mathrm{d}\theta_\rho} \right) \mathrm{d}\theta_\rho$$

where $\Psi^*(z) = e^z + e^{-z} - 2$. This can be shown to be equivalent to the framework for the FKE of [PRST22], after incorporating the fact that the net fluxes in their setting differ by a factor 2 from the net fluxes used in this work.

10.5.3 Force structures

Below we will make an additional decomposition of $\mathcal{R}$ and show how the so-called *force structures* of [Ren18, KJZ18] for the net flux (see also [PRS21] for a generalization to arbitrary quasi-potentials) can be generalized to one-way fluxes.

Assume for simplicity that $\mathcal{X}$ is finite, that π is positive and is an invariant measure, and that κ is bounded from below. In particular, we have the *weak reversibility* condition $\kappa(x,y) > 0 \iff \kappa(y,x) > 0$. Let the *force* $F(\rho) = F_S(\rho) + F_A$ be defined as

$$F = \frac{1}{2}\log\frac{\rho(x)\kappa(x,y)}{\rho(y)\kappa(y,x)}, \quad F_S = -\frac{1}{2}\overline{\nabla}\log u, \quad F_A(x,y) = \frac{1}{2}\log\frac{\pi(x)\kappa(x,y)}{\pi(y)\kappa(y,x)}.$$

It is easily checked that the equivalent force F^* for the backward-kernel $\kappa^\dagger$ corresponding to the adjoint process is equivalent to $F^* = F_S - F_A$, and hence F_S, F_A can be seen as (anti)symmetrizations with respect to taking adjoints. Moreover, note that

$$\frac{1}{2}\left(\mathcal{L}(\rho_t, j_t) - \mathcal{L}(\rho_t, \mathsf{s}_\# j_t)\right) = -\langle F, j_t\rangle.$$

Now, let $\mathcal{R}_S^*$ be the w-symmetrized dual dissipation potential

$$\mathcal{R}_S^*(\rho, w) = \frac{1}{2}\int_E \left(e^{w(x,y)} + e^{w(y,x)} - 2\right)\mathrm{d}\theta_\rho,$$

and $\mathcal{R}_{\mathrm{tsym}}, \mathcal{R}^*_{\mathrm{tsym}}$ the primal/dual dissipation potential one obtains after time symmetrization of $\mathcal{L}$, i.e.

$$\begin{aligned}
\mathcal{R}_{\mathrm{tsym}}(\rho, j) &:= \frac{1}{2}\left(\mathcal{L}(\rho, j) + \mathcal{L}(\rho, \mathsf{s}_\# j)\right) - \inf_{\hat{j}} \frac{1}{2}\left(\mathcal{L}(\rho, \hat{j}) + \mathcal{L}(\rho, \mathsf{s}_\# \hat{j})\right)\\
&= \mathcal{E}\mathrm{nt}(j|\sqrt{\nu_\rho \mathsf{s}_\# \nu_\rho})\\
\mathcal{R}^*_{\mathrm{tsym}}(\rho, w) &:= \int_E (e^w - 1)\sqrt{\nu_\rho \mathsf{s}_\# \nu_\rho})\\
&= \frac{1}{2}\int_E \left(e^{w(x,y)} + e^{w(y,x)} - 2\right)\mathrm{d}\sqrt{\nu_\rho \mathsf{s}_\# \nu_\rho}.
\end{aligned}$$

Note that $\mathcal{R}_{\mathrm{tsym}} \geq 0$ and $\mathcal{R}^*_{\mathrm{tsym}}, \mathcal{R}^*_S$ are nonnegative for antisymmetric w. Moreover, it is straightforward to check that

$$\mathcal{R}(\rho, j) = \mathcal{R}_{\mathrm{tsym}}(\rho, j) - \langle F_A, j\rangle + \mathcal{R}^*_{\mathrm{tsym}}(\rho, F_A), \quad \mathcal{D}(\rho) = \mathcal{R}^*_S(D\mathcal{F}).$$

Together with the fact that $\partial_t \mathcal{F} = -\langle F_S, j\rangle$ and $\mathcal{L} = \mathcal{R} + \partial_t\mathcal{F} + \mathcal{D}$ the above statement is a generalization of [KJZ18, Corollary 4]. Finally, related to similar expressions in [Ren18] for net fluxes, note that we have the relations

$$(\partial_j L)(\rho, j_0) = -F, \qquad (\partial_w H)(\rho, -F) = j_0,$$

where $j_0 := \sqrt{\nu_\rho \mathsf{s}_\# \nu_\rho}$ is the flux that minimizes $\mathcal{R}_{\mathrm{tsym}}$ when $\partial_t \rho = 0$.

10.5.4 Time-dependent π

The framework outlined in this section can easily be generalized to the case where π depends explicitly on time as well. Namely, suppose that $t \mapsto \pi_t \in \mathcal{M}^+(\mathcal{X})$ is differentiable, and

$$\sup_{x\in\mathcal{X}} \int_{\mathcal{X}} \kappa(x, \mathrm{d}y), \quad \sup_{t\in[0,T]} \sup_{x\in\mathrm{supp}(\pi)} \int_{\mathcal{X}} \kappa_t^\dagger(x, \mathrm{d}y) < \infty.$$

where for π_t-a.e. x

$$\kappa_t^\dagger(x, \mathrm{d}y) = \pi_t(\mathrm{d}y)\frac{\mathrm{d}\kappa(y, \mathrm{d}x)}{\mathrm{d}\pi_t(\mathrm{d}x)}.$$

The final condition is for example satisfied if $\mathrm{d}\pi_t/\mathrm{d}\tilde{\pi}$ is uniformly bounded from above and below for a $\tilde{\pi}$ such that κ satisfies the detailed balance condition with respect to $\tilde{\pi}$.

Now let $\nu_{t,\rho}, \mathcal{R}_t, \mathcal{D}_t$, be time-dependent generalizations of their time-independent counterparts. Then one can show that for any curve $(\rho, j) \in \mathsf{CE}$ such that $u_t := \mathrm{d}\rho_t/\mathrm{d}\pi_t$ is uniformly bounded from above and below we have for a.e. $t \in [0, T]$

$$\mathcal{L}(\rho_t, j_t) = \mathcal{R}_t(\rho_t, j_t) + \mathcal{D}_t(\rho_t) + \tfrac{1}{2}\partial_t \mathcal{E}\mathrm{nt}(\rho_t|\pi_t) + \frac{1}{2}\int_{\mathcal{X}} u(x)\,\mathrm{d}(\partial_t \pi_t).$$

Suppose now that π_t itself is a solution to (FKE), i.e. $\partial_t \pi_t = Q^* \pi_t$. Then

$$\mathcal{L}(\rho_t, j_t) = \mathcal{R}_t(\rho_t, j_t) + \frac{1}{2}\int_E \left(\sqrt{u_t(y)} - \sqrt{u_t(x)}\right)^2 \pi_t(\mathrm{d}x)\kappa(x, \mathrm{d}y) + \tfrac{1}{2}\partial_t \mathcal{E}\mathrm{nt}(\rho_t|\pi_t).$$

In particular, we obtain the inequality

$$\mathcal{L}(\rho_t, j_t) \geq \frac{1}{2}\int_E \left(\sqrt{u_t(y)} - \sqrt{u_t(x)}\right)^2 \pi_t(\mathrm{d}x)\kappa(x, \mathrm{d}y) + \tfrac{1}{2}\partial_t \mathcal{E}\mathrm{nt}(\rho_t|\pi_t)$$

This is a special case of the so-called *generalized FIR inequalities* with parameter $\lambda = \frac{1}{2}$, introduced in [HPST20] for jump processes over countable state spaces. Note that one can replicate the case of arbitrary $\lambda \in (0, 1)$ by substituting $\lambda D\mathcal{F}$ for $\frac{1}{2}D\mathcal{F}$ in the dual approach of Remark 10.11.

This also shows that our representation leads to new inequalities for when π_t satisfies the FKE corresponding to a *different* jump kernel, with possible consequences to coarse-graining (e.g. see [HS22]) or quantitative propagation of chaos for Part II.B.

10.5.5 Dirichlet forms

Suppose for simplicity that $\mathcal{X}$ is locally compact and the kernel κ is bounded, and let us define the bilinear form

$$\mathcal{E}(f,g) := -(Qf,g)_\pi = \int_E (f(x)-f(y))g(x)\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) \tag{10.24}$$

It is straightforward to check that $\mathcal{E}$ is a regular semi-Dirichlet form in the sense of [Osh13] with norm

$$\mathcal{E}_\alpha(f,f) := \mathcal{E}(f,f) + \alpha\|f\|_{L^2(\pi)}$$

for any $\alpha \geq c_\kappa$. Namely, lower-boundedness of $\mathcal{E}$ is provided by the identity

$$\begin{aligned}\mathcal{E}(f,f) &= \int_E \Gamma(f)\,\mathrm{d}\pi - \frac{1}{2}\int_{\mathcal{X}} Q(f^2)\mathrm{d}\pi \\ &= \frac{1}{2}\int_E (\overline{\nabla} f)^2\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) + \frac{1}{2}f(x)^2\,\mathrm{d}\overline{\nabla}\cdot(\pi\kappa),\end{aligned} \tag{10.25}$$

with $\Gamma(f) := \frac{1}{2}Q(f^2) - fQf$, and the fact that

$$(\overline{\nabla}\cdot(\pi\kappa))(\mathrm{d}x) \geq -\sup_{x'}\int_{y\in\mathcal{X}} \kappa^\dagger(x',\mathrm{d}y)\pi(\mathrm{d}x).$$

Similarly, for each β,

$$\begin{aligned}\mathcal{E}(f,g) &\leq \frac{\beta}{2}\int (f(y)-f(x))^2\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) + \frac{1}{2\beta}\int g^2(x)\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) \\ &\leq \frac{\beta}{2}\mathcal{E}_\alpha(f,f) + \frac{1}{2\beta}\|g\|_{L^2(\pi)}\sup_x\int_{\mathcal{X}}\kappa(x,\mathrm{d}y),\end{aligned}$$

with the final term being bounded by $\frac{1}{2\beta}\mathcal{E}_\alpha(g,g)$, and hence after optimizing over β we obtain

$$\mathcal{E}(f,g) \leq \sqrt{\mathcal{E}_\alpha(f,f)\mathcal{E}_\alpha(f,f)},$$

and hence the sector condition is satisfied.

Moreover, density in $L^2(\pi)$ implies density w.r.p. $\mathcal{E}_\alpha$ and one can verify that $\mathcal{E}$ is regular, ensuring the existence of a corresponding Hunt process in $\mathcal{X}$, see [Osh13, Theorem 3.3.4].

Now, let $\mathrm{d}\rho = u\mathrm{d}\pi$. It turns out that, similar to the reversible case as in [LMPR17], one can represent the dual dissipation potential and the Fisher information in terms of $\mathcal{E}$:

$$\mathcal{R}^*(\rho,\overline{\nabla} f) = \mathcal{E}\left(\sqrt{u},\sqrt{u}\right) - \mathcal{E}\left(e^f\sqrt{u}, e^{-f}\sqrt{u}\right) \tag{10.26}$$

$$\mathcal{D}(\rho) = \mathcal{E}\left(\sqrt{u}, \sqrt{u}\right) = \int_{\mathcal{X}} \Gamma\left(\sqrt{u}\right) \mathrm{d}\pi + \frac{1}{2}\int_{\mathcal{X}} u(x)\, \mathrm{d}\overline{\nabla}\cdot(\pi\kappa) \tag{10.27}$$

By a dual approach this implies that for any $(\rho, j) \in \mathsf{CE}$

$$\int_0^T \mathcal{R}(\rho_t, j_t)\, \mathrm{d}t \geq \int_0^T \mathcal{R}_{\mathrm{net2}}(\rho_t, \partial_t\rho_t)\, \mathrm{d}t,$$

where the dual dissipation potential of $\mathcal{R}_{\mathrm{net2}}$ (now depending on the net change $\partial_t\rho_t$ instead of net flux j_t^{net}) is given by

$$\mathcal{R}^*_{\mathrm{net2}}(\rho, f) := \mathcal{E}\left(\sqrt{u}, \sqrt{u}\right) - \mathcal{E}\left(e^f\sqrt{u}, e^{-f}\sqrt{u}\right),$$

and that under suitable conditions the solution curve is a null minimizer of

$$\mathcal{J}_{\mathrm{net2}}(\rho) := \int_0^T \left(\mathcal{R}_{\mathrm{net2}}(\rho_t, \partial_t\rho_t) + \mathcal{E}\left(\sqrt{u_t}, \sqrt{u_t}\right)\right) \mathrm{d}t + \frac{1}{2}\left(\mathcal{E}\mathrm{nt}(\rho_T|\pi) - \mathcal{E}\mathrm{nt}(\rho_0|\pi)\right).$$

See also [DO21], where under the assumption of π being an invariant measure and additional locality and orthogonality conditions they consider (pre-)GENERIC variational formulations for a large class of linear Markov processes.

Remark 10.13. Note that one can again let π depend on time. In the case that π_t is sufficiently regular and satisfies $\partial_t\pi_t = Q^*\pi_t$ we have instead the functional

$$\mathcal{J}_{\mathrm{net2}}(\rho) := \int_0^T \left(\mathcal{R}_{\mathrm{net2},t}(\rho_t, \partial_t\rho_t) + \int_{\mathcal{X}} \Gamma\left(\sqrt{u_t}\right) \mathrm{d}\pi_t\right) \mathrm{d}t + \frac{1}{2}\left(\mathcal{E}\mathrm{nt}(\rho_T|\pi_T) - \mathcal{E}\mathrm{nt}(\rho_0|\pi_0)\right).$$

□

Chapter 11

Preliminaries

Below we will collect various definitions, and elementary and preliminary results that will repeatedly be used throughout Part II.

11.1 Topologies on the space of measures

Let Y be a Polish space with corresponding Borel σ-algebra $\mathcal{B}(Y)$. We will denote by $\mathcal{M}(Y)$ the space of signed Borel measures μ with finite total variation $\|\mu\|_{TV}$

$$\|\mu\|_{TV} := \int_Y \mathrm{d}(\mu^+ - \mu^-) = \mu^+(Y) - \mu^-(Y),$$

where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ. We have

$$\|\mu\|_{TV} = \sup_{f \in B_b(Y)} \int_Y f \,\mathrm{d}\mu,$$

with $B_b(Y)$ the space of bounded Borel measurable functions over Y. Moreover, let $\mathcal{M}^+(Y)$ be the space of finite nonnegative Borel measures over Y, $\mathcal{P}(Y)$ the space of probability measures. Finally, we consider the spaces of signed or nonnegative locally finite measures $\mathcal{M}_{\mathrm{loc}}(Y)$ and $\mathcal{M}^+_{\mathrm{loc}}(Y)$, which coincide with the spaces of signed or nonnegative Radon measures.

We consider the following topologies and notions of convergence on these spaces:

1. *setwise* convergence in $\mathcal{M}(Y)$, where $\mu^n \to \mu$ setwise as $n \to \infty$ if

$$\lim_{n\to\infty} \mu^n(A) = \mu(A), \qquad \text{for all } A \in \mathcal{B}(Y),$$

2. the *narrow* topology, the weak topology $\sigma(\mathcal{M}(Y), C_b(Y))$, induced by the space of bounded continuous functions $C_b(Y)$, and the corresponding notion of narrow convergence.

For various equivalent formulations of setwise convergence and related topologies, see [PRST22, Section 2.2]. In particular, setwise convergence and sequentially setwise compact sets coincide with convergence and compact sets for the weak topology $\sigma(\mathcal{M}(Y), B_b(Y))$, induced by the space of Borel measurable and bounded functions $B_b(Y)$.

When in addition Y is locally compact we also use

1. setwise convergence on relatively compact sets in $\mathcal{M}_{\mathrm{loc}}(Y)$,

2. the *vague* topology on $\mathcal{M}_{\mathrm{loc}}(Y)$, the weak $\sigma(\mathcal{M}_{\mathrm{loc}}(Y), C_c(Y))$, induced by the space of continuous functions with compact support $C_c(Y)$, and the corresponding notion of vague convergence.

Note that setwise convergence coincides with convergence in the weak topology $\sigma(\mathcal{M}_{\mathrm{loc}}(Y), B_c(Y))$ induced by the space of Borel measurable functions with compact support $B_c(Y)$. Moreover, setwise convergence implies narrow convergence, and setwise convergence on relatively compact sets implies vague convergence.

The spaces $\mathcal{P}(Y)$ and $\mathcal{M}^+(Y)$ are Polish if equipped with the narrow topology [Bog07, Theorem 8.94], and in the case of Y locally compact the space $\mathcal{M}^+_{\mathrm{loc}}(Y)$ (in fact, even the whole space $\mathcal{M}_{\mathrm{loc}}(Y)$) equipped with the vague topology is Polish as well [Kal17, Theorem 4.1].

11.2 Concave and convex transformations

For any two measures $\mu, \nu \in \mathcal{M}^+(Y)$ (or $\mathcal{M}^+_{\mathrm{loc}}(Y)$) we define their geometric average $\sqrt{\mu\nu} \in \mathcal{M}^+(Y)$ (or $\mathcal{M}^+_{\mathrm{loc}}(Y)$) as the measure

$$\mathrm{d}\sqrt{\mu\nu} := \sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}\nu}{\mathrm{d}\sigma}}\mathrm{d}\sigma, \tag{11.1}$$

for some common dominating measure σ. By 1-homogeneity the definition is independent of the choice of σ, and due to Jensen's inequality and concavity of $(a, b) \mapsto \sqrt{ab}$ we have $\sqrt{\mu\nu} \ll \mu, \nu$ and

$$\sqrt{\mu\nu}(A) \le \sqrt{\mu(A)\nu(A)}, \qquad \text{for all } A \in \mathcal{B}(Y). \tag{11.2}$$

Moreover, we recall the relative entropy $\mathcal{E}\mathrm{nt} : \mathcal{M}^+_{\mathrm{loc}}(Y)^2 \to [0, +\infty]$

$$\mathcal{E}\mathrm{nt}(\mu|\nu) := \begin{cases} \displaystyle\int_Y \phi\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\mathrm{d}\nu, & \text{if } \mu \ll \nu, \\ +\infty, & \text{otherwise.} \end{cases} \tag{11.3}$$

where $\phi : [0, +\infty) \to [0, +\infty)$ is the entropy function ϕ and its Legendre dual ϕ^*

$$\phi(z) := z \log z - z + 1, \qquad \phi^*(s) := e^s - 1. \tag{11.4}$$

In addition, we introduce the monotone relaxation of ϕ

$$\tilde{\phi}(z) := \phi(s \vee 1), \tag{11.5}$$

and the function $\Psi^* : \mathbb{R} \to \mathbb{R}$

$$\Psi^*(z) = e^z + e^{-z} + 1, \tag{11.6}$$

with associated dual $\Psi := (\Psi^*)^*$

$$\Psi(s) = s \log\left(\frac{s + \sqrt{s^2 + 4}}{2}\right) - \sqrt{s^2 + 4} + 2. \tag{11.7}$$

Note that $\tilde{\phi}$ is still convex and superlinear, $0 \le \tilde{\phi} \le \phi$, and $\tilde{\phi}(0) = 0$. Moreover, Ψ, Ψ^* are even, convex and superlinear, and from monotonicity of $\Psi^*(z)/z^2$ on $[0, +\infty)$ it is clear that

$$\Psi^*(az) \le a^2 \Psi^*(z), \qquad \text{for } 0 \le a \le 1,\ z \in \mathbb{R}. \tag{11.8}$$

Finally, the squared Hellinger distance $H^2 : \mathcal{M}^+_{\mathrm{loc}}(Y)^2 \to [0, +\infty]$ is

$$H^2(\mu, \nu) := \frac{1}{2}\int_Y \left(\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\sigma}} - \sqrt{\frac{\mathrm{d}\nu}{\mathrm{d}\sigma}}\right)^2 \mathrm{d}\sigma, \tag{11.9}$$

for any arbitrary dominating measure σ.

Lemma 11.1. *The squared Hellinger distance is jointly convex, and jointly lower semicontinuous on $\mathcal{M}^+(Y)^2$ with respect to setwise and narrow convergence.*

If Y is locally compact it is jointly lower semicontinuous on $\mathcal{M}^+_{\mathrm{loc}}(Y)^2$ with respect to vague convergence and setwise convergence on relatively compact sets.

Proof. Joint convexity and lower semicontinuity with respect to setwise/narrow convergence follow directly from the 1-homogeneity, lower semicontinuity and convexity of the function $(a, b) \mapsto (\sqrt{a} - \sqrt{b})^2 = a - 2\sqrt{ab} + b$, and [PRST22, Lemma 2.2]. For the lower semicontinuity with respect to vague convergence (and hence setwise convergence on relatively compact sets) see [But98, Theorem 3.4.3]. □

Let $\Upsilon : [0, +\infty)^2 \to [0, +\infty]$ be the jointly convex and lower semicontinuous function

$$\Upsilon(w, u, v) := \begin{cases} \phi\left(w/\sqrt{uv}\right)\sqrt{uv}, & \text{if } u, v > 0, \\ \sqrt{uv}, & \text{if } w = 0, \\ +\infty, & \text{if } w > 0 \text{ and either } u = 0 \text{ or } v = 0. \end{cases} \tag{11.10}$$

We then have the following results for the relative entropy.

Lemma 11.2.

1. *For any $\eta, \mu, \nu \in \mathcal{M}^+(Y), \mathcal{M}^+_{\text{loc}}(Y)$ and a common dominating measure σ:*

$$\mathcal{E}\text{nt}(\eta|\sqrt{\mu\nu}) = \int \Upsilon(\mathrm{d}\eta/\mathrm{d}\sigma, \mathrm{d}\mu/\mathrm{d}\sigma, \mathrm{d}\nu)\,\mathrm{d}\sigma \tag{11.11}$$

2. *Both $(\eta, \mu, \nu) \to \mathcal{E}\text{nt}(\eta|\sqrt{\mu\nu})$ and $(\mu, \nu) \to \mathcal{E}\text{nt}(\mu|\nu)$ are jointly convex, and jointly lower semicontinuous on $\mathcal{M}^+(Y)^3, \mathcal{M}^+(Y)^2$ with respect to setwise convergence and narrow convergence.*

   If Y is locally compact they are jointly lower semicontinuous on $\mathcal{M}^+_{\text{loc}}(Y)^3$, $\mathcal{M}^+_{\text{loc}}(Y)^2$ with respect to vague convergence and setwise convergence on relatively compact sets.

3. *For $\mu, \nu \in \mathcal{M}^+(Y)$*

$$\mathcal{E}\text{nt}(\mu|\nu) = \sup_{f \in B_b} \int_Y f\,\mathrm{d}\mu - \int_Y \phi^*(f)\,\mathrm{d}\nu. \tag{11.12}$$

   Moreover, if Y is locally compact and $\mu, \nu \in \mathcal{M}^+_{\text{loc}}(Y)$

$$\mathcal{E}\text{nt}(\mu|\nu) = \sup_{f \in B_c} \int_Y f\,\mathrm{d}\mu - \int_Y \phi^*(f)\,\mathrm{d}\nu. \tag{11.13}$$

4. *Let $\mathcal{K} \subset \mathcal{M}^+(Y)$ be a sequentially compact set with respect to either setwise or narrow convergence. Then the sets*

$$\{\eta : \mathcal{E}\mathrm{nt}(\eta|\sqrt{\mu\nu}) \leq L, \mu, \nu \in \mathcal{K}\}, \{\eta : \mathcal{E}\mathrm{nt}(\eta|\nu) \leq L, \nu \in \mathcal{K}\}$$

are sequentially compact as well.

Similarly, if Y is locally compact and $\mathcal{K} \subset \mathcal{M}^+_{\mathrm{loc}}(Y)$ is a sequentially compact set with respect to either vague convergence or setwise convergence on relatively compact sets, then the above sets are sequentially compact as well.

Proof. The first two properties follow directly from the properties of Υ and [PRST22, Lemma 2.2], [But98, Theorem 3.4.3], with suitable dual formulations for $C_b(Y)$ or $C_c(Y)$ cited or used therein as well, leading to the third property.

We now consider the compactness result. Suppose first that $\mathcal{K} \subset \mathcal{M}^+(Y)$ is sequentially compact with respect to setwise convergence. By [PRST22, Theorem 2.1] a set M is compact if there exists a measure $\sigma \in \mathcal{M}^+(Y)$ such that for all $\mu \in M$

$$\forall \varepsilon > 0\, \exists \delta > 0 : \qquad A \in \mathcal{B}(Y),\, \sigma(A) \leq \delta \implies \mu(A) \leq \varepsilon.$$

Therefore, fixing any $L \geq 0$, it is clear that it is sufficient to show for every $\varepsilon > 0$ that there exists a $\varepsilon' > 0$ such that for all $A \in \mathcal{B}(Y)$ and $\mu, \nu \in \mathcal{K}$

$$\mathcal{E}\mathrm{nt}(\eta|\sqrt{\mu\nu}) \leq L, \max(\mu(A), \nu(A)) \leq \varepsilon' \implies \eta(A) \leq \varepsilon.$$

To show this, fix $\varepsilon > 0$ and let $\varepsilon' > 0$ be such that $\tilde{\phi}(\varepsilon/\varepsilon')\varepsilon' = L$. Note that such ε' can be found for every $\varepsilon > 0$ due to superlinearity of $\tilde{\phi}$. Thus, by Jensen's inequality

$$\begin{aligned}\mathcal{E}\mathrm{nt}(\eta|\sqrt{\mu\nu}) &\geq \phi\left(\frac{\eta(A)}{\sqrt{\mu\nu}(A)}\right)\sqrt{\mu\nu}(A) \geq \tilde{\phi}\left(\frac{\eta(A)}{\sqrt{\mu\nu}(A)}\right)\sqrt{\mu\nu}(A)\\ &\geq \tilde{\phi}\left(\frac{\eta(A)}{\varepsilon'}\right)\varepsilon',\end{aligned}$$

where the final inequality follow from the inequality (11.2) and the fact that $\tilde{\phi}$ is convex and $\tilde{\phi}(0) = 0$. In particular, we obtain by monotonicity of $\tilde{\phi}$ the desired inequality $\eta(A) \leq \varepsilon$.

Similar arguments follow for the other three cases: if $\mathcal{K}$ is sequentially compact with respect to the narrow topology, via Prokhorov's theorem; for setwise convergence on relatively compact sets one can repeat the above argument using the compact exhaustion $\{K_m\}_m$ of the locally compact space Y; and finally, note that for vague convergence we have sequential compactness if and only if $\sup_{\mu\in\mathcal{K}} \mu(K) < \infty$ for every compact set $K \subset Y$.

□

11.3 Chain rule for regular functionals

In proving chain rules for entropy functionals along suitable curves we repeatedly lift from corresponding results for regular functionals, which are of the form

$$F(u) := \int_Y f\zeta(u)\,\mathrm{d}\mu, \tag{11.14}$$

for some fixed $f \in C_b(Y)$, a σ-finite measure μ over Y, and a function $\zeta \in C^1([0,+\infty))$ such that $\|\zeta'\|_\infty < \infty$. Note that, in particular, $\zeta' \in C_b([0,+\infty))$. We then have the following standard result.

Lemma 11.3. *Let $u : t \mapsto L^1(\mu)$ be an absolutely continuous and a.e. differentiable mapping with derivative $\partial_t u_t$. Then*

$$\int_Y f\zeta(u_t)\,\mathrm{d}\mu - \int_Y f\zeta(u_s)\,\mathrm{d}\mu = \int_s^t \int_Y f\zeta'(u_r)\partial_r u_r\,\mathrm{d}\mu\,\mathrm{d}r, \quad \textit{for all } s,t \in [0,T].$$

Proof. It is clear the mapping $t \mapsto \int_Y f\zeta(u_t)\,\mathrm{d}\mu$ is absolutely continuous. Now take any point of differentiability t for the differentiable mapping $t \mapsto u_t$ and consider a sequence t_n such that $t_n \to t$ as $n \to \infty$. Without loss of generality one can assume that as $n \to \infty$ we have for μ-a.e. x the limits $u_{t^n}(x) \to u_t(x)$ and

$$\lim_{n\to\infty} \frac{u_{t^n}(x) - u_t(x)}{t_n - t} = (\partial_t u_t)(x).$$

Approximating $\zeta(u_{t^n}(x))$ by $\zeta(u_t(x)) + (t^n - t)\zeta'(v^n(x))$ for some $v^n(x) \in [u_t, u_{t^n}]$, the result follows from a modified dominated convergence argument and the fact that $\zeta' \in C_b([0,+\infty))$. □

We will apply the above results for appropriate regularizations of ϕ. Namely, for every $n > 0$ we define the regularized entropy functions

$$\phi_n(z) := \int_1^z [\phi']_n(s)\,\mathrm{d}s \tag{11.15}$$

with the truncations $[a]_n := \max(-n, \min(n, a))$. Note that for all n we have $\phi_n \geq 0$, with $\phi'_n = [\phi']_n$ and, in particular, $|\phi'_n| \leq \min(n, |\phi_n|)$. Moreover, $\phi_n \uparrow \phi$ and $\phi'_n \to \phi'$ pointwise. Since $\phi_n \in C^1([0,+\infty))$ and $\|\phi'_n\|_\infty \leq n$ it is clear that $\zeta = \phi_n$ satisfies the necessary assumptions.

For most of our purposes the curve u_t is defined via a continuity equation in some suitable weak way. To still obtain differentiable curves, let us state the following two results.

Lemma 11.4. *Let $g : [0,T] \to L^1(\mu)$ weakly measurable with $t \mapsto \|g_t\|_{L^1(\mu)}$ integrable. Then for a.e. $t \in [0,T]$*

$$\lim_{s\to t} \frac{1}{s-t}\int_t^s g_r \,\mathrm{d}r = g_t \qquad \textit{in } L^1(\mu). \tag{11.16}$$

Note that a sufficient condition for the weak measurability of g is if $g = \mathrm{d}\nu_t/\mathrm{d}\mu$ for some Borel family $(\nu_t)_{t\in[0,T]}$.

Proof. Since Y is Polish the measure space $(Y, \mathcal{B}(Y), \mu)$ is separable and hence $L^1(\mu)$ is separable, see [Coh13, Proposition 3.4.5]. By standard results on Bochner integration g is strongly measurable, and the results follows from the fact that a.e. $t \in [0,T]$ is a Lebesgue point [HNVW16, Theorem 2.3.4]. □

Now consider a curve $u_t : [0,T] \to L^1(\mu)$. We directly have the following.

Lemma 11.5. *Suppose that there exists a weakly measurable family $(g_t)_{t\in[0,T]} \subset L^1(\mu)$ with*

$$u_t - u_s = \int_s^t g_r \,\mathrm{d}r, \qquad \textit{for all } s,t \in [0,T],$$

and the bound

$$\int_0^T \|g_t\|_{L^1(\mu)} < \infty.$$

Then $t \mapsto u_t$ is absolutely continuous and a.e. differentiable in $L^1(\mu)$ with derivative $\partial_t u_t$, and

$$\partial_t u_t = g_t, \qquad \textit{for a.e. } t \in [0,T].$$

Chapter 12

Bounded kernel

In this chapter, we will give a variational formulation for the forward Kolmogorov equation (FKE) involving unidirectional fluxes and the entropy with respect to some arbitrary reference measure π, in the case that the underlying jump kernel and the backward kernel are bounded, see Assumption 10.7. In particular, we assume throughout

$$\sup_{x\in\mathcal{X}} \int_{\mathcal{X}} \kappa(x, \mathrm{d}y) < \infty. \tag{12.1}$$

It can be shown that Q is a bounded operator on $B_b(\mathcal{X})$ and that we indeed can write for $\rho \in \mathcal{P}(\mathcal{X})$,

$$(Q^*\rho)(\mathrm{d}x) = -\overline{\nabla} \cdot \nu_\rho = -\left(\int_{y\in\mathcal{X}} \kappa(x, \mathrm{d}y)\right) \rho(\mathrm{d}x) + \int_{y\in\mathcal{X}} \rho(\mathrm{d}y)\kappa(y, \mathrm{d}x),$$

and that $\nu_\rho = \rho(\mathrm{d}x)\kappa(x, \mathrm{d}y) \in \mathcal{M}^+(E)$ for every $\rho \in \mathcal{P}(\mathcal{X})$. Moreover, solutions to (FKE) in $(\mathcal{P}(\mathcal{X}), \|\cdot\|_{TV})$ exist and are unique for any initial datum $\bar{\rho} \in \mathcal{P}(\mathcal{X})$, as will be shown in Proposition 12.12.

We will consider families of curves (ρ, j) with finite fluxes satisfying the continuity equation with respect to bounded functions, in the following sense.

Definition 12.1 (Continuity equation)**.** A pair $(\rho, j) \in \mathsf{CE}_b$ if

1. the curve $[0,T] \ni t \mapsto \rho_t \in \mathcal{P}(\mathcal{X})$ is absolutely continuous with respect to $\|\cdot\|_{TV}$,
2. the Borel family $(j_t)_{t\in[0,T]} \subset \mathcal{M}^+(E)$ satisfies $\int_0^T \|j_t\|_{TV}\,\mathrm{d}t < \infty$,
3. for every $s, t \in [0,T]$ and all $f \in B_b(\mathcal{T})$

$$\int_{\mathcal{X}} f\,\mathrm{d}\rho_t - \int_{\mathcal{X}} f\,\mathrm{d}\rho_s = \int_s^t \int_E (f(y) - f(x))\,\mathrm{d}j_r\,\mathrm{d}r. \tag{12.2}$$

□

Note that $\overline{\nabla} \cdot j$ is well-defined for any $j \in \mathcal{M}^+(E)$. Moreover, the continuity equation $\partial_t \rho = -\overline{\nabla} \cdot j$ holds in a strong sense if ρ_t is differentiable in $(\mathcal{P}(\mathcal{X}), \|\cdot\|_{TV})$. Various conditions for such differentiability will be stated in Lemma 12.5, and from Proposition 12.12 it will be clear that ρ_t is a solution to (FKE) if and only if $(\rho, \nu_\rho) \in$ CE.

Now, we will assume the backward kernel is bounded as well, i.e. there exists a constant $c_\kappa < \infty$ such that

$$\sup_{x \in \mathcal{X}} \int_{\mathcal{X}} \kappa(x, \mathrm{d}y), \ \sup_{x \in \mathcal{X}} \int_{\mathcal{X}} \kappa^\dagger(x, \mathrm{d}y) \leq c_\kappa. \tag{12.3}$$

Recall that

$$\nu_\rho^\dagger(\mathrm{d}x, \mathrm{d}y) := \rho(\mathrm{d}y)\kappa^\dagger(y, \mathrm{d}x), \quad \theta_\rho := \sqrt{\nu_\rho \nu_\rho^\dagger} \tag{12.4}$$

where the latter is a geometric average in the sense of (11.1). Note that $\nu_\rho(E),\ \nu_\rho^\dagger(E) \leq c_\kappa$ and hence by Jensen's inequality $\theta_\rho(E) \leq c_\kappa$ as well.

Recall the Lagrangian $\mathcal{L}$, dissipation potential $\mathcal{R}$, Fisher information $\mathcal{D}$, free energy $\mathcal{F}$, EDP-functional $\mathcal{J}$ and rate functional $\mathcal{I}$, as defined in Section 10.3. We restrict the domain of $\mathcal{R}$ and $\mathcal{L}$ to $\mathcal{P}(\mathcal{X}) \times \mathcal{M}^+(E)$, and note that the Fisher information is well-defined and bounded, since

$$-c_\kappa \leq \frac{1}{2} \int_E \mathrm{d}(\nu_\rho - \nu_\rho^\dagger) \leq \mathcal{D}(\rho) \leq \frac{1}{2} \int_E \mathrm{d}(3\nu_\rho + \nu_\rho^\dagger) \leq 2c_\kappa$$

Remark 12.2. Moreover, note that all calculations from Remark 10.6 are valid, i.e. if $\mathrm{d}\rho = u\mathrm{d}\pi$ then θ_ρ and $\mathcal{D}$ reduce to

$$\begin{aligned}
\theta_\rho(\mathrm{d}x, \mathrm{d}y) &= \sqrt{u(x)u(y)}\, \nu_\pi(\mathrm{d}x, \mathrm{d}y), \\
\mathcal{D}(\rho) &= \frac{1}{2} \int_E \left(\sqrt{u(y)} - \sqrt{u(x)}\right)^2 \nu_\pi(\mathrm{d}x, \mathrm{d}y) + \int_{\mathcal{X}} u(x)(\overline{\nabla} \cdot \nu_\pi)(\mathrm{d}x) \\
&= \int_E \left(u(x) - \sqrt{u(x)u(y)}\right) \nu_\pi(\mathrm{d}x, \mathrm{d}y).
\end{aligned}$$

In particular, if $\rho \ll \pi$ and π is an invariant measure it follows that

$$\mathcal{D}(\rho) = \frac{1}{2} \int_E \left(\sqrt{u(y)} - \sqrt{u(x)}\right)^2 \nu_\pi(\mathrm{d}x, \mathrm{d}y) \geq 0.$$

□

We can now state the main result of this chapter.

Theorem 12.3. *For any* $(\rho, j) \in \mathsf{CE}$ *with* $\mathcal{F}(\rho_0) < \infty$

$$\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j) \tag{12.5}$$

In particular, $\mathcal{J}(\rho, j) \geq 0$ *and*

$$\mathcal{J}(\rho, j) = 0 \iff \begin{cases} \rho_t \text{ is the unique solution to } (\mathsf{FKE}), \\ j_t = \nu_{\rho_t} \text{ for a.e. } t \in [0, T]. \end{cases} \tag{12.6}$$

Moreover, if $\mathcal{J}(\rho, j) < \infty$ *the following chain rule holds for the free energy* $\mathcal{F}$*, i.e. the map* $t \mapsto \mathcal{F}(\rho_t)$ *is absolutely continuous and a.e. differentiable with*

$$\partial_t \mathcal{F}(\rho_t) = \frac{1}{2} \int_E \overline{\nabla} \phi' \left(u_t\right) \mathrm{d}j_t, \qquad \text{for a.e. } t \in [0, T], \text{ and} \tag{12.7}$$

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t \mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t), \qquad \text{for a.e. } t \in [0, T]. \tag{12.8}$$

Remark 12.4. Note that we indeed have the equality $2(\mathcal{R}(\rho, j) + \mathcal{D}(\rho) = \mathcal{K}(\nu_\rho, \nu_\rho^\dagger)$, leading to the energy balance of Theorem 10.10. Namely, if either one is finite we have $\nu_{\rho_t} \ll \nu_{\rho_t}^\dagger$ for a.e. $t \in [0, T]$, and for such t

$$\begin{aligned} &\frac{1}{2} \partial_t \mathcal{E}\mathrm{nt}(\rho_t) = \mathcal{R}(\rho_t, \nu_{\rho_t}) + \mathcal{D}(\rho_t) \\ &= \int_E \left(\log \left(\frac{\mathrm{d}\nu_{\rho_t}}{\mathrm{d}\sqrt{\nu_{\rho_t}^\dagger \nu_{\rho_t}}} \right) \mathrm{d}\nu_{\rho_t} - \mathrm{d}\nu_{\rho_t} + \mathrm{d}\sqrt{\nu_{\rho_t}^\dagger \nu_{\rho_t}} \right) + \int_E \left(\mathrm{d}\nu_{\rho_t} - \mathrm{d}\sqrt{\nu_{\rho_t}^\dagger \nu_{\rho_t}} \right) \\ &= \frac{1}{2} \int_E \log \left(\frac{\mathrm{d}\nu_{\rho_t}}{\mathrm{d}\nu_{\rho_t}^\dagger} \right) \mathrm{d}\nu_{\rho_t}. \end{aligned}$$

□

For technical purposes, let us also restrict the domains and ranges of the dual dissipation potential $\mathcal{R}^* : \mathcal{P}(\mathcal{X}) \times B_b(E) \to \mathbb{R}$ and the Hamiltonian $\mathcal{H} : \mathcal{P}(\mathcal{X}) \times B_b(E) \to \mathbb{R}$, with natural extensions to functionals over $\mathcal{P}(\mathcal{X}) \times B(E)$ to $[-c_\kappa, +\infty]$.

The proof of Theorem 12.3 is postponed until Section 12.3, where we will prove the chain rule and the equivalence $\mathcal{I} = \mathcal{J}$. First, we establish necessary preliminary estimates under finite $\mathcal{I}, \mathcal{J}$ and spatial/time-regularity of the corresponding curves in Section 12.1, and state the definition, existence, and uniqueness of solutions in Section 12.2. Moreover, even though it is not necessary for our subsequent results, in Section 12.4 we prove a compactness result under finite action and show that the action is lower semicontinuous for the corresponding notion of convergence, in line with similar results of [PRST22].

12.1 Estimates and regularity

We briefly cover the regularity of curves satisfying the continuity equation, or in addition having finite action $\int_0^T \mathcal{R}\,\mathrm{d}t$, or finite rate $\mathcal{I}$, encompassed by the following three Lemmas.

Lemma 12.5. *Let $(\rho, j) \in \mathsf{CE}$. Then*

$$d_{TV}(\rho_s, \rho_t) \le 2 \int_s^t \|j_r\|_{TV}\,\mathrm{d}r, \tag{12.9}$$

and, in particular, there exists a measure $\sigma \in \mathcal{M}^+(\mathcal{X})$ such that $\rho_t \ll \sigma$ for all $t \in [0,T]$.

Moreover, suppose that in addition $j_t \ll \nu_{\rho_t}$ for a.e. $t \in [0,T]$. Then there exists a $\sigma \in \mathcal{M}(\mathcal{X})$ such that $\rho_t \ll \sigma$ for all t and $j_t \ll \nu_\sigma$ for a.e. $t \in [0,T]$, and $t \mapsto v_t := \mathrm{d}\rho_t/\mathrm{d}\sigma$ is an absolutely continuous and a.e. differentiable map in $L^1(\sigma)$.

The continuity equation then holds in a strong sense: $t \mapsto \rho_t \in (\mathcal{P}, \|\cdot\|_{TV})$ is a.e. differentiable with derivative in $\mathcal{M}(\mathcal{X})$ and

$$\partial_t \rho_t = -\overline{\nabla} \cdot j_t, \qquad \text{for a.e. } t \in [0,T]. \tag{12.10}$$

Finally, if $\rho_t \ll \pi$ for a.e. $t \in [0,T]$ one can take $\sigma = \pi$.

Proof. The bound (12.9) follows directly from the continuity equation (12.2) and taking the supremum over f with $\|f\|_\infty \le 1$. Now, take a sequence of times $t_n \in [0,T]$ such that $\{t_n\}_{n\in\mathbb{N}}$ is dense in $[0,T]$, and define

$$\sigma = \frac{1}{2^n} \sum_{n=1} \rho_{t_n}.$$

Clearly $\rho_{t_n} \ll \sigma$ for every $n \in \mathbb{N}$. Thus, by a density argument and the TV-regularity of the map $t \mapsto \rho_t$, we find that $\rho_t \ll \sigma$ for every $t \in [0,T]$.

The common dominating measure σ is then obtained by first constructing a common dominating measure over the countable set $\mathbb{Q} \cap [0,T] \subset [0,T]$, following a density argument using the TV-regularity of ρ_t.

Next, assume that $j_t \ll \nu_{\rho_t}$ for a.e. $t \in [0,T]$. It is straightforward to check that by the TV-regularity of ρ_t the measures $\int_{y\in\mathcal{X}} \kappa(x, \mathrm{d}y)\rho_t(\mathrm{d}x)$ and $\int_{y\in\mathcal{X}} \kappa(y, \mathrm{d}x)\rho(\mathrm{d}y)$ are TV-regular in time as well and, in particular, there exists a common dominating measure σ. Therefore, we find that $\rho_t \ll \sigma$ for all t, and since $j_t \ll \nu_\rho$ for a.e. $t \in [0,T]$ there exist measurable maps $g^\pm : [0,T] \to L^1(\sigma)$ with

$$\int_{y\in\mathcal{X}} j_t(\mathrm{d}x, \mathrm{d}y) = g_t^-(x)\sigma(\mathrm{d}x), \int_{y\in\mathcal{X}} j_t(\mathrm{d}y, \mathrm{d}x) = g_t^+(x)\sigma(\mathrm{d}x), \qquad \text{for a.e. } t \in [0,T].$$

Thus, by the continuity equation and Lemma 11.5, the map $t \mapsto v_t := \rho_t/\mathrm{d}\sigma$ is absolutely continuous and a.e. differentiable in $L^1(\sigma)$, with

$$\partial_t v_t(x) = -g_t^-(x) + g_t^+(y) \qquad \text{for a.e. } t \in [0, T].$$

Integrating the expression by σ we derive (12.10).

Finally, note if $\rho_t \ll \pi$ then clearly $\nu_{\rho_t} \ll \nu_\pi$ and $\mathsf{s}_\# \nu_{\rho_t} \ll \mathsf{s}_\# \nu_\pi = \mathsf{s}_\# \nu_\pi^\dagger$ and we find $\int_{y\in\mathcal{X}} \kappa(x, \mathrm{d}y)\rho_t(\cdot), \int_{y\in\mathcal{X}} \kappa(y, \cdot)\rho_t(\mathrm{d}y) \ll \pi$. □

Lemma 12.6. *Let $\rho \in \mathcal{P}(\mathcal{X})$, $j \in \mathcal{M}^+(E)$, and suppose that either $\mathcal{R}(\rho, j) < \infty$ or $\mathcal{L}(\rho, j) < \infty$. Then $j \ll \nu_\rho$ and*

$$c_\kappa \tilde{\phi}\left(c_\kappa^{-1} \|j\|_{TV}\right) \le \min\left(\mathcal{R}(\rho, j), \mathcal{L}(\rho, j)\right), \tag{12.11}$$

Recall from (11.5) that $\tilde{\phi}(s) = \phi(s \vee 1)$ is the monotone relaxation of ϕ.

Proof. The first statement follows from the fact that $\theta_\rho \ll \nu_\rho, \nu_\rho^\dagger$ and $\mathcal{R}(\rho, j) = \mathcal{E}\mathrm{nt}(j|\theta_\rho)$, $\mathcal{L}(\rho, j) = \mathcal{E}\mathrm{nt}(j|\nu_\rho)$. For the inequality, by Jensen's inequality

$$\mathcal{R}(\rho, j) = \mathcal{E}\mathrm{nt}(j|\theta_\rho) \ge \phi\left(\frac{j(E)}{\theta_\rho(E)}\right)\theta_\rho(E) \ge \tilde{\phi}\left(\frac{j(E)}{\theta_\rho(E)}\right)\theta_\rho(E).$$

Using the bound $\theta_\rho(E) \le c_\kappa$, and the monotonicity and convexity of $\tilde{\phi}$, we find

$$\mathcal{R}(\rho, j) \ge \tilde{\phi}\left(c_\kappa^{-1}\, j(E)\right) c_\kappa,$$

and a similar calculation follows for $\mathcal{L}$. □

Lemma 12.7. *Let $(\rho, j) \in \mathsf{CE}$, and $\rho_t = \rho_t^a + \rho_t^\perp$ the decomposition with respect to π, where $\mathrm{d}\rho_t^a = u_t \mathrm{d}\pi$ and $\rho_t^\perp \perp \pi$. Then*

$$\mathcal{I}(\rho, j) < \infty \implies \qquad t \mapsto \rho_t^\perp(\mathcal{X}) \text{ is decreasing,}$$
$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t < \infty \implies \qquad t \mapsto \rho_t^\perp(\mathcal{X}) \text{ is constant.}$$

We will also sometimes refer to the quantity $\int \mathcal{R}\,\mathrm{d}t$ as the *action* for a curve.

Remark 12.8. In the case where $\mathcal{X}$ is a finite the above statement for $\mathcal{I}$ is related to the fact that supp(π) is an absorbing set for the corresponding Markov chain.

Moreover, the statement for $\mathcal{R}$ follows from the fact that finite action implies $j_t \ll \nu_{\rho_t}, \nu_{\rho_t}^\dagger$ for a.e. $t \in [0, T]$, and in a sense one can then repeat the argument used for $\mathcal{L}$, now both going forward and backward in time, as will be shown below. □

Proof. By Lemma 12.5 there exists a measure σ' dominating ρ_t for all $t \in [0,T]$. Now set $\sigma := \pi + \sigma'$ and $h := \mathrm{d}\pi/\mathrm{d}\sigma$. It is clear that $\mathrm{d}\rho_t^a = 1_{h>0}\mathrm{d}\rho_t$ and $\mathrm{d}\rho_t^\perp = 1_{h=0}\mathrm{d}\rho_t$. Via the continuity equation we can therefore write for all $s,t \in [0,T]$

$$\rho_t^\perp(\mathcal{X}) - \rho_s^\perp(\mathcal{X}) = \int_s^t \int_E (1_{h(y)=0} - 1_{h(x)=0}) j_r(\mathrm{d}x, \mathrm{d}y)\,\mathrm{d}r.$$

However, note that $\nu_{\rho_r}(\{h(x) > 0, h(y) = 0\}) = 0$ since $k(x,\cdot) \ll \pi$ for π-a.e. x, and similarly $\nu_{\rho_r}^\dagger(\{h(x) = 0, h(y) > 0\}) = 0$. Hence if $j_t \ll \nu_{\rho_t}$ for a.e. $t \in [0,T]$ we obtain that $t \mapsto \rho_t^\perp(\mathcal{X})$ is decreasing, and if in addition $j_t \ll \nu_{\rho_t}^\dagger$ for a.e. $t \in [0,T]$ then the map is constant. The result now follows from the fact that finite $\mathcal{I}$ implies $j_t \ll \nu_{\rho_t}$ for a.e. $t \in [0,T]$, and finite action implies $j_t \ll \nu_{\rho_t}, \nu_{\rho_t}^\dagger$ for a.e. $t \in [0,T]$. □

Lemmas 12.5, 12.6, and 12.7 have the following direct consequence.

Corollary 12.9. *Suppose that $(\rho, j) \in \mathsf{CE}$ and either $\mathcal{I}(\rho, j) < \infty$ or*

$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t < \infty.$$

Then $t \mapsto \rho_t$ is a.e. differentiable in $(\mathcal{P}(\mathcal{X}), \|\cdot\|_{TV})$ with derivative in $\mathcal{M}(\mathcal{X})$ and

$$\partial_t \rho_t = -\overline{\nabla} \cdot j_t, \qquad \textit{for a.e. } t \in [0,T].$$

Moreover,

$$\int_0^T c_\kappa \tilde{\Phi}\left(\tfrac{1}{2} c_\kappa^{-1} \|\partial_t \rho_t\|_{TV}\right)\,\mathrm{d}t \le \min\left(\mathcal{I}(\rho, j),\ \int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t\right).$$

Finally, if $\rho_0 \ll \pi$ then $\rho_t \ll \pi$ for all $t \in [0,T]$, and $\partial_t \rho_t \ll \pi$ for a.e. $t \in [0,T]$.

Remark 12.10. Recall that $\tilde{\Phi}$ is convex and superlinear, and hence for a sequence (ρ^n, j^n) such that

$$\limsup_{n\to\infty} \int_0^T \mathcal{R}(\rho_t^n, j_t^n)\,\mathrm{d}t < \infty,$$

it follows that the sequence of curves $\{(\|\partial_t \rho_t^n\|_{TV})_{t\in[0,T]}\}_{n\in\mathbb{N}}$ is equiintegrable, and therefore the sequence of curves $\{(\rho_t^n)_{t\in[0,T]}\}_{n\in\mathbb{N}}$ is equicontinuous with respect to d_{TV}. □

12.2 Solutions

Recall the forward Kolmogorov equation (FKE), written in the following form

$$\partial_t \rho_t = Q^* \rho_t = -\overline{\nabla} \cdot v_{\rho_t}, \qquad \rho_t \in \mathcal{P}(\mathcal{X}),$$

Definition 12.11. A solution to (FKE) is any TV-absolutely continuous and a.e. differentiable mapping $\rho : [0,T] \to (\mathcal{P}(\mathcal{X}), \|\cdot\|_{TV})$ satisfying

$$\partial_t \rho_t(\mathrm{d}x) = -\overline{\nabla} \cdot v_{\rho_t}, \qquad \text{for a.e. } t \in [0,T] \tag{12.12}$$

□

Various classical tools exists for establishing existence and uniqueness of solutions to (FKE) in either TV or suitable L^1 spaces. However, due to the fact that we will consider birth/death in Chapter 16 we follow the usual constructive approach in $(\mathcal{M}^+(\mathcal{X}), \|\cdot\|_{TV})$, using the fact that $-\overline{\nabla} \cdot v_{\rho_t}$ is Lipschitz in ρ.

Proposition 12.12. *For any $\bar{\rho} \in \mathcal{P}(\mathcal{X})$ there exists a unique solution ρ_t with $\rho_0 = \bar{\rho}$. Moreover, ρ_t is a solution if and only if $(\rho, v_\rho) \in$ CE.*

Proof. We will first construct a unique solution in $\mathcal{M}^+(\mathcal{X})$. From the continuity equation it is clear that mass is conserved, establishing that the solution is in $\mathcal{P}(\mathcal{X})$.

Now, fix an arbitrary time-window $[0,T]$, and consider any Borel families $(\lambda_t^+)_{t\in[0,T]} \in \mathcal{M}^+(\mathcal{X})$ and $(c_t)_{t\in[0,T]} \in B_b(\mathcal{X})$ with c_t uniformly bounded, and λ_t such that $\int_0^T \|\lambda_t\|_{TV}\,\mathrm{d}t < \infty$ and a common dominating measure exist. Then a unique nonnegative solution to

$$\partial_t \rho_t(\mathrm{d}x) = \lambda_t(\mathrm{d}x) - c_t(x)\rho_t(\mathrm{d}x)$$

in $\mathcal{M}^+(\mathcal{X})$ exists, and is given by

$$\rho_t := e^{-\int_0^t c_s \mathrm{d}s} \left(\int_0^t \lambda_s e^{\int_0^s c_r \,\mathrm{d}r} \mathrm{d}s + \rho_0 \right).$$

We now set $\rho_t^0 := \bar{\rho}$ for all $t \in [0,T]$ and perform the implicit Picard iteration

$$\partial_t \rho_t^{n+1}(\mathrm{d}x) = \int_{y\in\mathcal{X}} \kappa(y,\mathrm{d}x)\rho_t^n(\mathrm{d}y) - \left(\int_{y\in\mathcal{X}} \kappa(x,\mathrm{d}y) \right) \rho_t^{n+1}(\mathrm{d}x),$$

i.e. $\rho^{n+1} = \mathcal{G}(\rho^n)$ with

$$(\mathcal{G}\rho)_t(\mathrm{d}x) := e^{-t\kappa(x,\mathcal{X})} \left(\int_0^t e^{s\kappa(x,\mathcal{X})} \int_{y\in\mathcal{X}} \kappa(y,\mathrm{d}x)\rho_s(\mathrm{d}y)\,\mathrm{d}s + \bar{\rho}(\mathrm{d}x) \right).$$

It is straightforward to check that for all $t \in [0,T]$

$$\sup_{n\geq 1} \rho_t^n(\mathcal{T}) \leq e^{c_\kappa t}\bar{\rho}(\mathcal{T}) \leq e^{c_\kappa t}\bar{\rho}(\mathcal{T}) =: C,$$

and that on the set of measures with mass bounded by C

$$\|(\mathcal{G}\rho)_t - (\mathcal{G}\tilde{\rho})_t\|_{TV} \leq K\int_0^t \|\rho_s - \tilde{\rho}_s\|_{TV}\,\mathrm{d}s, \qquad \text{for all } t \in [0,T].$$

for some constant $K > 0$. Hence by a Gronwall-type argument we find that for any $\varepsilon > 0$ for all $t \in [0,T]$

$$\|(\mathcal{G}\rho)_t - \mathcal{G}(\tilde{\rho})_t\|_{TV} e^{-(K+\varepsilon)t} \leq \frac{K}{K+\varepsilon}\left(\sup_{s\in[0,T]} \|\rho_s - \tilde{\rho}_s\|_{TV} e^{-(K+\varepsilon)s}\right).$$

Applying Banach fixed-point theorem we find that there exists a unique absolutely continuous curve $\rho_t : t \mapsto (\mathcal{M}^+(\mathcal{X}), \|\cdot\|_{TV})$ such that

$$\rho_t - \rho_s = \int_s^t \left(\int_{y\in\mathcal{X}} \kappa(y,\mathrm{d}x)\rho_r(\mathrm{d}y) - \left(\int_{y\in\mathcal{X}} \kappa(x,\mathrm{d}y)\right)\rho_r(\mathrm{d}x)\right)\mathrm{d}r, \qquad (12.13)$$

for all $s,t \in [0,T]$. Finally, note that ρ_t satisfies (12.13) if and only if $(\rho, \nu_\rho) \in \mathsf{CE}$. Arguing as in Lemma (12.5) we obtain sufficient regularity in time to state that such a curve is indeed a solution. □

The following consequence is obtained from Lemma 12.7 and the fact that $\nu_\rho(\mathrm{d}y,\mathrm{d}x) = u(y)\nu_\pi(\mathrm{d}y,\mathrm{d}x) = u(y)\pi(\mathrm{d}x)\kappa^\dagger(x,\mathrm{d}y)$.

Corollary 12.13. *Suppose that* $\mathrm{d}\bar{\rho} = u_0\mathrm{d}\pi$. *Then* ρ_t *is a solution to* (FKE) *with initial datum* $\bar{\rho}$ *if and only if* $\mathrm{d}\rho_t = u_t\mathrm{d}\pi$ *for all* $t \in [0,T]$ *and* $(u_t)_{t\in[0,T]}$ *is an absolutely continuous and a.e. differentiable curve in* $L^1(\pi)$ *satisfying*

$$\partial_t u_t(x) = \int_{\mathcal{X}} u_t(y)\kappa^\dagger(x,\mathrm{d}y) - u_t(x)\int_{\mathcal{X}} \kappa(x,\mathrm{d}y), \qquad \text{for a.e. } t \in [0,T],\ \pi\text{-a.e. } x. \qquad (12.14)$$

Note that (12.14) can be written as

$$\partial_t u_t(x) = \int_{\mathcal{X}} (u_t(y) - u_t(x))\kappa^\dagger(x,\mathrm{d}y) - u_t(x)\left(\int_{\mathcal{X}} \kappa(x,\mathrm{d}y) - \int_{\mathcal{X}} \kappa^\dagger(x,\mathrm{d}y)\right),$$

and therefore if π is an invariant measure the generator corresponding to $\kappa^\dagger$ is simply the $L^2(\pi)$-adjoint of the generator corresponding to κ. See also Section 12.5 for the relation to (semi)-Dirichlet forms.

12.3 Chain rule and equivalence

In this section, we will discuss the chain rule for the entropy, and the equivalence between the EDP-functional $\mathcal{J}$ and the rate functional $\mathcal{I}$.

Recall the regularized entropy functions ϕ_n of (11.15), and let $\mathrm{d}\rho = u\mathrm{d}\pi$. We consider the functionals

$$S_n(\rho) := \int_{\mathcal{X}} \phi_n(u)\,\mathrm{d}\pi. \tag{12.15}$$

It is straightforward to check that indeed $S_n(\rho) \to S(\rho) := \mathcal{E}\mathrm{nt}(\rho|\pi)$ for every $\rho \in \mathcal{P}(\mathcal{X})$ with $\rho \ll \pi$ as $n \to \infty$.

Theorem 12.3 will now follow directly from the following two Propositions, one establishing the chain rule and the relation $\mathcal{I} = \mathcal{J}$ if $\mathcal{J}$ is finite, and the other stating the inequality $\mathcal{J} \le \mathcal{I}$.

Proposition 12.14. *Let $(\rho, j) \in \mathsf{CE}$ such that $\mathcal{F}(\rho_0) < \infty$ and*

$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t < \infty. \tag{12.16}$$

Then for a.e. t we have $j_t(\{u(x), u(y) > 0\}^c) = 0$. Moreover, for the free energy the chain rule holds: $t \mapsto \mathcal{F}(\rho_t)$ is absolutely continuous and a.e. differentiable with

$$\partial_t \mathcal{F}(\rho_t) = \frac{1}{2}\int_E \overline{\nabla}\phi'\left(u_t\right)\mathrm{d}j_t, \qquad \text{for a.e. } t \in [0,T]. \tag{12.17}$$

Finally,

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t \mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t) \qquad \text{for a.e. } t \in [0,T]. \tag{12.18}$$

and, in particular, $\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j)$.

Remark 12.15. In fact, as the proof below will show, we can modify the assumption on ρ_0 to finite entropy for ρ_t^* for some fixed t^*. Moreover, we have the uniform bound

$$\frac{1}{2}\left|\mathcal{E}\mathrm{nt}(\rho_t|\pi) - \mathcal{E}\mathrm{nt}(\rho_s|\pi)\right| \le \int_s^t \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t + 2|t-s|c_\kappa, \text{ for all } s,t \in [0,T]. \tag{12.19}$$

□

Proof. Consider a curve $(\rho, j) \in \mathsf{CE}$ with $\mathcal{F}(\rho_0) < \infty$ and

$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t = \int_0^T \mathcal{E}\mathrm{nt}(j_t|\theta_{\rho_t})\,\mathrm{d}t < \infty.$$

Since $\rho_0 \ll \pi$ we find by Lemma 12.7 that $\rho_t \ll \pi$ for every $t \in [0,T]$. Moreover, the map $t \mapsto u_t := \mathrm{d}\rho_t/\mathrm{d}\pi$ is absolutely continuous and a.e. differentiable in $L^1(\pi)$, and for every $f \in B(\mathcal{X})$

$$\int_{\mathcal{X}} f\,\partial_t u_t \,\mathrm{d}\sigma = \int_E \overline{\nabla} f \,\mathrm{d}j_t, \qquad \text{for a.e. } t \in [0,T].$$

Note that $\theta_{\rho_t}(\mathrm{d}x,\mathrm{d}y) = \sqrt{u_t(x)u_t(y)}\nu_\pi(\mathrm{d}x,\mathrm{d}y)$ and hence $\theta_{\rho_t}(\{u(x) > 0, u(y) > 0\}^c) = 0$. Since $j_t \ll \theta_{\rho_t}$ for a.e. $t \in [0,T]$ the same statement applies for j_t for such t.

Now, recall the regularized entropy functionals $S_n(\rho)$. By Lemmas 11.3 and 12.5,

$$S_n(\rho_t) - S_n(\rho_s) = \int_s^t \int_E \overline{\nabla}\phi_n'(u_r)\,\mathrm{d}j_r\,\mathrm{d}r, \quad \text{for all } s,t \in [0,T].$$

Recall that $S_n \to \mathcal{E}\mathrm{nt}(\rho|\pi)$ as $n \to \infty$ for all ρ, and $\phi_n'(u) \to \phi(u)$ pointwise. Moreover, by duality of $\mathcal{R}, \mathcal{R}^*$, one can bound

$$\frac{1}{2}\int_E |\overline{\nabla}\phi_n'(u)|\,\mathrm{d}j_r \leq \mathcal{R}(\rho, j) + \mathcal{R}^*(\rho, |\overline{\nabla}\phi_n'(u)|).$$

We will show that for every n and ρ,

$$\mathcal{R}^*(\rho, |\overline{\nabla}\phi_n'(u)|) \leq 2c_\kappa, \tag{12.20}$$

which implies that $S(\rho_t) < \infty$ for all $t \in [0,T]$ and $t \mapsto S(\rho_t)$ is absolutely continuous, and via a dominated convergence argument establishes

$$S(\rho_t) - S(\rho_s) = \int_s^t \int_E \overline{\nabla}\phi'(u_r)\,\mathrm{d}j_r\,\mathrm{d}_r, \quad \text{for all } s,t \in [0,T].$$

Recall that $u_r(x), u_r(y) > 0$ for j_r-a.e. x, y and a.e. $r \in [0,T]$, so that the above expression is well-defined. To verify (12.20), we use the truncation inequality $|\overline{\nabla}\phi_n'(u)| \leq |\overline{\nabla}\phi'(u)|$ and the inequality $e^{|z|} \leq e^z + e^{-z}$ to derive

$$\begin{aligned}
\mathcal{R}^*(\rho, \tfrac{1}{2}|\overline{\nabla}\phi_n'(u)|) &= \int_E \left(e^{\frac{1}{2}|\overline{\nabla}\phi_n'(u)|} - 1\right)\sqrt{u(x)u(y)}\,\mathrm{d}\nu_\pi \\
&\leq \int_{u(x),u(y)>0} \left(e^{\frac{1}{2}\overline{\nabla}\phi_n(u)} + e^{-\frac{1}{2}\overline{\nabla}\phi_n(u)}\right)\sqrt{u(x)u(y)}\,\mathrm{d}\nu_\pi \\
&= \int_{u(x),u(y)>0} \left(\frac{\sqrt{u(y)}}{\sqrt{u(x)}} + \frac{\sqrt{u(x)}}{\sqrt{u(y)}}\right)\sqrt{u(x)u(y)}\,\mathrm{d}\nu_\pi \\
&= \int_{u(x),u(y)>0} (u(y) + u(x))\,\mathrm{d}\nu_\pi.
\end{aligned}$$

However, since $u(x)\nu_\pi(\mathrm{d}x,\mathrm{d}y) = \nu_\rho(\mathrm{d}x,\mathrm{d}y)$ and $u(y)\nu_\pi(\mathrm{d}x,\mathrm{d}y) = \nu_\rho^\dagger(\mathrm{d}x,\mathrm{d}y)$, the desired bound follows.

With the chain rule in hand we can now simplify the EDP-functional to write

$$\mathcal{J}(\rho,j) = \int_0^T \left(\mathcal{R}(\rho_t, j_t) + \frac{1}{2}\int_E \overline{\nabla}\phi'(u_t)\,\mathrm{d}j_t + \mathcal{D}(\rho_t) \right) \mathrm{d}t.$$

Setting $g := \mathrm{d}j/\mathrm{d}\nu_\pi$ and using $\mathrm{d}\theta_\rho = \sqrt{u(x)u(y)}\mathrm{d}\nu_\pi$ the inner term reduces to

$$\begin{aligned}
&\mathcal{R}(\rho,j) + \int_E \overline{\nabla}\phi'(u)\,\mathrm{d}j + \mathcal{D}(\rho)\\
&\quad= \int_E \left(\phi\left(\frac{\mathrm{d}j}{\mathrm{d}\theta_\rho}\right)\mathrm{d}\theta_\rho + (\log u(y) - \log u(x))\,\mathrm{d}j + \left(u(x) - \sqrt{u(x)u(y)}\right)\mathrm{d}\nu_\pi \right)\\
&\quad= \int_E \left(g(x,y)\log\frac{g(x,y)}{u(x)} - g(x,y) + \sqrt{u(x)u(y)} + \left(u(x) - \sqrt{u(x)u(y)}\right)\right)\mathrm{d}\nu_\pi\\
&\quad= \int_E \phi\left(\frac{g(x,y)}{u(x)}\right) u(x)\,\mathrm{d}\nu_\pi = \mathcal{E}\mathrm{nt}(j|\nu_\rho).
\end{aligned}$$

Hence we conclude

$$\mathcal{R}(\rho_t,j_t) + \int_E \overline{\nabla}\phi'(u_t)\,\mathrm{d}j_t + \mathcal{D}(\rho_t) = \mathcal{E}\mathrm{nt}(j_t|\nu_{\rho_t}), \qquad \text{for a.e. } t\in[0,T]. \quad (12.21)$$

and, in particular, $\mathcal{J}(\rho,j) = \mathcal{I}(\rho,j)$. □

Proposition 12.16. *For any $(\rho,j)\in\mathsf{CE}$ such that $\mathcal{F}(\rho_0) < \infty$,*

$$\mathcal{I}(\rho,j) \geq \mathcal{J}(\rho,j).$$

Let us first discuss the strategy of the proof of Proposition 12.16. Instead of making the first calculation of Remark 10.11 rigorous, as in the previous proof, we establish the inequality via the second approach that was outlined, namely a dual approach. For any $w \in B_b(E)$ and ρ, j such that $\rho \ll \pi$ and u is bounded from above and below

$$\mathcal{L}(\rho,j) \geq \int_E \left(w + \tfrac{1}{2}\overline{\nabla}\phi'(u)\right)\mathrm{d}j - \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla}\phi'(u))$$

where the right-most expression can be reduced as

$$\begin{aligned}
\mathcal{H}(\rho, w + \tfrac{1}{2}\phi'(u)) &= \int_E \left(e^{w(x,y)}\frac{\sqrt{u(y)}}{\sqrt{u(x)}} - 1\right) u(x)\nu_\pi\\
&= \int_E \left(e^{w(x,y)}\sqrt{u(x)u(y)} - u(x)\right)\nu_\pi\\
&= \mathcal{R}^*(\rho,w) - \mathcal{D}(\rho).
\end{aligned}$$

However a priori we do not have a chain rule for the entropy yet under finite $\mathcal{I}$, which is where the regularized entropy functionals S_n come in.

Remark 12.17. The two approaches for Propositions 12.14 and 12.16 are not interchangeable. Although we can write via a similar calculation as above

$$\begin{aligned}\mathcal{R}^*(\rho, w - \tfrac{1}{2}\overline{\nabla}\phi'(u)) &= \int_{u(x),u(y)>0} \left(e^{w} \frac{\sqrt{u(x)}}{\sqrt{u(y)}} - 1 \right) \sqrt{u(x)u(y)}\, \mathrm{d}\nu_\pi \\ &= \int_{u(x),u(y)>0} (e^{w} - 1)u(x) + (u(x) - \sqrt{u(x)u(y)})\, \mathrm{d}\nu_\pi,\end{aligned}$$

yet this is not equal to $\mathcal{H}(\rho, w) + \mathcal{D}(\rho)$ for arbitrary ρ, and hence a dual approach would not work for Proposition 12.14. Similarly, there is no β such that for every ρ

$$\mathcal{H}(\rho, \beta|\overline{\nabla}\phi|) < \infty,$$

and thus a chain-rule approach does not work for Proposition 12.16. □

Remark 12.18. Due to the equivalence $\mathcal{I} = \mathcal{J}$ of Theorem 12.3 we will indeed have a posteriori that $j_t(\{u(x), u(y) > 0\}^c) = 0$ for almost every $t \in [0, T]$. In particular, this implies for the solution flux $j_t = \nu_{\rho_t}$ that for such t that $\nu_{\rho_t}(\{u(x), u(y) > 0\}^c)$ and

$$\mathcal{D}(\rho_t) = \int_{u_t(x),u_t(y)>0} (u_t(x) - \sqrt{u_t(x)u_t(y)})\, \mathrm{d}\nu_\pi = \mathcal{R}^*(\rho_t, -\tfrac{1}{2}\overline{\nabla}\phi'(u_t)).$$

Note that this equality does not hold for arbitrary ρ, which is not a contradiction because along the solution curve ρ_t it might only hold for a.e. $t \in [0, T]$. □

Proof of Proposition 12.16. Consider a $(\rho, j) \in \mathsf{CE}$ such that $\mathcal{I}(\rho, j) < \infty$. Similar as in Prop 12.14, by Lemma 12.7 $\rho_t \ll \pi$ for every $t \in [0, T]$, the map $t \mapsto u_t := \mathrm{d}\rho_t/\mathrm{d}\pi$ is absolutely continuous and a.e. differentiable in $L^1(\pi)$, and for every $f \in B(\mathcal{X})$

$$\int_{\mathcal{X}} f\, \partial_t u_t\, \mathrm{d}\sigma = \int_E \overline{\nabla} f\, \mathrm{d}j_t, \qquad \text{for a.e. } t \in [0, T].$$

Moreover, for the regularized entropy functionals

$$S_n(\rho_t) - S_n(\rho_s) = \int_s^t \overline{\nabla}\phi_n'(u_r)\, \mathrm{d}j_r\, \mathrm{d}_r, \quad \text{for all } s, t \in [0, T],$$

$S_n(\rho) \to \mathcal{E}\mathrm{nt}(\rho|\pi)$ as $n \to \infty$ for all ρ, and $\phi_n'(u) \to \phi'(u)$ pointwise. Finally, define the variational derivatives $DS_t^n(x) := \phi'(u_t(x))$.

We can now use the dual approach sketched above to give a lower bound of $\mathcal{I}$. We will first use it to show that $\mathcal{E}\mathrm{nt}(\rho_t|\pi) < \infty$ for all $t \in [0,T]$, and then verify

$$\mathcal{I} \geq \tfrac{1}{2}(\mathcal{E}\mathrm{nt}(\rho_T|\pi) - \mathcal{E}\mathrm{nt}(\rho_0\pi) + \int_0^T \left(\int_E w_t \,\mathrm{d}j_t - \mathcal{R}^*(\rho_t, w_t) + \mathcal{D}(\rho_t) \right) \mathrm{d}t. \tag{12.22}$$

Using the duality (11.12) of Lemma 11.2 with $Y = \mathbb{E} \times [0,T]$ and taking the supremum over all bounded $(w_t)_{t\in[0,T]})$ we obtain the desired inequality.

Note that for every $t \in [0,T]$ and curve $w : [0,T] \to B_b(\mathcal{X})$ with $\sup_t \|w_t\|_\infty < \infty$, we have

$$\begin{aligned}
\mathcal{I} &\geq \int_0^t \mathcal{L}(\rho_r, j_r)\,\mathrm{d}r \\
&\geq \int_0^t \left(\int_E (w_r + \overline{\nabla}\tfrac{1}{2}DS_r^n)\,\mathrm{d}j_r - \mathcal{H}(\rho_r, w_r + \tfrac{1}{2}\overline{\nabla}DS_r^n) \right) \mathrm{d}r \\
&= \tfrac{1}{2}(S_n(\rho_t) - S_n(\rho_0)) + \int_0^t \left(\int_E w_r \,\mathrm{d}j_r - \mathcal{H}(\rho_r, w_r + \tfrac{1}{2}\overline{\nabla}DS_r^n) \right) \mathrm{d}r.
\end{aligned}$$

To establish finiteness of $\mathcal{E}\mathrm{nt}(\rho_t|\pi)$ and (12.22) it is sufficient to show that for every ρ, w (dropping the r-dependence for simplicity)

$$\mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla}DS^n) \leq 2e^{\|w\|_\infty} c_\kappa \tag{12.23}$$

$$\limsup_{n\to\infty} \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla}DS^n) \leq \mathcal{R}^*(\rho, w) - \mathcal{D}(\rho). \tag{12.24}$$

We will employ the truncation inequality $e^{[b]_n - [a]_n} \leq e^{b-a} + 1$ for finite a, b to derive that for any $0 < u, v < \infty$, $|z| < \infty$:

$$\left(e^{z + \frac{1}{2}(\phi_n'(v) - \phi_n'(u))} - 1 \right) u \leq e^z \left(e^{\frac{1}{2}(\phi_n'(v) - \phi_n'(u))} + 1 \right) u = e^z(\sqrt{uv} + u) \leq \frac{e^z}{2}(3u + v), \tag{12.25}$$

and moreover

$$\lim_{n\to\infty} \left(e^{z + \frac{1}{2}(\phi_n'(v) - \phi_n'(u))} - 1 \right) u = e^z\sqrt{uv} - u = (e^z - 1)\sqrt{uv} - (u - \sqrt{uv}). \tag{12.26}$$

Using $\phi_n'(0) = -n$ it is straightforward to check that the statements even hold when $u = 0$ and/or $v = 0$. We can verify (12.23) since the above implies

$$\begin{aligned}
\mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla}DS^n) &= \int_E \left(e^{w(x,y) + \frac{1}{2}\overline{\nabla}DS^n(x,y)} - 1 \right) u(x)\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) \\
&\leq \int_E \frac{1}{2} e^{w(x,y)}(3u(x) + u(y))\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) \leq 2e^{\|w\|_\infty} c_\kappa.
\end{aligned}$$

Moreover, by the bounds above and a dominated convergence argument we conclude the proof. □

Proof of Theorem 12.3. Let $(\rho, j) \in \mathsf{CE}$ be such that $\mathcal{F}_0(\rho) < \infty$. By Proposition 12.14 $\mathcal{I}(\rho, j) = \mathcal{J}(\rho, j)$ if $\mathcal{J}(\rho, j) < \infty$ and thus $\mathcal{I}(\rho, j) \leq \mathcal{J}(\rho, j)$, and we obtain the reverse inequality by Proposition 12.16. We can therefore directly conclude that $\mathcal{I}(\rho, j) = \mathcal{J}(\rho, j)$, and that if either is finite the chain rule holds for $\mathcal{F}$ with

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t \mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t) \qquad \text{for a.e. } t \in [0, T]. \tag{12.27}$$

Next, from Prop 12.12 it is clear $\mathcal{I}(\rho, j) = 0$ if and only if ρ_t is the unique solution to (FKE) and $j_t = \nu_{\rho_t}$ for a.e. $t \in [0, T]$. □

12.4 Compactness and lower semicontinuity

Below we state a compactness result under bounded entropy and action, and that all the ingredients of the EDP-functional are lower semicontinuous with respect to the corresponding notion of convergence. While we will not apply this result to our study of interacting particle systems, we will use the notion of *EDP-convergence* repeatedly in Part II.B. Due to the similarity of proving EDP-convergence to the main result of this section, and the fact that it implies existence of minimizers and well-defined minimizing movement schemes, we include it for completeness.

Let us introduce the following notion of convergence in CE.

Definition 12.19. A sequence $(\rho^n, j^n) \in \mathsf{CE}$ is said to converge to some (ρ, j) if

$$\begin{aligned} \rho_t^n &\to \rho_t && \text{setwise in } \mathcal{P}(\mathcal{X}), \text{ for every } t \in [0, T], \\ j_t^n(\mathrm{d}x, \mathrm{d}y)\,\mathrm{d}t &\to j_t(\mathrm{d}x, \mathrm{d}y)\,\mathrm{d}t && \text{setwise in } \mathcal{M}(\mathbb{E} \times [0, T]). \end{aligned}$$

□

We then have the following compactness and lower semicontinuity result.

Theorem 12.20. *Consider a sequence* $(\rho^n, j^n) \in \mathsf{CE}$ *with*

$$\begin{gathered} \limsup_{n\to\infty} \int_0^T \mathcal{R}(\rho_t^n, j_t^n)\,\mathrm{d}t < \infty, \\ \limsup_{n\to\infty} \mathcal{E}\mathrm{nt}(\rho_{t^*}^n | \pi) < \infty \end{gathered} \tag{12.28}$$

for some $t^* \in [0, T]$. *Then*

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \mathcal{E}\mathrm{nt}(\rho_t^n | \pi) < \infty,$$

and there exists a subsequence (non-relabeled) (ρ^n, j^n) *and a* $(\rho, j) \in \mathsf{CE}$ *such that* $(\rho^n, j^n) \to (\rho, j)$*. Moreover, for any such converging sequence*

$$\liminf_{n\to\infty} \int_0^T \mathcal{R}(\rho_t^n, j_t^n)\,\mathrm{d}t \geq \int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t,$$
$$\liminf_{n\to\infty} \int_0^T \mathcal{D}(\rho_t^n)\,\mathrm{d}t \geq \int_0^T \mathcal{D}(\rho_t)\,\mathrm{d}t,$$
$$\liminf_{n\to\infty} \mathcal{E}\mathrm{nt}(\rho_t^n) \geq \mathcal{E}\mathrm{nt}(\rho_t^n) \quad \textit{for all } t \in [0,T].$$

Proof. The uniform bound on the entropy directly follows from Proposition 12.14 and Remark 12.15 and, in particular, the set $\{\rho_t\}_{t\in[0,T],n\in\mathbb{N}}$ is sequentially compact with respect to setwise convergence. Moreover, due to Corollary 12.9 and Remark 12.10 the sequence of curves $(\rho_t^n)_{t\in[0,T]})_{n\in\mathbb{N}}$ is equicontinuous with respect to d_{TV}, with d_{TV} lower semicontinuous with respect to setwise convergence.

By an Arzelà-Ascoli argument we can now find a curve $(\rho_t)_{t\in[0,T]} \subset \mathcal{P}(\mathcal{X})$ and a subsequence (non-relabeled) such that

$$\rho_t^n \to \rho_t \qquad \text{setwise in } \mathcal{P}(\mathcal{X}), \text{ for every } t \in [0,T].$$

It follows that

$$\nu_{\rho_t^n} \to \nu_{\rho_t} \qquad \text{setwise in } \mathcal{M}^+(E), \text{ for every } t \in [0,T].$$

and therefore $\nu_{\rho_t^n}(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t \to \nu_{\rho_t}(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t$ setwise in $\mathcal{M}^+(\mathbb{E}\times[0,T])$, and similarly $\nu^\dagger_{\rho_t^n}(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t \to \nu^\dagger_{\rho_t}(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t$. Since, in shorthand,

$$\int_0^T \mathcal{R}(\rho_t^n, j_t^n)\,\mathrm{d}t = \mathcal{E}\mathrm{nt}\left(j_t^n(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t \middle| \sqrt{\nu_{\rho_t^n}(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t\, \nu^\dagger_{\rho_t^n}(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t} \right),$$

we obtain by Lemma 11.2 uniform absolute continuity of $j^n\mathrm{d}t$, and hence we can find a measure $J \in \mathcal{M}^+(\mathbb{E}\times[0,T])$ and a subsequence (non-relabeled) such that $j^n(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t \to J$ setwise. It is clear that there exists a Borel family of measures $(j_t)_{t\in[0,T]}$ such that $J(\mathrm{d}x,\mathrm{d}y,\mathrm{d}t) = j_t(\mathrm{d}x,\mathrm{d}y)\,\mathrm{d}t$, and one can pass to the limit in the continuity equation to show that indeed $(\rho, j) \in \mathsf{CE}$.

Finally, lower semicontinuity directly follows from Lemma 11.2. □

12.5 Comments

12.5.1 Choice of π

Lemma 12.7 can provide us with a simple example where finite $\mathcal{I}$ does not imply finite $\int \mathcal{R}\,\mathrm{d}t$, depending on the choice of π. Namely, one can take $\mathcal{X} := [0,1]$,

$\kappa(x, \mathrm{d}y) := \mathscr{L}|_{\mathcal{X}}(\mathrm{d}y)$, $\rho_0 = \delta_{x^*}$ for some $x^* \in [0,1]$, and finally ρ_t the solution curve, so that $\mathcal{I} = 0$. Set $\pi := \mathscr{L}|_{\mathcal{X}}$, such that κ satisfies the detailed balance condition with respect to π. By Lemma 12.7 finite action would imply that $\rho_t^\perp(\mathcal{X})$ is constant in time, which it clearly is not, and hence $\int \mathcal{R}\,\mathrm{d}t = +\infty$.

Now, instead set $\pi = \mathscr{L}|_{\mathcal{X}} + \alpha\delta_{x^*}$ for some $\alpha > 0$, $\kappa(x, \mathrm{d}y) := 1_{y \neq x^*}\pi(\mathrm{d}y)$ and $\kappa^\dagger(x, \mathrm{d}y) := 1_{x \neq x^*}\pi(\mathrm{d}y)$. Again let $\rho_0 = \delta_{x^*}$ and ρ_t the solution curve, then clearly $\mathcal{E}\mathrm{nt}(\rho_0|\pi) < \infty$ and therefore $\mathcal{J} = \mathcal{I} = 0$ by Theorem 12.3.

12.5.2 Locally finite π

While throughout we consider π to be a finite Borel measure, many of the listed results might carry over to the case of locally finite π (and hence Radon) under suitable moment assumptions. Namely, define the modified entropy functional

$$S(\rho) := \begin{cases} \displaystyle\int_{\mathcal{X}} u \log u \,\mathrm{d}\pi, & \text{if } \mathrm{d}\rho = u\,\mathrm{d}\pi,\ \int_{\mathcal{X}} \max(u\log u, 0)\,\mathrm{d}\pi < \infty, \\ +\infty & \text{otherwise} \end{cases}$$

and suppose that there exists a function $V : \mathcal{X} \to [0, +\infty]$ with

$$0 < \int_{\mathcal{X}} e^{-V}\,\mathrm{d}\pi < \infty.$$

Let $\mathrm{d}\pi_V := \frac{1}{Z_V} e^{-V}\mathrm{d}\pi$ with $Z_V := \int e^{-V}\mathrm{d}\pi$. Then one can show that $S(\rho) < \infty$ if and only if $\mathcal{E}\mathrm{nt}(\rho|\pi_V) < +\infty$, for any ρ with $\int V\mathrm{d}\rho < \infty$, and in that case

$$S(\rho) = \mathcal{E}\mathrm{nt}(\rho|\pi_V) - \int_{\mathcal{X}} V\mathrm{d}\rho - \log \int_{\mathcal{X}} e^{-V}\mathrm{d}\pi.$$

In particular, there exists a constant C such that

$$\int_{\mathcal{X}} \min(u\log u, 0)\,\mathrm{d}\pi > C - \int_{\mathcal{X}} V\mathrm{d}\rho,$$

and therefore S is coercive on the sets $A_L := \{\rho : \int V\mathrm{d}\rho \leq L\}$ for all $L > 0$. Moreover, proven similar to the argument in Lemma 19.11, S is setwise and narrowly lower semicontinuous on A_L if in addition

$$\int_{\mathcal{X}} e^{\beta V}\mathrm{d}\pi < \infty, \text{ for all } \beta \geq 0.$$

Let us introduce

$$s_n(z) := \int_0^z ([\log r]_n + 1)\,\mathrm{d}r,\ s(z) := z\log z, \qquad z \geq 0.$$

Note as $n \to \infty$ we have $s_n(1) \downarrow 0$, $s_n(z) - s_n(1) \uparrow s(z) - s(1) = s(z)$ for $z \geq 1$, $\min(s_n(z), 0) \downarrow \min(s(z), 0)$ for $z \in [0, 1]$, and $0 \leq \max(s_n(z), 0) \leq s_n(1)$ for $z \in [0, 1]$. Using the elementary inequality $\pi(\{u(x) \geq a\}) \leq a^{-1}$ for every $a > 0$ we can derive that $S_n(\rho) < \infty$ and $S_n(\rho) \to S(\rho)$ as $n \to \infty$ for every $\rho \ll \pi$ with $\int V \mathrm{d}\rho < \infty$, where

$$S_n(\rho) := \int_{\mathcal{X}} s_n(u) \, \mathrm{d}\pi.$$

Recall the regularized entropy functions ϕ_n, and note $s_n'(a) - s_n'(b) = \phi_n'(a) - \phi_n'(b)$. Therefore one can repeat the calculations of Theorem 12.3 to prove that the chain rule for $S(\rho)$ holds for curves $(\rho, j) \in \mathsf{CE}$ with $\rho_t \ll \pi$ and $\rho_t \in A_L$ for all $t \in [0, T]$. If one can show that these properties are propagated along the solution, we obtain a variational representation along the line of Theorem 13.2.

12.5.3 Minimizing movement scheme

Let CE_τ be the set of curves satisfying the continuity equation over a time window $[0, \tau]$. Then we define the following variational transport cost

$$\mathscr{W}(\tau, \rho^1, \rho^2) := \inf_{(\rho, j) \in \mathsf{CE}_\tau} \left\{ \int_0^\tau \mathcal{R}(\rho_t, j_t) \, \mathrm{d}t \ : \ \rho_0 = \rho^1, \rho_\tau = \rho^2 \right\}. \tag{12.29}$$

While we will not use it further in this work, a natural object of study is to consider the generalized minimizing movement scheme outlined in [PRST22] based on $\mathcal{F}$ and this transport cost, i.e. a scheme involving a time-step τ and the minimization problem

$$\rho^n \in \operatorname{argmin}_\rho \left\{ \mathscr{W}(\tau, \rho^{n-1}, \rho) + \mathcal{F}(\rho) \right\} \tag{12.30}$$

A close look at the proof of [PRST22, Theorem 4.26] would confirm that on sub-level sets of $\mathcal{F}$ the cost $\mathcal{W}$ satisfies all of the properties listed, *except* the non-degeneracy condition since

$$\mathscr{W}(\tau, \rho_0, \rho_1) = 0$$

is not equivalent to $\rho_0 = \rho_1$. This is because zero $\mathcal{W}$ is equivalent to $j_t = \theta_{\rho_t} = \sqrt{\nu_\rho \nu_\rho^\dagger}$ for a.e. t, and $\overline{\nabla} \cdot \theta_{\rho_t}$ might not be zero, unless $\kappa^\dagger = \mathsf{s}_\# \kappa$. However, note that one can still bound the increase of entropy in the minimization step by τc_κ. With this in mind, aside from the possible increase of the free energy $\mathcal{F}$, the non-degeneracy is similar to minimizing movement schemes for GENERIC or non-gradient systems as in [JST22, ADdR22], and might not be a problem.

Chapter 13

Unbounded kernel I: bounded fluxes

In the previous section we discussed the case of a bounded kernel, but many of the results directly carry over to the case where all the fluxes involved are finite, exchanging Assumption 10.7 for Assumption 10.8. Since we assume less on the kernel κ we have to to require more on the curves (ρ, j) themselves, which will be encoded in the continuity equation.

Definition 13.1 (Continuity equation)**.** A pair $(\rho, j) \in \mathsf{CE}_b^*$ if

1. the curve $[0,T] \ni t \mapsto \rho_t \in \mathcal{P}(\mathcal{X})$ is absolutely continuous with respect to $\|\cdot\|_{TV}$,
2. the Borel family $(j_t)_{t\in[0,T]} \subset \mathcal{M}^+(E)$ satisfies $\int_0^T \|j_t\|_{TV}\,\mathrm{d}t < \infty$,
3. for every $s, t \in [0,T]$ and all $f \in B_b(X)$
$$\int_{\mathcal{X}} f\,\mathrm{d}\rho_t - \int_{\mathcal{X}} f\,\mathrm{d}\rho_s = \int_s^t \int_E (f(y) - f(x))\,\mathrm{d}j_r\,\mathrm{d}r, \tag{13.1}$$
4. there exists a $c_{\kappa,\rho} < \infty$ depending on (ρ, j) such that
$$\sup_{t\in[0,T]} \int_E \rho_t(\mathrm{d}x)\kappa(x,\mathrm{d}y),\ \sup_{t\in[0,T]} \int_E \rho_t(\mathrm{d}x)\kappa^\dagger(x,\mathrm{d}y) \le c_{\kappa,\rho}. \tag{13.2}$$

□

Note that the above implies
$$\sup_{t\in[0,T]} \max\left(\nu_{\rho_t}(E), \nu_{\rho_t}^\dagger(E), \theta_{\rho_t}(E)\right) \le c_{\kappa,\rho} < \infty. \tag{13.3}$$

We now directly have the counterpart of Theorem 12.3:

Theorem 13.2. *For any* $(\rho, j) \in \mathsf{CE}_b^*$ *with* $\mathcal{F}(\rho_0) < \infty$

$$\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j).$$

In particular, $\mathcal{J}(\rho, j) \geq 0$ *and*

$$\mathcal{J}(\rho, j) = 0 \iff \begin{cases} \rho_t \text{ is the unique solution to (FKE) that satisfies (13.2)} \\ j_t = \nu_{\rho_t} \text{ for a.e. } t \in [0, T]. \end{cases}$$

Moreover, if $\mathcal{J}(\rho, j) < \infty$ *the following chain rule holds for the free energy* $\mathcal{F}$*, i.e. the map* $t \mapsto \mathcal{F}(\rho_t)$ *is absolutely continuous and a.e. differentiable with*

$$\partial_t \mathcal{F}(\rho_t) = \frac{1}{2} \int_E \overline{\nabla} \phi' \left(u_t\right) \mathrm{d}j_t, \qquad \text{for a.e. } t \in [0, T], \text{ and}$$

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t \mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t), \qquad \text{for a.e. } t \in [0, T].$$

Note that under the assumption of (13.2) we still have the equality

$$\mathcal{K}(\nu_{\rho_t}, \nu_{\rho_t}^\dagger) = 2\left(\mathcal{R}(\rho_t, \nu_{\rho_t}) + \mathcal{D}(\rho_t)\right), \qquad \text{for all } t \in [0, T].$$

Moreover, the underlying subtlety is that the equivalence $\mathcal{J} = \mathcal{I}$, the chain rule and the uniqueness of EDP-solutions follow directly under the integrability condition (13.2). However, to obtain existence of solutions and propagation of suitable moments we will need additional assumptions.

Assumption 13.3. *There exists a measurable function* $\mathsf{f} \in B(\mathcal{X})$ *with* $\mathsf{f} \geq 1$ *with the following properties:*

(B1) There exist constants $A, B \geq 0$ *such that*

$$\int_{\mathcal{X}} (\mathsf{f}(y) - \mathsf{f}(x))_+ \kappa(x, \mathrm{d}y) \leq A\mathsf{f}(x) + B, \quad \text{for all } x \in \mathcal{X} \tag{13.4}$$

(B2) For all $x \in \mathcal{X}$

$$\int_{\mathcal{X}} \kappa(x, \mathrm{d}y), \int_{\mathcal{X}} \kappa^\dagger(x, \mathrm{d}y) \leq \mathsf{f}(x) \tag{13.5}$$

(B3) If $\rho^n \to \rho$ *setwise (see Section 11.1) in* $\mathcal{P}(\mathcal{X})$ *and*

$$\sup_n \int_{\mathcal{X}} \mathsf{f} \, \mathrm{d}\rho^n < \infty,$$

then $\nu_\rho^n \to \nu_\rho$ *setwise in* $\mathcal{M}^+(E)$*.*

Here Assumption (B1) is to guarantee the propagation of f-moments along solutions, the condition (13.5) in turn implies the estimate (13.2), and Assumptions (B1) and (B3) allow us to use a regularization approach to obtain existence. Together this will imply the following.

Proposition 13.4. *Suppose that Assumption (B1) holds and $\bar\rho$*

$$\int_{\mathcal{X}} \mathsf{f}\,\mathrm{d}\bar\rho < \infty.$$

Then any solution to (FKE) *with initial datum $\bar\rho$ satisfies* (13.2). *If $\mathcal{F}(\bar\rho_0) < \infty$, such a solution exists.*

Remark 13.5. Even in the detailed balance case, if π is only locally finite and κ grows too unbounded, solutions might only exist in $\mathcal{M}^+(\mathcal{X})$ instead of $\mathcal{P}(\mathcal{X})$, with the mass $\rho_t(\mathcal{X})$ decreasing over time. This corresponds to escaping to infinity of the associated process, see Section 14.4.1 for a discussion on this and the related notion of stochastic completeness. □

Remark 13.6. Note that existence of suitable solutions for a bounded kernel would also follow from the theory of Dirichlet forms, see Section 10.5.5, but for irreversible processes with unbounded kernels this is less straightforward. Typically one requires better integrability estimates on the antisymmetric part $\kappa - \kappa^\dagger$ than the symmetric part $\kappa + \kappa^\dagger$, see for example [Osh13, Section 1.5.2] and [SW15], which we do not expect for our setting. □

13.1 Solutions

Recall the notion of solutions to (FKE) of Definition 12.11, i.e. a TV-absolutely continuous and a.e. differentiable mapping $\rho : [0,T] \to (\mathcal{P}(\mathcal{X}), \|\cdot\|_{TV})$ satisfying

$$\partial_t \rho_t(\mathrm{d}x) = -\overline{\nabla}\cdot \nu_{\rho_t}, \qquad \text{for a.e. } t \in [0,T],$$

where, to make it well-defined, we assume that a priori

$$\int_0^T \nu_{\rho_t}(\mathrm{d}x)\,\mathrm{d}t < \infty.$$

Note that for any solution satisfying (13.2) automatically $(\rho, \nu_\rho) \in \mathsf{CE}_b^*$, and vice versa. Moreover, we will consider the notion of EDP-solutions satisfying (13.2): i.e. curves ρ_t that are solutions to (FKE) such that $\mathcal{E}\mathrm{nt}(\rho_0|\pi) < \infty$, (13.2) is satisfied, and the energy-dissipation balance holds: for all $t \in [0,T]$

$$\mathcal{E}\mathrm{nt}(\rho_0|\pi) - \mathcal{E}\mathrm{nt}(\rho_t|\pi) = \int_0^t \mathcal{K}(\nu_{\rho_s}, \nu_{\rho_s}^\dagger)\,\mathrm{d}s.$$

Since $\mathcal{K}(\nu_{\rho_s}, \nu^\dagger_{\rho_s}) \geq -c_{\kappa,\rho}$ for all $s \in [0,T]$ by assumption of (13.2) this expression is well-defined, and $\mathcal{E}\mathrm{nt}(\rho_t|\pi) < \infty$ for all $t \in [0,T]$.

The following is a well-known property for gradient flow solutions and entropy solutions, and is a consequence of the convexity of the dissipation, and the strict convexity of the entropy.

Lemma 13.7. *Assume that the following chain-rule inequality holds, i.e. for any* $(\rho, j) \in \mathsf{CE}^*_b$ *and for all* $t \in [0,T]$*:*

$$\tfrac{1}{2}\mathcal{E}\mathrm{nt}(\rho_t|\pi) - \tfrac{1}{2}\mathcal{E}\mathrm{nt}(\rho_0|\pi) + \int_0^t \big(\mathcal{R}(\rho_s, j_s) + \mathcal{D}(\rho_s)\big)\,\mathrm{d}s \geq 0,$$

then EDP-solutions for given initial data are unique.

Note that the chain rule inequality holds if the map $t \mapsto \mathcal{F}(\rho_t)$ is absolutely continuous and a.e. differentiable with

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t\mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t), \qquad \text{for a.e. } t \in [0,T].$$

Proof. Let ρ^1_t and ρ^2_t be two EDP-flow solutions satisfying (13.2), for common initial datum $\bar\rho \in \mathcal{P}(\mathcal{X})$ with $\mathcal{E}\mathrm{nt}(\bar\rho|\pi) < \infty$. Fix t and set

$$\mathcal{J}_t(\rho) := \mathcal{E}\mathrm{nt}(\rho_t) - \mathcal{E}\mathrm{nt}(\bar\rho) + \int_0^t \mathcal{K}(\nu_{\rho_s}, \nu^\dagger_{\rho_s})\,\mathrm{d}s,$$

and note $\mathcal{J}_t(\rho) \geq 0$ for all $t \in [0,T]$ by the chain-rule inequality.

Now assume that $\rho^1_t \neq \rho^2_t$, and take the convex combination $\hat\rho_t = \frac{1}{2}\rho^1_t + \frac{1}{2}\rho^2_t$, which by linearity of the continuity equation satisfies $(\hat\rho, \nu_{\hat\rho}) \in \mathsf{CE}^*_b$, and clearly $\hat\rho_0 = \bar\rho$. Since $\mathcal{K}$ is convex in its arguments and $\rho \mapsto \mathcal{E}\mathrm{nt}(\rho|\pi)$ is strictly convex we obtain

$$\mathcal{J}_t(\hat\rho) < \frac{1}{2}(\mathcal{J}_t(\rho^1) + \mathcal{J}_t(\rho^2)) = 0,$$

which leads to a contradiction. □

Now let us establish the existence of a solution and propagation of moments under the additional constraints Assumptions (B1)-(B3). We will use the regularized kernels

$$\kappa_\varepsilon(x, \mathrm{d}y) := \frac{1}{1 + \varepsilon \mathsf{f}(x)\mathsf{f}(y)}\kappa(x, \mathrm{d}y), \quad \kappa^\dagger_\varepsilon(x, \mathrm{d}y) := \frac{1}{1 + \varepsilon \mathsf{f}(x)\mathsf{f}(y)}\kappa^\dagger(x, \mathrm{d}y).$$

Note that $\kappa_\varepsilon, \kappa^\dagger_\varepsilon$ are bounded kernels satisfying Assumption 10.7 with $c_{\kappa_\varepsilon} = \varepsilon^{-1}$ and, in particular, we have the equality

$$\pi(\mathrm{d}x)\kappa_\varepsilon(x, \mathrm{d}y) = \pi(\mathrm{d}y)\kappa^\dagger_\varepsilon(y, \mathrm{d}x).$$

By Theorem 12.3 solutions to the associated FKEs exist, and Assumption 10.8 allows us to control the dissipation of $\mathcal{E}\mathrm{nt}(\rho|\pi)$ and obtain limits of these regularized problems.

Proof of Proposition 13.4. For any $\varepsilon \geq 0$ let ρ_t^ε be the solution to the regularized problem corresponding to the kernels κ_ε, for the initial datum $\bar\rho$ satisfying

$$\mathcal{E}\mathrm{nt}(\bar\rho|\pi) < \infty, \int_{\mathcal{X}} \mathsf{f}\,\mathrm{d}\bar\rho < \infty.$$

We will first show that these bounds are propagated in time, uniformly in ε. Namely, consider the truncations $\mathsf{f}_m(x) := \min(\mathsf{f}(x), m)$ for $m \geq 0$. It is straightforward to check that via the elementary inequality $(b_m - a_m)_+ \leq (b - a)_+$ that f_m satisfies the inequality (13.4) as well, i.e for all $\varepsilon, m \geq 0$

$$\int_{\mathcal{X}} (\mathsf{f}_m(y) - \mathsf{f}_m(x))_+ \kappa_\varepsilon(x, \mathrm{d}y) \leq \int_{\mathcal{X}} (\mathsf{f}_m(y) - \mathsf{f}_m(x))_+ \kappa(x, \mathrm{d}y) \leq A\mathsf{f}_m(x) + B,$$

From the continuity equation and a Gronwall-type argument we derive that

$$\int_{\mathcal{X}} \mathsf{f}_m\,\mathrm{d}\rho_t^\varepsilon \leq e^{tA}\left(tB + \int_{\mathcal{X}} \mathsf{f}_m\,\mathrm{d}\bar\rho\right) \leq e^{TA}\left(TB + \int_{\mathcal{X}} \mathsf{f}\,\mathrm{d}\bar\rho\right) =: c_\kappa < \infty, \quad (13.6)$$

and, in particular, after taking the limit $m \to \infty$,

$$\sup_{t\in[0,T]} \int_E \rho_t^\varepsilon(\mathrm{d}x)\kappa_\varepsilon(x, \mathrm{d}y),\ \sup_{t\in[0,T]} \int_E \rho_t^\varepsilon(\mathrm{d}x)\kappa_\varepsilon^\dagger(x, \mathrm{d}y) \leq c_\kappa.$$

Therefore, by the energy-dissipation balance,

$$\sup_{t\in[0,T]} \mathcal{E}\mathrm{nt}(\rho_t^\varepsilon|\pi) \leq Tc_\kappa + \mathcal{E}\mathrm{nt}(\bar\rho|\pi) < \infty.$$

Moreover, shown similar as in Lemma 12.5, we have the bound

$$d_{TV}(\rho_t^\varepsilon, \rho_s^\varepsilon) \leq c_\kappa |t - s|.$$

As in the proof of Theorem 12.20 we can now employ Arzelà-Ascoli to extract a subsequence (non-relabeled) and a curve ρ_t such that $\rho_t^\varepsilon \to \rho_t$ setwise in $\mathcal{P}(\mathcal{X})$ as $\varepsilon \to 0$, for every $t \in [0, T]$. By assumption this implies setwise convergence of $\nu_{\rho_t^\varepsilon} \to \nu_{\rho_t}$, and hence by the monotone convergence of $(1 + \varepsilon f(x)f(y))^{-1} \uparrow 1$ as $\varepsilon \to 0$

$$\rho_t^\varepsilon(\mathrm{d}x)\kappa_\varepsilon(x, \mathrm{d}y) \to \rho_t(\mathrm{d}x)\kappa(x, \mathrm{d}y) \quad \text{setwise in } \mathcal{M}^+(E), \text{ for all } t \in [0, T].$$

Therefore, one can pass to the limit in the continuity equation to establish that $(\rho, j) \in \mathsf{CE}_b$. Passing to the limit in (13.6) as well we conclude that in addition

$$\sup_{t\in[0,T]} \int_E \rho_t(\mathrm{d}x)\kappa(x,\mathrm{d}y), \ \sup_{t\in[0,T]} \int_E \rho_t(\mathrm{d}x)\kappa^\dagger(x,\mathrm{d}y) \le c_\kappa < \infty,$$

and thus $(\rho, j) \in \mathsf{CE}_b^*$.

Finally, the propagation of f-moments for solution follows straightforwardly from Gronwall's inequality and the sequence of functions f_m above. □

Remark 13.8. In the case that $\mathcal{X}$ is locally compact one can delete the Assumption (B3), since one can pass to the limit in the continuity equation with respect to compactly supported functions, similar as in Proposition 14.12 and use the bounds above to lift the continuity equation to bounded functions.

However, it is not always true that if a solution is defined weakly with respect to compactly supported functions one can show propagation of f-moments. This plays a role for example in Chapter 17. □

13.2 Chain rule and equivalence

We are now in a position to prove our main result Theorem 13.2

Proof of Theorem 13.2. Under the bound (13.3) it is clear that Lemma 12.5 still applies if the dominating measure is π, i.e. $\rho_t \ll \pi$ for all $t \in [0,T]$, and a similar statement hold for Lemmas 12.6 and 12.7. In particular, we directly obtain the fact that if $(\rho, j) \in \mathsf{CE}_b^*$, $\rho_0 \ll \pi$, and either $\mathcal{I}(\rho, j) < \infty$ or

$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t < \infty,$$

then $t \mapsto \rho_t$ is a.e. differentiable in $(\mathcal{P}(\mathcal{X}), \|\cdot\|_{TV})$ with derivative in $\mathcal{M}(\mathcal{X})$ and

$$\partial_t \rho_t = -\overline{\nabla} \cdot j_t, \qquad \text{for a.e. } t \in [0,T].$$

Moreover, $\rho_t \ll \pi$ for all $t \in [0,T]$, and $\partial_t\rho_t \ll \pi$ for a.e. $t \in [0,T]$.

The equivalence and chain rule now follows similarly as in the proof of Theorem 12.3, using the bounds of (13.3). This implies that solutions satisfying (13.3) are null-minimizers of $\mathcal{J}$ and vice versa, and, in particular, satisfy the chain rule and are EDP-solutions, which are unique by Lemma 13.7. □

Chapter 14

Unbounded kernel II: detailed balance

In the previous chapters we either considered the case of bounded kernels, or a setting in which all the fluxes involved were bounded. We will now drop these assumptions and construct a variational framework where the fluxes are merely Radon measures over a locally compact space.

This is reminiscent of the work of [Erb14] where singular kernels over $\boldsymbol{R}^d$ are considered and the metric structure of the Euclidean space is exploited, and in fact we will give an example where some of our assumptions reduce to the existence of a Heine–Borel metric that is adapted to the jump kernel.

The consequence to working with infinite fluxes is that to control the entropy production we require the reference measure π to be an invariant measure of the jump kernel κ, and even more so, to establish a chain rule we require κ to satisfy the detailed balance condition. This identity, and suitable moment and approximation assumptions, allow us to again rigorously identify the solution to the FKE as the minimizer of the corresponding EDP-functional.

Throughout we will assume Assumption 10.9. Namely, we assume that $\mathcal{X}$ is not just Polish but in addition locally compact, and that the detailed balance condition holds:

$$\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) = \pi(\mathrm{d}y)\kappa(y,\mathrm{d}x). \tag{14.1}$$

Observe that we can now set $\kappa^\dagger = \kappa$ and $\nu_\rho^\dagger = \mathsf{s}_\# \nu_\rho$. In particular, $\nu_\rho^\dagger$ no longer depends explicitly on the invariant measure π, and $\nu_\pi^\dagger = \mathsf{s}_\# \nu_\pi = \nu_\pi$.

Moreover, there exists a function a such that a measurable function $a : \mathcal{X} \to (0, 1/2]$, bounded from below on compact sets, with the property that

$$\max\left(\sup_{x\in\mathcal{X}} a(x)\int_{\mathcal{X}} \kappa(x,\mathrm{d}y), \sup_{x\in\mathcal{X}}\int_{\mathcal{X}} a(y)\kappa(x,\mathrm{d}y)\right) =: \tfrac{1}{2}c_{\kappa,a} < \infty. \tag{14.2}$$

Here the factors 1/2 are for convenience, but can always be absorbed in the constant c_κ. Note that the assumption implies that for any compact set $K \subset \mathcal{X}$,

$$\sup_{x\in K}\int_{\mathcal{X}}\kappa(x,\mathrm{d}y),\ \sup_{x\in\mathcal{X}}\int_K\kappa(x,\mathrm{d}y)<\infty. \tag{14.3}$$

With a little abuse of notation we write $a(x,y) = a(x)+a(y)$, where $0 \le a(x,y) \le 1$. We introduce the weighted total variation over $\mathcal{M}_{\mathrm{loc}}(\mathcal{X}), \mathcal{M}_{\mathrm{loc}}(E)$ respectively,

$$\begin{aligned} \|\mu\|_{a,TV} &:= \int_{\mathcal{X}} a(x)\,\mathrm{d}|\mu|, \qquad \text{for all } \mu\in\mathcal{M}(\mathcal{X}),\\ \|\nu\|_{a,TV} &:= \int_E a(x,y)\,\mathrm{d}|\nu|, \qquad \text{for all } \nu\in\mathcal{M}(E), \end{aligned} \tag{14.4}$$

and let $d_{a,TV}$ the corresponding distance on $\mathcal{M}(\mathcal{X})$.

One can show that under the above assumptions and initial data with finite entropy solutions to (FKE) exist and are absolutely continuous curves in the space $(\mathcal{P}(\mathcal{X}), d_{a,TV})$, as we will show in Proposition 14.12. Because of this, let us redefine the notion of the continuity equation.

Definition 14.1 (Continuity equation)**.** A pair $(\rho, j)\in \mathsf{CE}_a^*$ if

1. the curve $[0,T]\ni t\mapsto \rho_t\in\mathcal{P}(\mathcal{X})$ is absolutely continuous with respect to $\|\cdot\|_{a,TV}$,
2. the Borel family $(j_t)_{t\in[0,T]}\subset\mathcal{M}^+_{\mathrm{loc}}(E)$ satisfies $\int_0^T \|j_t\|_{a,TV}\,\mathrm{d}t<\infty$,
3. for every $s,t\in[0,T]$ and all $f\in B_c(X)$

$$\int_{\mathcal{X}} f\,\mathrm{d}\rho_t - \int_{\mathcal{X}} f\,\mathrm{d}\rho_s = \int_s^t\int_E (f(y)-f(x))\,\mathrm{d}j_r\,\mathrm{d}r, \tag{14.5}$$

4. For the sequence of function ξ_m of Assumption (B2) it holds that

$$\sup_{t\in[0,T]}\int_E |\overline{\nabla}\xi_m|\rho_t(\mathrm{d}x)\kappa(x,\mathrm{d}y) =: L_\rho<\infty, \tag{14.6}$$

□

Note that for any function $f\in B_c$ supported inside the compact set K the integral in 14.5 the integral is well-defined, since finite $\|j\|_{a,TV}$ would imply

$$\int_{K\times\mathcal{X}} j(\mathrm{d}x,\mathrm{d}y),\ \int_{\mathcal{X}\times K} j(\mathrm{d}x,\mathrm{d}y)<\infty.$$

In particular, note that for any $f \in B_c$, $j \in (\mathcal{M}_{\mathrm{loc}}(E), \|\cdot\|_{a,TV})$ the expression $\langle f, \overline{\nabla} j\rangle$ is well-defined and

$$\langle f, \overline{\nabla} j\rangle = \int_E \overline{\nabla} f \, \mathrm{d}j.$$

Define $\mathcal{R}(\rho, j)$, $\mathcal{L}(\rho, j)$ now straightforwardly as functionals over $\mathcal{P}(\mathcal{X})\times\mathcal{M}^+_{\mathrm{loc}}$, and $\mathcal{R}^*(\rho, w)$, $\mathcal{H}(\rho, w)$ as functionals over $\mathcal{P}(\mathcal{X})\times B_c(E)$ with extensions to $\mathcal{P}(\mathcal{X})\times B(E)$ if the integrals involved are finite. Here the expected flux ν_ρ and the geometric average $\theta_\rho = \sqrt{\nu_\rho \mathsf{s}_\# \nu_\rho}$ are now contained in $(\mathcal{M}^+_{\mathrm{loc}}(E), \|\cdot\|_{a,TV})$ with

$$\|\nu_\rho\|_{a,TV}, \|\theta_\rho\|_{a,TV} \le c_{\kappa,a} \tag{14.7}$$

for every $\rho \in \mathcal{P}(\mathcal{X})$, which is easily derived from (14.2). Moreover, the Fisher information $\mathcal{D} : \mathcal{P}(\mathcal{X}) \to [0, +\infty]$ is now defined without the excess term $\nu_\rho - \nu^\dagger_\rho$,

$$\mathcal{D}(\rho) := H^2(\nu_\rho, \nu^\dagger_\rho).$$

and similarly for $\mathcal{K} : (\mathcal{M}^+_{\mathrm{loc}}(\mathcal{X}))^2 \to [0, +\infty]$,

$$\mathcal{K}(\nu, \mu) := \mathcal{E}\mathrm{nt}(\nu|\mu).$$

Remark 14.2. Note that the some of the manipulations for the Fisher information no longer apply, since θ_ρ is no longer a finite measure. However, if $\rho \ll \pi$ we do have the equalities

$$\begin{aligned}\theta_\rho(\mathrm{d}x, \mathrm{d}y) &= \sqrt{u(x)u(y)}\, \nu_\pi(\mathrm{d}x, \mathrm{d}y),\\ \mathcal{D}(\rho) &= \frac{1}{2}\int_E \left(\sqrt{u(y)} - \sqrt{u(x)}\right)^2 \nu_\pi(\mathrm{d}x, \mathrm{d}y)\end{aligned}$$

□

With all these assumptions in hand, we can state the analog of Theorems 12.3 and 13.2 in the current setting.

Theorem 14.3. *For any* $(\rho, j) \in \mathsf{CE}^*_a$ *with* $\mathcal{F}(\rho_0) < \infty$

$$\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j).$$

In particular, $\mathcal{J}(\rho, j) \ge 0$ *and*

$$\mathcal{J}(\rho, j) = 0 \iff \begin{cases} \rho_t \text{ is the unique solution to (FKE) that satisfies (14.6)},\\ j_t = \nu_{\rho_t} \text{ for a.e. } t \in [0, T].\end{cases}$$

Moreover, if $\mathcal{J}(\rho, j) < \infty$ the following chain rule holds for the free energy $\mathcal{F}$, i.e. the map $t \mapsto \mathcal{F}(\rho_t)$ is absolutely continuous and a.e. differentiable with

$$\partial_t \mathcal{F}(\rho_t) = \frac{1}{4} \int_E \overline{\nabla} \phi'(u_t)\, \mathrm{d}j_t^{\mathrm{net}}, \qquad \textit{for a.e. } t \in [0,T], \textit{ and}$$

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t \mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t), \qquad \textit{for a.e. } t \in [0,T].$$

Remark 14.4. Note that even though we do not have all the equivalent definitions of the Fisher information $\mathcal{D}$ as in the case of bounded fluxes, we still have the equality

$$\mathcal{R}(\rho, \nu_\rho) + \mathcal{D}(\rho) = \frac{1}{2} \mathcal{K}(\nu_\rho, \nu_\rho^\dagger). \tag{14.8}$$

This can be seen by symmetrizing the integrands of $\mathcal{R}$ and $\mathcal{K}$ and using the non-negativity of $\mathcal{D}, \mathcal{R}, \mathcal{K}$, since by the symmetry $\mathsf{s}_\# \nu_\rho = \nu_\rho^\dagger$

$$\mathcal{K}(\nu_\rho, \nu_\rho^\dagger) = \frac{1}{2} \left(\mathcal{E}\mathrm{nt}(\nu_\rho | \nu_\rho^\dagger) + \mathcal{E}\mathrm{nt}(\nu_\rho^\dagger | \nu_\rho) \right), \quad \mathcal{D}(\rho) = H^2(\nu_\rho, \nu_\rho^\dagger),$$
$$\mathcal{R}(\rho, \nu_\rho) = \frac{1}{2} \left(\mathcal{E}\mathrm{nt}\left(\nu_\rho \middle| \sqrt{\nu_\rho \nu_\rho^\dagger} \right) + \mathcal{E}\mathrm{nt}\left(\nu_\rho^\dagger \middle| \sqrt{\nu_\rho \nu_\rho^\dagger} \right) \right).$$

□

Remark 14.5. To show the chain rule under finite $\mathcal{J}$ we in fact require a slightly weaker assumption on the functions ξ_m, namely

$$\sup_m \sup_{t \in [0,T]} \int_E |\overline{\nabla} \xi_m|^2 \rho_t(\mathrm{d}x) \kappa(x, \mathrm{d}y) < \infty. \tag{14.9}$$

This gives us a variational formulation of a unique EDP-flow solution to (FKE), if it exists. It is only for the full equivalence that the condition (14.6) comes into play. □

Remark 14.6. Suppose that there exists a Heine–Borel metric $d_\mathcal{X}$ over $\mathcal{X}$, i.e. a metric with compact sublevel sets, that is compatible with the topology of $\mathcal{X}$ and is adapted to the kernel κ in the sense that

$$\sup_{x \in \mathcal{X}} \int_\mathcal{X} (d_\mathcal{X}^2(x, y) \vee 1) \kappa(x, \mathrm{d}y) < \infty. \tag{14.10}$$

Now fixing a point $x^* \mathcal{X}$ and some $\varepsilon > 0$ we can choose a sequence of balls $\{B_m\}_{m \in \mathbb{N}}$ such that $d(B_m, B_{m+1}^c) > \varepsilon$ for all $m \in \mathbb{N}$. This allows us to construct a sequence of nonnegative functions ξ_m that are equal to 1 on B_m, supported in B_{m+1}, and are uniformly Lipschitz with Lipschitz constant ε^{-1}. In particular, (14.9) is satisfied, and (14.6) if we require the slightly stronger condition

$$\sup_{x\in\mathcal{X}}\int_{\mathcal{X}}(d_{\mathcal{X}}(x,y)\vee 1)\kappa(x,\mathrm{d}y)<\infty. \tag{14.11}$$

However, it should be noted that such a metric might not always exist if $\int_{\mathcal{X}}\kappa(x,\mathrm{d}y)$ grows too rapidly as $x\to\infty$. For example, take $\mathcal{X}=\mathbb{N}$ and $\kappa(x,x+1)=k(x+1,x)=x^{(1+\varepsilon)}$ for some $\varepsilon>0$. Then for a metric d to be satisfy (14.11) one would require $d(x,x+1)\sim 1/x^{1+\varepsilon}$ for large x. However, it is straightforward to check that this implies that $d(0,x)<M$ for all x for some fixed $M>0$, and hence d is not a Heine–Borel metric.

See also Section 14.4.1, where we discuss how the properties of d are related to stochastic completeness of the associated Dirichlet form, and the possible escape to infinity for the associated process. □

Similar to the previous chapter uniqueness only holds in the class of solutions satisfying (14.6). In contrast however, existence of solutions only depend on Assumption (B1). Yet, to obtain that solutions propagate suitable moments we do require similar assumptions as in Chapter 13.

Assumption 14.7. *There exists a measurable function* $\mathsf{f}\in B(\mathcal{X})$ *with* $\mathsf{f}\geq 1$ *with the following properties:*

(C1) There exist constants $A,B\geq 0$ *such that*

$$\int_{\mathcal{X}}(\mathsf{f}(y)-\mathsf{f}(x))\kappa(x,\mathrm{d}y)\leq A\mathsf{f}(x)+B,\quad \textit{for all } x\in\mathcal{X} \tag{14.12}$$

(C2) For all $x\in\mathcal{X}$

$$\sup_m\int_{\mathcal{X}}|\overline{\nabla}\xi_m|\kappa(x,\mathrm{d}y)\leq \mathsf{f}(x). \tag{14.13}$$

In light of Proposition 13.4 and the upcoming existence proof Proposition 14.12 we state the following result without proof.

Proposition 14.8. *Suppose that Assumption (C1) holds and* $\bar{\rho}$ *is such that*

$$\int_{\mathcal{X}}\mathsf{f}\,\mathrm{d}\bar{\rho}<\infty.$$

Then any solution to (FKE) *with initial datum* $\bar{\rho}$ *satisfies* (14.6)*, and such a solution exist.*

14.1 Estimates and regularity

Below we will collect the analogues of the time-regularity estimates of Chapters 13 and 12 to the case of weighted TV-spaces. Due to the similarities of the proofs we will only give brief sketches.

Lemma 14.9. *For any $\rho \in \mathcal{P}(\mathcal{X})$, $j \in \mathcal{M}_{\mathrm{loc}}(E)$ finite $\mathcal{R}$ or $\mathcal{L}$ imply that $j \ll \nu_\rho$ and*

$$c_{\kappa,a}\tilde{\phi}\left(c_{\kappa,a}^{-1}\|j\|_{TV,a}\right) \le \mathcal{R}(\rho, j),\ \mathcal{L}(\rho, j), \tag{14.14}$$

For any $(\rho, j) \in \mathsf{CE}_a^*$

$$\|\rho_t - \rho_s\|_{TV,a} \le \int_s^t \|j_r\|_{TV,a}\,\mathrm{d}r \tag{14.15}$$

Finally, suppose that suppose that $(\rho, j) \in \mathsf{CE}_a^*$, $\rho_0 \ll \pi$ and either $\mathcal{I} < \infty$ or

$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t < \infty. \tag{14.16}$$

Then $\rho_t \ll \pi$ for every $t \in [0, T]$ and $t \mapsto u_t := \mathrm{d}\rho_t/\mathrm{d}\pi_t \in L^1(a\pi)$ is absolutely continuous and a.e. differentiable with

$$\partial_t u_t = \int_{\mathcal{X}} (g_t(y, x) - g_t(x, y))\kappa(x, \mathrm{d}y), \qquad \textit{for a.e. } t \in [0, T]$$

where $g_t(x, y) = \mathrm{d}j/\mathrm{d}\nu_\pi$. In particular, the continuity equation holds in a strong sense: $t \mapsto \rho_t \in (\mathcal{P}, \|\cdot\|_{TV,a})$ is absolutely continuous and a.e. differentiable and

$$\partial_t\rho_t = -\overline{\nabla}\cdot j_t, \qquad \textit{for a.e. } t \in [0, T]. \tag{14.17}$$

Proof. The first statement follows from Jensen's inequality and $0 \le a(x, y) \le 1$, since

$$\mathcal{R}(\rho, j) = \int_E \phi\left(\frac{\mathrm{d}j}{\mathrm{d}\theta_\rho}\right)\mathrm{d}\theta_\rho \ge \int_E \tilde{\phi}\left(\frac{\mathrm{d}(aj)}{\mathrm{d}(a\theta_\rho)}\right)\mathrm{d}(a\theta_\rho) \ge \tilde{\phi}\left(\frac{\|j\|_{a,TV}}{\|\theta_\rho\|_{a,TV}}\right)\|\theta_\rho\|_{a,TV}.$$

We then obtain the desired bound after substituting the bound form (14.7).

Next, let $(\rho, j) \in \mathsf{CE}_a^*$ and $f \in B_b(\mathcal{X})$ with $|f|_\infty \le 1$. Then after a density argument via the sequence $f_m := f\mathbb{1}_{K_m}$ and K_m a compact exhaustion of $\mathcal{X}$, we find

$$\int_{\mathcal{X}} af\,\mathrm{d}(\rho_t - \rho_s) \le \int_s^t \int_E |\overline{\nabla}(af)|\,\mathrm{d}j_r\,\mathrm{d}r \le \int_s^t \|j_r\|_{a,TV}\,\mathrm{d}r$$

Taking the supremum over such f we obtain (14.15).

Now, suppose in addition that $\rho_0 \ll \pi$ and either $\mathcal{I} < \infty$ or

$$\int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t < \infty,$$

and, in particular, $j \ll \nu_{\rho_t}$ for a.e. $t \in [0,T]$. Similarly as in Lemma 12.7 we can conclude that $\rho_t \ll \pi$ for all $t \in [0,T]$, which implies $\nu_{\rho_t} \ll \nu_\pi$. Therefore, set $u_t := \mathrm{d}\rho_t/\mathrm{d}\pi$, $\ell := a\pi$, $\Sigma =: a\nu_\pi$, and $g_t := \mathrm{d}j_t/\mathrm{d}\nu_\pi$. It is clear that $u_t : [0,T] \to L^1(\ell)$ is absolutely continuous and $g_t : t \mapsto L^1(\Sigma)$ is integrable on $[0,1]$. Moreover, since $\mathsf{s}_\# j_t = g_t(y,x)\mathsf{s}_\#\nu_\pi = g_t(y,x)\nu_\pi$, we obtain via a density argument over $B_c(\mathcal{X})$,

$$u_t - u_s = \int_s^t \int_{\mathcal{X}} (g_r(y,x) - g_r(x,y))\kappa(x,\mathrm{d}y)\,\mathrm{d}r, \qquad \text{for all } s,t \in [0,T]$$

and the result follows from Lemma 11.5. □

For technical reasons involving the chain rule, we also introduce some estimates on the net-flux, which directly follow from the arguments of Section 10.5.2 and the evenness of Ψ^*.

Lemma 14.10. *Consider $\rho \in \mathcal{P}(\mathcal{X})$, $j \in \mathcal{M}^+_{\mathrm{loc}}$ with $\mathcal{R}(\rho,j) < \infty$ and set $j^{\mathrm{net}} := j - \mathsf{s}_\# j$. Then for any $w \in B(E)$,*

$$\int_E |w|\,\mathrm{d}|j^{\mathrm{net}}| \le \mathcal{R}(\rho,j) + \mathcal{R}^*_{\mathrm{net}}(\rho,w), \tag{14.18}$$

where

$$\mathcal{R}^*_{\mathrm{net}}(\rho,w) = \frac{1}{2}\int_E \Psi^*(2w)\,\mathrm{d}\theta_\rho.$$

14.2 Solutions

Below we discuss the various notions of solutions in our setting, and discuss their existence/uniqueness. Both properties follow similarly as in Section 13.1, with the difference that for the existence we can take the limit in the continuity equation (14.5) with respect to functions with compact support, which frees us from needing to use additional moment estimates. This is made possible due to the detailed balance condition, since the excess term $\int_{\mathcal{X}^2} \nu_\rho - \int_{\mathcal{X}^2} \nu_\rho^\dagger$, which is zero if finite, is removed and hence no longer plays a role in the entropy production.

Definition 14.11. A solution to (FKE) is any narrowly continuous mapping $\rho : [0,T] \to \mathcal{P}(\mathcal{X})$ satisfying

$$\int_{\mathcal{X}} f\,\partial\rho_t(\mathrm{d}x) - \int_{\mathcal{X}} f\,\partial\rho_r(\mathrm{d}x) = \int_s^t \int_E \overline{\nabla} f\,\mathrm{d}\nu_{\rho_r}\,\mathrm{d}r \tag{14.19}$$

for a.e. $t \in [0,T]$, for all $f \in B_c(\mathcal{X})$. □

It is clear after an argument as in Lemma 14.9 that solutions are absolutely continuous with respect to $d_{a,TV}$, and differentiable for a.e. t if $\rho_0 \ll \pi$. Moreover, a curve ρ_t is a solution satisfying (14.6) if and only if $(\rho, \nu_\rho) \in \mathsf{CE}^*_a$. In addition, we can establish existence using the properties of the function a.

Proposition 14.12. *For any $\bar{\rho} \in \mathcal{P}(\mathcal{X})$ with*

$$\mathcal{E}\mathrm{nt}(\bar{\rho}|\pi) < \infty$$

there exists a solution ρ_t to (FKE) *with initial datum $\bar{\rho}$.*

Proof. The proof is similar to the one for Proposition 14.12. We introduce the regularized bounded kernels

$$\kappa_\varepsilon(x, \mathrm{d}y) = \frac{1}{1 + \varepsilon a(x)a(y)} \kappa(x, \mathrm{d}y), \tag{14.20}$$

and note that for the solutions ρ_t^ε to the regularized problems we have the uniform bound

$$d_{a,TV}(\rho_t^\varepsilon, \rho_s^\varepsilon) \le c_\kappa |t - s|.$$

Moreover, from the energy-dissipation balance and the detailed balance condition we obtain

$$\mathcal{E}\mathrm{nt}(\rho_t^\varepsilon|\pi) \le \mathcal{E}\mathrm{nt}(\bar{\rho}_t|\pi).$$

The metric $d_{a,TV}$ is lower semicontinuous with respect to setwise convergence and convergence in $d_{a,TV}$ implies setwise convergence on the sublevel sets of $\mathcal{E}\mathrm{nt}(\cdot|\pi)$. We can therefore again subtract a subsequence (non-relabeled) such that $\rho_t^\varepsilon \to \rho_t$ setwise in $\mathcal{P}(\mathcal{X})$ as $\varepsilon \to 0$ for every $t \in [0,T]$, for some curve ρ_t.

It remains to show that for every $t \in [0,T]$ and every $f \in B_c(\mathcal{X})$

$$\int_E f(x)\mathrm{d}\nu_{\rho_t^\varepsilon} \to \int_E f(x)\mathrm{d}\nu_{\rho_t}, \int_E f(y)\mathrm{d}\nu_{\rho_t^\varepsilon} \to \int_E f(y)\mathrm{d}\nu_{\rho_t},$$

as $\varepsilon \to 0$, since combined with the monotone convergence of $(1 + \varepsilon a(x)a(y))^{-1}$ we can then pass to the limit in the continuity equation (14.5) to conclude that ρ_t is a solution.

However, this convergence is easily seen from the fact that $x \mapsto f(x)\kappa(x, \mathrm{d}y)$ and $x \mapsto f(y)\kappa(x, \mathrm{d}y)$ are bounded mappings due to (14.2) and the compact support of f. □

Remark 14.13. Suppose that for every $f \in C_c(\mathcal{X})$ the map

$$x \mapsto (f(y) - f(x)) \int_{\mathcal{X}} \kappa(x, \mathrm{d}y) \tag{14.21}$$

is itself in $C_c(\mathcal{X})$. Then one can repeat the above argument, replacing the setwise topology by the vague topology, to conclude that for any initial data in $\mathcal{P}(\mathcal{X})$ there exists a curve $(\rho_t)_{t\in[0,T]} \subset \mathcal{M}^+(\mathcal{X})$ with $\rho_t(\mathcal{X}) \le 1$ satisfying (14.19). However, it is clear that there is no guarantee that the solution actually conserves mass. □

We revisit again the notion of EDP-solution, i.e. a solution to (FKE) with $\mathcal{E}\mathrm{nt}(\rho_0|\pi) < \infty$, and the energy-dissipation balance holds: for all $t \in [0, T]$

$$\mathcal{E}\mathrm{nt}(\rho_0|\pi) - \mathcal{E}\mathrm{nt}(\rho_t|\pi) = \int_0^t \mathcal{K}(\nu_{\rho_s}, \nu_{\rho_s}^{\dagger})\,\mathrm{d}s.$$

It is clear from Lemma 13.7 that we again have uniqueness in this class under a chain rule inequality, and we state the following without proof.

Lemma 14.14. *Assume that the following chain-rule inequality holds, i.e. for any* $(\rho, j) \in \mathsf{CE}_a^*$ *and for all* $t \in [0, T]$*:*

$$\tfrac{1}{2}\mathcal{E}\mathrm{nt}(\rho_t|\pi) - \tfrac{1}{2}\mathcal{E}\mathrm{nt}(\rho_0|\pi) + \int_0^t \big(\mathcal{R}(\rho_s, j_s) + \mathcal{D}(\rho_s)\big)\,\mathrm{d}s \ge 0,$$

Then EDP-flow solutions satisfying (14.6) *are unique.*

14.3 Chain rule and equivalence

This section is devoted to proving the chain rule and the equivalence between $\mathcal{J}$ and $\mathcal{I}$ in this setting of unbounded fluxes. Instead of the usual regularized entropies, we build double sequence

$$S_{m,n}(\rho_t) := \xi_m(x)\phi_n(u)\,\mathrm{d}\pi, \tag{14.22}$$

where the compactly supported multipliers $\xi_m(x)$ satisfy Assumption (B2). To control these terms we require the approximation properties of ξ_m and the detailed balance condition for κ.

As in Section 12.3, we split up the proof of Theorem 14.3 into the inequalities $\mathcal{J} \ge \mathcal{I}$ and $\mathcal{J} \le \mathcal{J}$.

Proposition 14.15. *Let* $(\rho, j) \in \mathsf{CE}_a^*$ *such that* $\mathcal{F}(\rho_0) < \infty$ *and* $\mathcal{J}(\rho, j) < \infty$.

Then for a.e. t we have $j_t(\{u(x), u(y) > 0\}^c) = 0$. Moreover, for the free energy the chain rule holds: $t \mapsto \mathcal{F}(\rho_t)$ is absolutely continuous and a.e. differentiable with

$$\partial_t \mathcal{F}(\rho_t) = \frac{1}{4} \int_E \overline{\nabla} \phi'(u_t) \, \mathrm{d} j_t^{\mathrm{net}}, \qquad \textit{for a.e. } t \in [0, T]. \tag{14.23}$$

Finally,

$$(\mathcal{R}(\rho_t, j_t) + \mathcal{D}(\rho_t)) + \partial_t \mathcal{F}(\rho_t) = \mathcal{L}(\rho_t, j_t) \qquad \textit{for a.e. } t \in [0, T]. \tag{14.24}$$

and, in particular, $\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j)$.

Proof. Consider a curve $(\rho, j) \in \mathsf{CE}_a^*$ such that $\mathcal{F}(\rho_0) < \infty$ and $\mathcal{J}(\rho, j) < \infty$. In particular,

$$\int_0^T \mathcal{R}(\rho_t, j_t) \, \mathrm{d}t < \infty, \int_0^T \mathcal{D}(\rho_t) \, \mathrm{d}t < \infty.$$

Moreover, recall that there exists a $L \geq 0$ is such that

$$\sup_{t \in [0,T]} \int_E |\overline{\nabla} \xi_m| \rho_t(\mathrm{d}x) \kappa(x, \mathrm{d}y) \leq L, \qquad \text{for all } m \geq 0, \, x \in \mathcal{X},$$

and note

$$\int_{\mathcal{X}} |\overline{\nabla} \xi_m|^2 \kappa(x, \mathrm{d}y) \leq 2L.$$

The proof now follows along the lines of that of Proposition 12.14. It is straightforward to check that $j_t(\{u(x), u(y) > 0\}^c) = 0$ for a.e. $t \in [0, T]$, and that the chain rule applies for the regularized entropy functionals $S_{m,n}$, i.e. for all $s, t \in [0, T]$

$$\begin{aligned} S_{m,n}(\rho_t) - S_{m,n}(\rho_s) &= \int_s^t \int_E \overline{\nabla}(\xi_m \phi_n') \, \mathrm{d} j_r \, \mathrm{d}r \\ &= \frac{1}{2} \int_s^t \int_E \overline{\nabla}(\xi_m \phi_n') \, \mathrm{d} j_r^{\mathrm{net}} \, \mathrm{d}r, \end{aligned} \tag{14.25}$$

where the last line follows from the fact that $\overline{\nabla} f$ is antisymmetric.

Moreover, $\lim_{n \to \infty} \lim_{m \to \infty} \overline{\nabla}(\xi_m \phi_n') = \overline{\nabla} \phi'$ pointwise for every $x, y \in \mathcal{X}$. We have

$$|\overline{\nabla} \xi_m \phi_n'| \leq |\xi_m(x) \overline{\nabla} \phi_n + \phi_n(y) \overline{\nabla} \xi_m| \leq |\overline{\nabla} \phi_n| + n |\overline{\nabla} \xi_m|$$

and it follows from a dominated convergence argument and Lemma 14.10 that it is sufficient to show that for any $\rho \in \mathcal{P}(\mathcal{X})$ we have the uniform bounds

$$\mathcal{R}_{\mathrm{net}}^* \left(\rho, \tfrac{1}{4} |\overline{\nabla} \phi_n'(u)| \right) \leq \mathcal{D}(\rho), \tag{14.26}$$

and

$$\mathcal{R}^*_{\text{net}}\left(\rho, \tfrac{1}{4}|\overline{\nabla}\xi_m|\right) \leq A, \tag{14.27}$$

for some constant A.

Indeed, since even though we only obtain integrability of $|\overline{\nabla}\phi'|$ with respect to $|j^{\text{net}}|$ instead of j we can still write out the same equivalence after symmetrizing. Namely, substituting $j = g\nu_\pi$, setting $g^\top(x,y) = g(y,x)$, and using the shorthands $u = u(x)$, $v = u(y)$, we can write

$$\begin{aligned}
\frac{1}{4}\int_E \overline{\nabla}\phi' \mathrm{d}j^{\text{net}} &= \frac{1}{4}\int_E (g - g^\top)\log v/y \,\mathrm{d}\nu_\pi, \\
\mathcal{R}(\rho, j) &= \int_E \phi\left(\frac{g}{\sqrt{uv}}\right)\sqrt{uv}\,\mathrm{d}\nu_\pi \\
&= \frac{1}{2}\int_E \phi\left(\frac{g}{\sqrt{uv}}\right)\sqrt{uv}\,\mathrm{d}\nu_\pi + \frac{1}{2}\int_E \phi\left(\frac{g^\top}{\sqrt{uv}}\right)\sqrt{uv}\,\mathrm{d}\nu_\pi, \\
&= \frac{1}{2}\int_E \left(\phi\left(\frac{g}{\sqrt{uv}}\right)\sqrt{uv} + \phi\left(\frac{g^\top}{\sqrt{uv}}\right)\sqrt{uv}\right)\mathrm{d}\nu_\pi \\
\mathcal{D}(\rho) &= \frac{1}{2}\left(\sqrt{u} - \sqrt{v}\right)^2 \mathrm{d}\nu_\pi.
\end{aligned}$$

Collecting all the terms together and repeating the same calculation as for the usual equivalence under finite action of Proposition 12.14, we derive that

$$\mathcal{R}(\rho, j) + \frac{1}{4}\int_E \overline{\nabla}\phi' \mathrm{d}j^{\text{net}} + \mathcal{D}(\rho) = \frac{1}{2}\int_{\mathcal{E}^2}\left(\phi\left(\frac{g}{u}\right)u + \phi\left(\frac{g^\top}{v}\right)v\right)\mathrm{d}\nu_\pi = \mathcal{L}(\rho, j),$$

where the final equality follows from the nonnegativity of ϕ.

Now, to establish (14.26), note that because of the monotonicity of Ψ^* on $[0, +\infty)$ and the truncation inequality $|\overline{\nabla}\phi_n'| \leq |\overline{\nabla}\phi'|$ we can bound

$$\begin{aligned}
\mathcal{R}^*_{\text{net}}\left(\rho, \tfrac{1}{4}|\overline{\nabla}\phi_n'(u)|\right) &= \frac{1}{2}\int_E \Psi^*(\tfrac{1}{2}\overline{\nabla}\phi_n'(u))\sqrt{u(x)u(y)}\,\mathrm{d}\nu_\pi \\
&\leq \frac{1}{2}\int_E \Psi^*(\tfrac{1}{2}|\overline{\nabla}\phi'(u)|)\sqrt{u(x)u(y)}\,\mathrm{d}\nu_\pi \quad = \mathcal{D}(\rho).
\end{aligned}$$

Moreover, by the inequality (11.8) it follows that

$$\begin{aligned}
\mathcal{R}^*_{\text{net}}\left(\rho, \tfrac{1}{4}|\overline{\nabla}\xi_m|\right) &= \frac{1}{2}\int_E \Psi^*(\tfrac{1}{2}\overline{\nabla}\xi_m)\,\mathrm{d}\theta_\rho \leq \frac{\Psi^*(1)}{2}\int_E |\overline{\nabla}\xi_m|^2 \mathrm{d}\theta_\rho \\
&\leq \frac{\Psi^*(1)}{2}\sqrt{\left(\int_E |\overline{\nabla}\xi_m|^2 \mathrm{d}\nu_\rho\right)\left(\int_E |\overline{\nabla}\xi_m|^2 \mathrm{d}\nu_\rho^\dagger\right)} \leq \Psi^*(1)L
\end{aligned}$$

which provides the desired bound (14.27). □

We now turn to the analogue of Proposition 12.16.

Proposition 14.16. *For any* $(\rho, j) \in \mathsf{CE}_a^*$ *such that* $\mathcal{F}(\rho_0) < \infty$,

$$\mathcal{I}(\rho, j) \geq \mathcal{J}(\rho, j)$$

The proof again relies on a dualization argument, but due to the infinite mass of the fluxes involved we have to use a slightly non-trivial truncation argument. Namely, we will employ the following Lemma.

Lemma 14.17. *Let* $u \in B(\mathcal{X})$ *with* $u \geq 0$, $f \in B_b(\mathcal{X})$ *with* $0 \leq f \leq 1$. *Then for any* $x, y \in \mathbb{E}$, $n \geq 0$.

$$\begin{aligned} -\left(\sqrt{u(x)} - \sqrt{u(y)}\right)^2 &\leq \left(e^{\frac{1}{2}\overline{\nabla}(f\phi_n')} - 1\right) u(x) + \left(e^{-\frac{1}{2}\overline{\nabla}(f\phi_n')} - 1\right) u(y) \\ &\leq \frac{ne^{2n}}{2} |\overline{\nabla} f|(u(x) + u(y)). \end{aligned}$$

Proof. The lower bound follows from the elementary inequality

$$(e^{w}-1)u+(e^{-w}-1)v = \left(e^{\frac{1}{2}w}\sqrt{u} - e^{-\frac{1}{2}w}\sqrt{v}\right)^2 - \left(\sqrt{u} - \sqrt{v}\right)^2 \geq -\left(\sqrt{u} - \sqrt{v}\right)^2,$$

for all $u, v \in [0, +\infty)$, $w \in \mathbb{R}$.

To establish the upper bound, first note that by convexity of e^z

$$e^{\frac{1}{2}\overline{\nabla}(f\phi_n')} \leq \tfrac{1}{2} e^{f(y)\overline{\nabla}\phi_n'} + \tfrac{1}{2} e^{\phi_n'(x)\overline{\nabla} f} \leq \tfrac{1}{2} e^{f(y)\overline{\nabla}\phi_n'} + \tfrac{1}{2} e^{n|\overline{\nabla} f|},$$

and similarly

$$e^{-\frac{1}{2}\overline{\nabla}(f\phi_n')} \leq \tfrac{1}{2} e^{-f(y)\overline{\nabla}\phi_n'} + \tfrac{1}{2} e^{n|\overline{\nabla} f|}.$$

Combining the terms, we find that

$$\begin{aligned} &\left(e^{\frac{1}{2}\overline{\nabla}(f\phi_n')} - 1\right) u(x) + \left(e^{-\frac{1}{2}\overline{\nabla}(f\phi_n')} - 1\right) u(y) \\ &\qquad \leq \frac{1}{2}\left((e^{f(y)\overline{\nabla}\phi_n'} - 1)u(x) + (e^{-f(y)\overline{\nabla}\phi_n'} - 1)u(y)\right) \\ &\qquad\quad + \frac{ne^{2n}}{2} |\overline{\nabla} f|(u(x) + u(y)). \end{aligned}$$

Since $0 \leq f(y) \leq 1$, by convexity it remains to show that for any $u, v \in [0, +\infty)$, $\alpha \in [0, 1]$

$$\left(e^{\alpha(\phi_n'(v) - \phi_n'(u))} - 1\right) u + \left(e^{\alpha(\phi_n'(v) - \phi_n'(u))} - 1\right) v \leq 0 \tag{14.28}$$

Since the inequality holds for $\alpha = 0$ it is sufficient to show it for $\alpha = 1$. Moreover, note that inequality holds if either $u = 0$ or $v = 0$, so let us assume without loss of generality that $v, u > 0$. Finally, whenever $\phi'(u), \phi'(v) \in [-n, n]$ the expression on the left-hand side of (14.28) reduces to

$$\left(e^{(\phi'(v)-\phi'(u))} - 1\right) u + \left(e^{(\phi'(v)-\phi'(u))} - 1\right) v = \left(\frac{v}{u} - 1\right) u + \left(\frac{u}{v} - 1\right) v = 0 \tag{14.29}$$

and, in particular, the inequality (14.28) is satisfied.

Now, let u_n^+ and u_n^- be defined by the relations $\phi_n'(u_n^+) = n, \phi_n'(u_n^-) = -n$, and suppose that $\phi'(u) \geq n$. Then $\phi_n'(u) = n = \phi'(u_n^+)$, $(\phi_n'(v) - \phi_n'(u)) \leq 0$ and hence

$$\left(e^{(\phi_n'(v)-\phi_n'(u))} - 1\right) u \leq \left(e^{(\phi_n'(v)-\phi_n'(u))} - 1\right) u_n^+ = \left(e^{(\phi'(v)-\phi'(u_n^+))} - 1\right) u_n^+$$

Moreover,

$$\left(e^{(\phi_n'(u)-\phi_n'(v))} - 1\right) v = \left(e^{(\phi'(v)-\phi'(u_n^+))} - 1\right) v$$

which implies the desired inequality via (14.29). The other cases follow similarly. □

Proof of Proposition 14.16. Consider a $(\rho, j) \in \mathsf{CE}_a^*$ such that $\mathcal{I}(\rho, j) < \infty$. Following the proof of Proposition 12.16, it is clear that for every $t \in [0, T]$ and curve $w : [0, T] \to B_c(E)$ with $w \in B_c(\mathbb{E} \times [0, T])$, i.e. with $\sup_{r\in[0,T]} \|w_t\|_\infty =: \|w\|_\infty < \infty$ and $\mathrm{supp}(w_t) \subset K$ for all t and some compact set $K \times K$ with $K \subset \mathcal{X}$ compact,

$$\begin{aligned} \mathcal{I} &\geq \int_0^t \mathcal{L}(\rho_r, j_r)\, \mathrm{d}r \\ &\geq \int_0^t \left(\int_E (w_r + \overline{\nabla} \tfrac{1}{2} DS_r^{m,n})\, \mathrm{d}j_r - \mathcal{H}(\rho_r, w_r + \tfrac{1}{2}\overline{\nabla} DS_r^{m,n}) \right) \mathrm{d}r \\ &= \tfrac{1}{2}(S^{m,n}(\rho_t) - S^{m,n}(\rho_0)) + \int_0^t \left(\int_E w_r\, \mathrm{d}j_r - \mathcal{H}(\rho_r, w_r + \tfrac{1}{2}\overline{\nabla} DS_r^{m,n}) \right) \mathrm{d}r, \end{aligned} \tag{14.30}$$

where $DS_r^{m,n} = \xi_m \phi_n'(u_r)$. It is therefore sufficient to show (dropping the r-dependence for simplicity) that for every $\rho \in \mathcal{P}(\mathcal{X})$, $w \in B_c(\mathcal{X})$ with support in the set $K \times K$,

$$\begin{aligned} \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla} DS^{m,n}) &\leq e^{\|w\|_\infty + n} \sup_{x\in K} \kappa(x, \mathcal{X}) + \frac{nLe^{2n}}{2} \\ \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla} DS^{n}) &\leq 2e^{\|w\|_\infty} \sup_{x\in K} \kappa(x, \mathcal{X}), \end{aligned} \tag{14.31}$$

$$\begin{aligned}\limsup_{m\to\infty} \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla} DS^n) &\le \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla} DS^n)\\ \limsup_{n\to\infty} \mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla} DS^n) &\le \mathcal{R}^*(\rho, w) - \mathcal{D}(\rho),\end{aligned} \tag{14.32}$$

since the desired inequality follows after Fatou's lemma and taking the supremum in (14.30) over all $w \in B_c(\mathcal{X} \times [0, T])$. Moreover, since the pointwise calculation for (14.32) is precisely the same as for Proposition 12.16, we will focus on bounds of the integrands. We can write, using the shorthand $u = u(x)$, $v = u(y)$, $w^\top(x, y) = w(y, x)$,

$$\begin{aligned}\mathcal{H}(\rho, w + \tfrac{1}{2}\overline{\nabla} DS^{m,n}) &= \int_E \left(e^{w + \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) u \,\mathrm{d}\nu_\pi\\ &= \frac{1}{2}\int_E \left(\left(e^{w + \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) u + \left(e^{w^\top - \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) v\right) \mathrm{d}\nu_\pi.\end{aligned}$$

We will consider the sets $K \times K$ and $(K \times K)^c$ separately. Note that $\sup_{x\in K} \kappa(x, \mathcal{X}) < \infty$ by the bound (14.3). Thus, on the set $K \times K$ we can proceed as in the case of a bounded kernel and, in particular, for $x, y \in K$

$$\begin{aligned}\left(e^{w + \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) u + \left(e^{w^\top - \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) v &\le e^{\|w\|_\infty + n}(u + v),\\ \left(e^{w + \frac{1}{2}\overline{\nabla} DS^{n}} - 1\right) u + \left(e^{w^\top - \frac{1}{2}\overline{\nabla} DS^{n}} - 1\right) v &\le 2e^{\|w\|_\infty}(u + v).\end{aligned}$$

Next we consider the set $(K \times K)^c$, upon which $w = 0$. Therefore, applying Lemma 14.17, we have for $(x, y) \in (K \times K)^c$

$$\begin{aligned}\left(e^{w + \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) u + \left(e^{w^\top - \frac{1}{2}\overline{\nabla} DS^{m,n}} - 1\right) v &\le \frac{ne^{2n}}{2} |\overline{\nabla} \xi_m| (u(x) + u(y)),\\ \left(e^{w + \frac{1}{2}\overline{\nabla} DS^{n}} - 1\right) u + \left(e^{w^\top - \frac{1}{2}\overline{\nabla} DS^{n}} - 1\right) v &\le 0.\end{aligned}$$

Integrating the quantities provides the bound (14.31). Moreover, the statement (14.32) follows from pointwise convergence and Fatou's Lemma, and thus we conclude the proof. □

Proof of Theorem 14.3. The proof is similar to that of 12.3 and 13.2. Namely, Propositions 14.15 and 14.16 together imply $\mathcal{J}(\rho, j) = \mathcal{I}(\rho, j)$ and the chain rule under finite $\mathcal{J}$ for any curve $(\rho, j) \in \mathsf{CE}_a^*$ with $\mathcal{F}(\rho_0) < \infty$. In particular, solutions to (FKE) satisfying (14.6) are null-minimizers of $\mathcal{J}$ and are EDP-solutions satisfying (14.6), and vice versa. Finally, by Lemma 14.14 EDP-solutions are unique. □

14.4 Comments

14.4.1 Stochastic completeness

Recall the setting of Remark 14.6, namely we consider a continuous Heine–Borel metric $d_{\mathcal{X}}$ that is in addition an intrinsic metric, i.e. it is adapted to κ in the sense that

$$\sup_{x\in\mathcal{X}} \int_{\mathcal{X}} (d_{\mathcal{X}}^2(x,y) \vee 1)\kappa(x, \mathrm{d}y) < \infty. \tag{14.33}$$

Moreover, recall the Dirichlet form $\mathcal{E}$ of Section 10.5.5 on $L^2(\pi)$, given by

$$\mathcal{E}(f,g) := \int_E (f(x) - f(y))g(x)\pi(\mathrm{d}x)\kappa(x,\mathrm{d}y) = \frac{1}{2}\int_E \overline{\nabla} f \overline{\nabla} g \pi(\mathrm{d}x)\kappa(x,\mathrm{d}y),$$

where the final equality follows from the symmetry of ν_π.

Let us for the moment assume that π is not necessarily finite, but merely some nonnegative Radon measure over $\mathcal{X}$. Following [GHM12], it easy to check that $\mathcal{E}(f,f)$ is finite for any $f \in C_c(\mathcal{X}) \cap BL(\mathcal{X})$, with $BL(\mathcal{X})$ the space of bounded $d_{\mathcal{X}}$-Lipschitz functions, and after taking an appropriate closure of this set we obtain a domain D such that the $(\mathcal{E}, D)$ is a regular Dirichlet form. Due to the boundedness of $\kappa(x,\mathcal{X})$ over compact sets in our setting it can be shown that in our case D is the maximal domain.

Moreover, a Dirichlet form is called *stochastically complete* if $P_t 1 = 1$ for all $t \geq 0$, where P_t is the corresponding semigroup, which corresponds to the conservation of mass of the corresponding forward Kolmogorov equation and hence the impossibility of escape to infinity of the associated Hunt process.

It can be shown that a Dirichlet form is stochastically complete if and only if there exists a there exists a sequence $\{f_m\}_{m\in N} \in D$ such that $0 \leq f_m \leq 1$, $f_m(x) \to 1$ for π-a.e. $x \in \mathcal{X}$ as $m \to \infty$, and finally that for any $g \in D \cap L^1(\mathcal{X},\pi)$

$$\lim_{m\to\infty} \mathcal{E}(f_n, g) = 0.$$

Moreover, it is shown in [GHM12] that there exist tight bounds on the growth of $\pi(B_r(x))$ for balls $B_r(x)$ around x with radius r that ensure $(\mathcal{E}, D)$ is stochastically complete, providing examples of stochastically incomplete graphs for which these bounds are not satisfied. In particular, the bounds fail if $d_{\mathcal{X}}$ is not a Heine–Borel but for example a bounded metric over a non-compact infinite space $\mathcal{X}$, in which case $\pi(B_r(x)) = +\infty$ for sufficiently large r.

Now note that any growth bounds on $\pi(B_r(x))$ hold if π is finite, and in fact one can take a sequence $f_m := \xi_m$ constructed in Remark 14.6 satisfying (14.9). In addition, if instead we have the stronger condition

$$\sup_{x\in\mathcal{X}} \int_{\mathcal{X}} (d_{\mathcal{X}}(x,y) \vee 1)\kappa(x, \mathrm{d}y) < \infty. \tag{14.34}$$

then $(\xi_m)_{m\in\mathbb{N}}$ satisfies (14.6), and the approximation holds even for σ-finite π.

It should be noted that for simple birth-death chains the above metric formulation is a bit overkill, and instead very precise conditions exist to characterize uniqueness and (non-)explosion. This is for example discussed in [MM20] and applied to to the Chemical Master Equation for chemical reaction networks, which is similar to our Bolker–Pacala model of Part II.B in the case of a finite trait space.

14.4.2 Chain rule for bounded u and regularization

In light of the previous section, it should be said that if π is finite the bounds (14.6), (14.9) are not necessary for stochastic completeness, since one only would require that

$$\sup_m \int_E |\overline{\nabla}\xi_m|^2 \pi(\mathrm{d}x)\kappa(x,\mathrm{d}y). \tag{14.35}$$

The latter is a strictly weaker assumption, and does not require the existence of an adapted Heine–Borel metric. The reason why in the chain rule we still require (14.6) or (14.9) is because $\mathcal{R}$, $\mathcal{D}$ in fact only depend on the kernel κ instead of π, and

$$\limsup_{n\to\infty}\limsup_{m\to\infty}\frac{1}{2}\int_{\mathcal{X}} |\overline{\nabla}\xi_m\phi_n'(\mathrm{d}\rho/\mathrm{d}\pi)|\,\mathrm{d}|j_{net}| \le \mathcal{R}(\rho,j)+\mathcal{D}(\rho), \tag{14.36}$$

holds even if π is merely σ-finite. That is, we are not using the finiteness of π explicitly, except in defining and bounding $\mathcal{E}\mathrm{nt}(\rho|\pi)$, which can be modified as in Section 12.5.2.

However, if we assume for the curve $(\rho,j)\in\mathsf{CE}_a^*$ a stronger regularity condition, namely that $\sup_t\|u_t\|_\infty<\infty$, then it is clear from the proof of Proposition 14.15 that one only requires the bound (14.35). To prove the chain rule for an arbitrary curve (ρ,j) one would need an additional regularization step via the sequence $\{(\rho^\varepsilon,j^\varepsilon)\}_\varepsilon\subset\mathsf{CE}_a^*$ with $\sup_t\|u_t^\varepsilon\|_\infty<\infty$ for all ε, with the following limsup estimates

$$\begin{aligned}\lim_{\varepsilon\to0}\mathcal{E}\mathrm{nt}(\rho_t^\varepsilon|\pi)&\le\mathcal{E}\mathrm{nt}(\rho|\pi),\ \text{ for all } t\in[0,T]\\ \lim_{\varepsilon\to0}\int_0^T\mathcal{R}(\rho_t^\varepsilon,j_t^\varepsilon)\,\mathrm{d}t&\le\int_0^T\mathcal{R}(\rho_t,j_t)\,\mathrm{d}t,\\ \lim_{\varepsilon\to0}\int_0^T\mathcal{D}(\rho_t^\varepsilon)\,\mathrm{d}t&\le\int_0^T\mathcal{D}(\rho_t)\,\mathrm{d}t.\end{aligned} \tag{14.37}$$

Unfortunately, preliminary calculations suggest that typical approaches such as employing regularizing semigroups (e.g. [Erb14]) or using truncations (e.g. [HPST20]) might not work in the case of unbounded kernels without additional assumptions.

14.4.3 Singular kernels

Throughout we considered unbounded but finite kernels, but in fact one can generalize much of the framework to *singular* kernels, building on for example the work of [Erb14] for Lévy-kernels over $\mathbb{R}^d$ and generalizing this to suitable metric spaces $(\mathcal{X}, d)$. Namely, assume again that $\mathcal{X}$ is locally compact and d is adapted in the sense of (14.33), where $\{\kappa(x,\cdot)\}_{x\in\mathcal{X}} \subset \mathcal{M}^+_{\mathrm{loc}}(\mathcal{X})$ is now no longer considered to be finite. Then the statements of [GHM12] on stochastic completeness and the regularity of the associated Dirichlet form $(\mathcal{E}, D)$ still apply.

Moreover, one can redefine the continuity equation CE^*_a in the sense that fluxes $j \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{E}/\mathbb{D})$ (with $\mathbb{D}$ the diagonal) are in a weighted TV-space with the norm $\int d^2 \wedge 1 \,\mathrm{d}j$. It is straightforward to show that finite $\mathcal{R}(\rho, j)$ in fact implies that for the net-fluxes $\int d \wedge 1 \,\mathrm{d}|j^{\mathrm{net}}| < \infty$ and hence the continuity equation is well-defined for all $f \in C_c(\mathcal{X}) \cap BL(\mathcal{X})$. Preliminary approximation arguments suggest that the chain rule still might hold for curves (ρ, j) such that $\sup_{t\in[0,T]} \|u_t\|_\infty < \infty$ and $\mathcal{E}(\sqrt{u_t}) < \infty$ for all $t \in [0, T]$. To extend this to more general curves one would have to do an extra regularization step and obtain (14.37).

In particular, this is the case if $\mathcal{X} = \mathbb{R}^d, \mathbb{T}^d$ with $\kappa(x, \mathrm{d}y)$ translation invariant, or if the kernel stems from a Cayley graph and is invariant under the corresponding symmetry group. Namely, in those settings one can simply use convolutions, as done for example for $\mathbb{R}^d$ in [Erb14].

Part II.B

Variational convergence for population dynamics

Chapter 15

Introduction

In the final part of this thesis, we come back to the setting of weakly interacting particle systems, investigating problems stemming from theoretical biology and ecology involving birth, death and mutation. The main goal is to reconstruct macroscopic equations from microscopic models, but instead of establishing large deviation principles directly as in Part I.B we consider its formal counterpart: proving convergence of variational structures.

The dynamics of these particle systems consist of particles being created, dispersed, and annihilated, and can be described as jump processes on the space of finite nonnegative measures over a *trait space*. Such measure-valued jump processes are expected to converge under a mean-field scaling and in the large-population limit to the *mean-field equation*, which can be represented as

$$\partial_t u_t = b^+[u_t] - b^-[u_t]u_t. \tag{15.1}$$

Here $u_t(x)$ represents the density of particles around the point $x \in \mathcal{T}$ with *trait space* $\mathcal{T}$, and b^+, b^- are density-dependent birth/death rates.

This connection has been made rigorous in various settings, as will be discussed in detail in Section 15.5, and historically involved the trait space being discrete or a closed set of $\mathbb{R}^d$. Of particular interest in this case is the *Bolker–Pacala–Dieckmann–Law* (BPDL) model, where

$$b^+[u](x) := \int_{\mathbb{R}^d} m(y,x)u(y)\,\mathrm{d}y \qquad b^-[u](x) := \int_{\mathbb{R}^d} c(x,y)u(y)\,\mathrm{d}y,$$

with the mutation rate $m(x, y)$ the rate that a particle with trait x creates a new particle with trait y, and the competition rate $c(x, y)$ is the rate with which particles with traits x, y compete for resources until one of them perishes.

However, in recent years there has been considerable interest in allowing for dynamics involving multiple species and combinations of discrete and continuous traits. Therefore, in this thesis, we consider the case of $\mathcal{T}$ an arbitrary compact Polish and the mutation/competition rates m, c merely bounded functions. We prove the mean-field limit and in addition establish entropic propagation of chaos, which controls the discrepancy between the microscopic and macroscopic models in a precise sense. To the author's knowledge, this is the first convergence result under such general assumptions.

To do so, we first apply the framework introduced in Part II.A, and equip the corresponding forward Kolmogorov equation (FKE) with a variational structure and a generalized Energy-Dissipation Principle (EDP) in the sense of Part II.A, after choosing a suitable Poisson point measure as a reference measure. We then show the EDP-convergence of these structures under a mean-field scaling and the large-population limit, where the product structure of the reference measure plays a strong role.

The limiting evolution equation is the Liouville equation corresponding to the mean-field equation, namely a transport equation that describes the evolution of the law of a process that follows deterministic mean-field dynamics but for possibly random initial data. This connection between the Liouville equation and the mean-field equation is made rigorous with the help of a new superposition principle.

In particular, we deduce that the laws determined by the FKE equation concentrate around the solution of the mean-field equation, which due to the convergence of the associated free energies translates into an entropic propagation of chaos result.

15.1 Measure-valued population dynamics and mean-field limits

We consider the forward Kolmogorov equation that corresponds to a generalized version of the BPDL model. In its classical form, the Bolker–Pacala model is a purely spatially-structured microscopic model for a population of plants involving the birth, dispersal, and either natural death or death by competition for resources, and can be modelled as a jump process in the space of nonnegative measures over $\mathbb{R}^d$ [BP97]. However, in certain models of adaptive evolution it is the mutation of traits that play a role, instead of spatial evolution (see [LD00, CFM06, CFM08]). Moreover, if one wants to model multiple interacting species or marked configuration spaces, more general spaces than $\mathbb{R}^d$ are needed [KLU99, FKK21]).

Therefore, let the trait space $\mathcal{T}$ be an arbitrary Polish space. We model the dynamics at any time t as an interacting particle system with particles with labels $A_t^1, \dots, A_t^{N_t}$ and traits $X_t^1, \dots, X_t^{N_t} \in \mathcal{T}$, where the number of particles N_t at time t is not fixed since particles can be removed from and added to the system.

Moreover, fix a positive parameter n, and let ν_t^n be the rescaled empirical measure at time t, defined as

$$\nu_t^n := \frac{1}{n}\sum_{i=1}^{N(t)} \delta_{X_t^i}.$$

Here $\nu_t^n \in \Gamma := \mathcal{M}^+(\mathcal{T})$, with Γ the space of nonnegative measures over the trait space $\mathcal{T}$. Note the difference with Part I.B where we considered the empirical measure over path space. Now, instead of u-dependent rates, we consider measure-dependent birth/death rates $b^\pm : \Gamma \to \mathcal{B}_b(\mathcal{T})$, $n > 0$ a positive parameter, and $\gamma \in \mathcal{M}_{\text{loc}}^+(\mathcal{T})$ a nonnegative reference measure.

Then, with a little abuse of notation, the dynamics can be described as follows:

- For each trait $x \in \mathcal{T}$ there is a birth clock with rate $nb^+[\nu_t^n](x)\gamma(x)$. If the birth clock rings, a new particle is added with trait x.
- For each particle, with trait $x \in \mathcal{T}$, there is a *death* clock with rate $nb^-[\nu_t^n](x)$. If the death clock rings, the particle is deleted.

The parameter $n > 0$ is called the *system size*, in the sense that the scalings guarantee that if the amount of particles in the system is of the order of n, the total rate of created or deleted particles is of the same order.

Similar as in Part I.B we can consider quite general measure-dependent birth and death rates, but for our guiding example, our generalization of the BPDL model, we take

$$b^+[\nu](x) := \int_{\mathcal{T}} m(y,x)\nu(\mathrm{d}y), \qquad b^-[\nu](x) := \int c(x,y)\nu(\mathrm{d}y), \tag{15.2}$$

with mutation rate $m(x, y)$ and competition rate $c(x, y)$.

Alternatively, we can describe these dynamics in the form of reacting particles. Namely, setting $m(x, y) := b(x)d(x, y)$, then with a little of abuse of notation we have

$$\begin{aligned} A_t^i &\to A_t^i + A_t^{N_t+1}, &&\text{with rate} \quad m\left(X_t^i, X_t^{N_1+1}\right)\gamma\left(X_t^{N_1+1}\right), \\ A_t^i + A_t^j &\to A_t^j, &&\text{with rate} \quad n^{-1}c\left(X_t^i, X_t^j\right). \end{aligned} \tag{15.3}$$

Instead of looking at the individual traits of the particles, it is common to only consider the measure-valued process ν_t^n. Under suitable assumptions the infinitesimal generator Q_n of this process is given for suitable $F \in C_b(\Gamma)$

$$(Q_nF)(\nu) = n\int_{\mathcal{T}} \left(F\left(\nu + \tfrac{1}{n}\delta_x\right) - F(\nu)\right)\chi^+[\nu](\mathrm{d}x) + n\int_{\mathcal{T}} \left(F\left(\nu - \tfrac{1}{n}\delta_x\right) - F(\nu)\right)\chi^-[\nu](\mathrm{d}x)$$

where $\chi^{\pm}[\nu] \in \Gamma$ are given by

$$\chi^{+}[\nu](\mathrm{d}x) := b^{+}[\nu](x)\gamma(\mathrm{d}x), \quad \chi^{-}[\nu](\mathrm{d}x) := b^{-}[\nu](x)\nu(\mathrm{d}x). \tag{15.4}$$

For simplicity we will also denote $b^{\pm}_{\nu}(x)$, $\chi^{\pm}_{\nu}(\mathrm{d}x)$. The law of the process is now given by the corresponding forward Kolmogorov equation

$$\partial_t \mathsf{P}^n_t = Q^*_n \mathsf{P}^n_t, \qquad \mathsf{P}^n_t \in \mathcal{P}(\Gamma). \tag{$\mathsf{FKE_n}$}$$

As the system size n converges to infinity, one would expect convergence of the measure-valued process ν^n_t to the mean-field equation

$$\partial_t \nu_t = \chi^{+}_{\nu} - \chi^{-}_{\nu}, \qquad \nu_t \in \Gamma. \tag{MF}$$

Note that if we take $\nu_t(\mathrm{d}x) = u_t(x)\gamma(\mathrm{d}x)$ and redefine the birth/death rates appropriately this is precisely equivalent to (15.1).

However, similar as in Part II.A, we do not consider the measure-valued process itself, but take the forward Kolmogorov equation ($\mathsf{FKE_n}$) as a starting point. We then show convergence to the mean-field equation in the sense that $\mathsf{P}^n_t \to \delta_{\nu_t}$ narrowly on $\mathcal{P}(\Gamma)$ under suitable initial data. Throughout we assume the following:

Assumption 15.1. *The trait space $\mathcal{T}$ is a compact Polish space, and moreover*

$\gamma \in \Gamma$	*(reference measure with finite mass)*
$m, c \in \mathcal{B}^{+}_{b}(\mathcal{T} \times \mathcal{T})$	*(bounded rates)*

Henceforth we equip the space Γ with the narrow topology. Moreover, the assumptions above guarantee that the Forward Kolmogorov equation falls into the setting of II.A, with reference measure $\Pi_n \in \mathcal{P}(\Gamma)$, which stems from a Poisson measure with intensity γ. Namely, set

$$\mathcal{P}\left(\coprod_{N\geq 0} \mathcal{T}^N\right) \ni \pi_n := \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^{\infty} \frac{n^N}{N!} \gamma^{\otimes N},$$

and consider the rescaled empirical measure mapping $L_n : \coprod_{N\geq 0} \mathcal{T}^N \to \Gamma$,

$$L_n(x_1, \ldots, x_N) := \frac{1}{n} \sum_{i=1}^{N} \delta_{x_i}, \quad N \in \mathbb{N}, \tag{15.5}$$

with the convention $L_n(\emptyset) = 0$. Then Π_n is defined as

$$\Pi_n := (L_n)_{\#} \pi_n.$$

This allows us to write the forward Kolmogorov equation as an EDP-solution with the relative entropy with respect to Π_n as driving energy functional, and equip it with a corresponding variational structure.

We will show, under suitable assumptions on the initial data, that as $n \to \infty$ the Forward Kolmogorov equation converges to the following *Liouville equation*

$$\partial_t \mathsf{P}_t + \mathrm{div}_\Gamma \left(\mathsf{P}_t \left(\chi_\nu^+ - \chi_\nu^-\right)\right) = 0, \qquad \mathsf{P}_t \in \mathcal{P}(\Gamma). \tag{Li}$$

It is a transport equation that can be interpreted as the lifting of mean-field dynamics in Γ to evolutions in $\mathcal{P}(\Gamma)$, and describes the evolution of the law of random measures ν_t that all satisfy (MF). In particular, if ν_t a solution of (MF) then $\mathsf{P}_t := \delta_{\nu_t}$ is itself a solution of (Li).

Letting $V[\nu] = \chi_\nu^+ - \chi_\nu^-$, we can therefore represent part of our results in Figure 15.1.

$$\begin{array}{llcll} (\mathsf{FKE_n}) & \partial_t \mathsf{P}^n_t = Q_n^* \mathsf{P}^n_t & \xrightarrow{n\to\infty} & (\mathsf{Li}) & \partial_t \mathsf{P}_t + \mathrm{div}_\Gamma \left(\mathsf{P}_t V[\nu]\right) = 0 \\ & & & & \qquad\qquad \big\downarrow\!\big\uparrow \\ & & & (\mathsf{MF}) & \qquad \partial_t \nu_t = V[\nu_t] \end{array}$$

Figure 15.1: Convergence in the large-population limit

This convergence is a direct consequence of the convergence of the dissipation structures, which we will describe below.

15.2 Variational and dissipation structures

The first main result of Part II.B concerns the variational formulation of the equations $(\mathsf{FKE_n})$, (MF), (Li), their specific dissipation structure, and the equivalence and relation to certain rate functionals stemming from large deviations. Due to the lack of detailed balance and the need to control the convergence of the driving energies in the large population limit, we adopt a similar strategy as in Part II.A. Namely, in all three cases, an Energy-Dissipation Principle in the generalized sense of Part II.A is satisfied, in that the driving energy functional stems from the relative entropy with respect to a chosen reference measure.

However, the precise continuity function relating particle densities or laws ρ_t and *generalized fluxes* j_t will depend on the setting. Therefore, since we will repeat the same concept three times on different levels and for different spaces, let us make the general and abstract concepts clear:

Definition 15.2. Given a *free energy* functional $\mathcal{F}(\rho)$, a *dissipation potential* $\mathcal{R}(\rho, j)$, a *Fisher information* functional $\mathcal{D}(\rho)$, and a linear operator B with dual B^*, we consider pairs of curves (ρ, j) satisfying the continuity equation

$$\partial_t \rho_t - B^* j_t = 0, \quad \text{for a.e. } t \in [0, T], \tag{CE}$$

and define the *EDP-functional*

$$\mathcal{J}(\rho, j) := \int_0^T \mathcal{R}(\rho_t, j_t)\,\mathrm{d}t + \mathcal{F}(\rho_T) - \mathcal{F}(\rho_0) + \int_0^T \mathcal{D}(\rho_t)\,\mathrm{d}t.$$

Moreover, a *EDP-solution* is a pair $(\hat\rho, \hat\jmath)$ satisfying (CE) with $I(\hat\rho, \hat\jmath) = 0$. □

In all three examples the generalized fluxes j consist of two parts: j^+ and j^-, corresponding to birth and death. The continuity equations depend on the setting and are summarized in Table 15.1, with $\mathcal{M}^+_{\mathrm{loc}}$ as the space of nonnegative Radon measures.

Remark 15.3. Note that the EDP-solution $(\hat\rho, \hat\jmath)$ is the null-minimizer of $\mathcal{J}$, and satisfies the *energy-dissipation balance*

$$\mathcal{F}(\hat\rho_T) + \int_0^T \left(\mathcal{R}(\hat\rho_t, \hat\jmath_t) + \mathcal{D}(\hat\rho_t)\right)\,\mathrm{d}t = \mathcal{F}(\hat\rho_0).$$

Moreover, for small $T \ll 1$ one would expect

$$\mathcal{J} \approx \mathcal{R}(\hat\rho, \hat\jmath) + \langle \hat\jmath, B\,\partial_\rho \mathcal{F}\rangle + \mathcal{D}(\hat\rho).$$

In light of the relation to minimizing movement schemes as in Section 12.5.3, a formal minimization procedure provides the EDP-solution

$$\begin{aligned} \partial_t \hat\rho_t - B^* \hat\jmath_t &= 0 \\ \hat\jmath_t &= (\partial_2 \mathcal{R}^*)(\hat\rho_t, -B\,\partial_\rho \mathcal{F}), \end{aligned}$$

and that along the solution

$$\mathcal{D}(\hat\rho_t) = \mathcal{R}^*(\hat\rho_t, -B\,\partial_\rho \mathcal{F}), \quad \partial_t \mathcal{F}(\hat\rho_t) = -\left(\mathcal{R}(\hat\rho_t, \hat\jmath_t) + \mathcal{D}(\hat\rho_t)\right) \tag{15.6}$$

where $\mathcal{R}^*(\rho, w)$ is the dual of the dissipation potential $\mathcal{R}$.

These identities will indeed shown to be valid in the settings that we consider. □

While in all cases the dissipation potential $\mathcal{R}$ is assumed to be nonnegative, this is *not* the case for the Fisher information and, in particular, the total dissipation of $\mathcal{F}$ along the solution, $\mathcal{R} + \mathcal{D}$, is allowed to be nonnegative. This goes against the

	CE	ρ	$j=(j^+,j^-)$	BF	
(MF)	$(\mathscr{C}\mathscr{E})$	$\nu\in\Gamma$	$(\lambda^+,\lambda^-)\in\Gamma^2$	$(F,-F)$	
($\mathsf{FKE_n}$)	(CE_n)	$\mathsf{P}\in\mathcal{P}(\Gamma)$	$(\mathsf{J}^+,\mathsf{J}^-)\in\mathcal{M}^+_{\mathrm{loc}}(\Gamma\times\mathcal{T})^2$	$(\overline{\nabla}^{n,+}F,\overline{\nabla}^{n,-}F)$	(15.9)
(Li)	(CE_∞)	$\mathsf{P}\in\mathcal{P}(\Gamma)$	$(\mathsf{J}^+,\mathsf{J}^-)\in\mathcal{M}^+_{\mathrm{loc}}(\Gamma\times\mathcal{T})^2$	$(\mathrm{grad}_\Gamma F,-\mathrm{grad}_\Gamma F)$	(15.13)

Table 15.1: Continuity equations

usual definition of Energy-Dissipation Principles, since the free energy no longer actually dissipates, but as in Part II.A we will still adopt the same language, due to the similarity of the variational structures.

Let us recall the Hellinger distance $H(\mu_1,\mu_2)$, see (11.9), and $\mathcal{E}\mathrm{nt}(\mu_1|\mu_2)$, the relative entropy of μ_1 with respect to μ_2 for two (possible infinite) locally finite Borel measures μ_1,μ_2:

$$\mathcal{E}\mathrm{nt}(\mu_1|\mu_2):=\begin{cases}\displaystyle\int\phi\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_2}\right)\mathrm{d}\mu_2, & \text{if } \mu_1\ll\mu_2,\\ +\infty, & \text{otherwise,}\end{cases}$$

where

$$\phi(s)=s\log s-s+1.$$

Moreover, let us introduce

$$\begin{aligned}\bar\chi^{n,+}_\nu(\mathrm{d}x)&=b^+[\nu-\tfrac{1}{n}\delta_x](x)\nu(\mathrm{d}x) & \bar\chi^{n,-}_\nu(\mathrm{d}x)&=b^-[\nu+\tfrac{1}{n}\delta_x](x)\gamma(\mathrm{d}x),\\ \bar\chi^{+}_\nu(\mathrm{d}x)&=b^+[\nu](x)\nu(\mathrm{d}x) & \bar\chi^{-}_\nu(\mathrm{d}x)&=b^-[\nu](x)\gamma(\mathrm{d}x),\end{aligned}\tag{15.7}$$

These will play a similar role as the backward kernels of Part II.A induced by the fixed reference measure $\gamma\in\Gamma$, and in fact the backward kernels for the FKE under the reference measure $\Pi\in\mathcal{P}(\Gamma)$ will be determined by them.

While the full technical details contained in Theorems 16.8, 17.11 and 18.9, we now discuss subsequently the mean-field equation, the Forward Kolmogorov equation, and the Liouville equation.

Mean-field equation

We consider triples $(\nu,\lambda^+,\lambda^-)$, with $\nu_t,\lambda^\pm_t\in\Gamma$, satisfying the mean-field continuity equation

$$\partial_t\nu_t=\lambda^+_t-\lambda^-_t.\tag{$\mathscr{C}\mathscr{E}$}$$

Let us define the dissipation potential $\mathcal{R}_{MF}$, free energy $\mathcal{F}_{MF}$, Fisher information $\mathcal{D}_{MF}$, and Lagrangian $\mathcal{L}_{MF}$ as

$$\begin{aligned} \mathcal{R}_{MF}(\nu, \lambda^+, \lambda^-) &:= \mathcal{E}\mathrm{nt}(\lambda^+|\theta_\nu^+) + \mathcal{E}\mathrm{nt}(\lambda^-|\theta_\nu^-), \\ \mathcal{F}_{MF}(\nu) &:= \tfrac{1}{2}\mathcal{E}\mathrm{nt}(\nu|\gamma), \\ \mathcal{D}_{MF}(\nu) &:= H^2(\chi_\nu^+, \bar{\chi}_\nu^+) + H^2(\chi_\nu^-, \bar{\chi}_\nu^-) + \tfrac{1}{2}\left(\chi_\nu^+ + \chi_\nu^- - \bar{\chi}_\nu^+ - \bar{\chi}_\nu^-\right) \\ \mathcal{L}_{MF}(\nu, \lambda^+, \lambda^-) &= \mathcal{E}\mathrm{nt}(\lambda^+|\chi_\nu^+) + \mathcal{E}\mathrm{nt}(\lambda^-|\chi_\nu^-), \end{aligned}$$

where $\theta_\nu^\pm$ are the geometric means

$$\theta_\nu^\pm := \sqrt{\chi_\nu^\pm \bar{\chi}_\nu^\pm}.$$

Moreover, let the corresponding EDP-functional I_{MF} be given as

$$\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) := \int_0^T \mathcal{R}_{MF}(\nu_t, \lambda_t^+, \lambda_t^-)\,\mathrm{d}t + \mathcal{F}_{MF}(\nu_T) - \mathcal{F}_{MF}(\nu_0) + \int_0^T \mathcal{D}_{MF}(\nu_t)\,\mathrm{d}t,$$

and the rate functional $\mathcal{I}_{MF}$ as

$$\mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-) = \int_0^T \mathcal{L}_{MF}(\nu_t, \lambda_t^+, \lambda_t^-)\,\mathrm{d}t.$$

We then have the following result.

Theorem 15.4 (Mean-field, cf. Theorem 16.8). *For any $(\nu, \lambda^+, \lambda^-) \in \mathscr{CE}$ with $\mathcal{F}(\nu_0) < \infty$*

$$\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) = \mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-). \tag{15.8}$$

In particular, $\mathcal{J}$ is nonnegative. Moreover, for any ν_0 with $\mathcal{F}_{MF}(\nu_0) < \infty$ a unique EDP-solution $(\hat{\nu}, \hat{\lambda}^+, \hat{\lambda}^-)$ exists, with $\hat{\nu}_t$ equal to the unique strong solution to (MF) *and $\hat{\lambda}_t^\pm = \chi_{\nu_t}^\pm$ for almost every $t \in [0, T]$.*

As in Part II.A for jump processes, the intuition underlying the equivalence (15.8) and the decomposition of the mean-field Lagrangian $\mathcal{L}_{MF}$ stems again from a symmetrization argument, which we will discuss in detail in Section 16.3. In essence, where for jump processes reversing time shows up as swapping x and y positions for edge fluxes, here it manifests as swapping birth and death rates, and the backward kernel for a pure birth process defined by χ is a pure death process determined by $\bar{\chi}^+$, and vice versa. This will be clear from the microscopic view, i.e. from the variational structure for the Forward-Kolmogorov equation below, and indeed the structure for the mean-field equation can be seen as a macroscopic consequence of this fact.

In the case of detailed balance, this is similar to the decompositions of LDP rate functionals for interacting jump processes or chemical reactions as in [PR19, BBBO21], and indeed, there is a strong connection to the variational formulations for jump processes arising from the large deviations of fluxes as seen and the equivalence of the EDP-functional to the expected rate functional, which we will discuss briefly in Appendix G.

Forward Kolmogorov equation

We consider triples $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$, with $\mathsf{P}_t \in \mathcal{P}(\Gamma)$ and $\mathsf{J}_t^\pm \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma \times \mathcal{T})$, satisfying the continuity equation

$$\langle F, \partial_t \mathsf{P}_t \rangle = \langle \overline{\nabla}^{n,+} F, \mathsf{J}_t^+ \rangle + \langle \overline{\nabla}^{n,-} F, \mathsf{J}_t^- \rangle, \qquad \forall F \in C_c(\Gamma), \tag{CE$_n$}$$

where

$$(\overline{\nabla}^{n,\pm} F)(\nu, x) := n \left(F(\nu \pm \tfrac{1}{n}\delta_x) - F(\nu) \right). \tag{15.9}$$

Let us define the n-dependent Fisher information $\mathcal{D}_n$ as stated in Definition 17.6, free energy

$$\mathcal{F}_n(\mathsf{P}) := \frac{1}{2n} \mathcal{E}\mathrm{nt}(\mathsf{P}|\Pi_n),$$

and dissipation potential $\mathcal{R}_n$ and Lagrangian $\mathcal{L}_n$

$$\begin{aligned} \mathcal{R}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) &:= \mathcal{E}\mathrm{nt}(\mathsf{J}^+|\Theta_{\mathsf{P}}^{n,+}) + \mathcal{E}\mathrm{nt}(\mathsf{J}^-|\Theta_{\mathsf{P}}^{n,-}), \\ \mathcal{L}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) &:= \mathcal{E}\mathrm{nt}(\mathsf{J}^+|\mathsf{P}\chi_\nu^+) + \mathcal{E}\mathrm{nt}(\mathsf{J}^-|\mathsf{P}\chi_\nu^-), \end{aligned}$$

where $(\mathsf{P}\chi_\nu^\pm)(\mathrm{d}\nu, \mathrm{d}x) := \chi_\nu^\pm(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu)$ and, with a little abuse of notation (see (17.19)),

$$\Theta_{\mathsf{P}}^{n,\pm}(\nu, x) := \sqrt{\left(\mathsf{P}(\nu)\chi_\nu^\pm\right)\left(\mathsf{P}(\nu \pm \tfrac{1}{n}\delta_x)\bar{\chi}^{n,\pm}[\nu \pm \tfrac{1}{n}\delta_x]\right)}.$$

Moreover, let the corresponding EDP-functional I_n be given as

$$\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{R}_n(\mathsf{P}_t, \mathsf{J}_t^+, \mathsf{J}_t^-)\,\mathrm{d}t + \mathcal{F}_n(\mathsf{P}_T) - \mathcal{F}_n(\mathsf{P}_0) + \int_0^T \mathcal{D}_n(\mathsf{P}_t)\,\mathrm{d}t,$$

and the rate functional

$$\mathcal{I}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{L}_n(\mathsf{P}_t, \mathsf{J}_t^+, \mathsf{J}_t^-)\,\mathrm{d}t.$$

Theorem 15.5 (Forward Kolmogorov, cf. Theorem 17.11)**.** *For any* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$ *with* $\mathcal{F}_n(\mathsf{P}) < \infty$ *and*

$$\sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^2 \, \mathrm{d}\mathsf{P}_t < \infty, \tag{15.10}$$

we have the equivalence

$$\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = \mathcal{I}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-). \tag{15.11}$$

In particular, $\mathcal{J}$ *is nonnegative for such curves. Moreover, for any* P_0 *with* $\mathcal{F}_n(\mathsf{P}_0) < \infty$ *and*

$$\int_\Gamma \nu(\mathcal{T})^2 \, \mathrm{d}\mathsf{P}_0 < \infty, \tag{15.12}$$

a unique EDP-solution $(\hat{\mathsf{P}}, \hat{\mathsf{J}}^\pm)$ *exists, with* $\hat{\mathsf{P}}_t$ *equal to the unique weak solution to* $(\mathsf{FKE_n})$ *and* $\hat{\mathsf{J}}_t^\pm = \hat{\mathsf{P}}_t \chi_\nu^\pm$ *for almost every* $t \in [0, T]$.

Similar to the mean-field case, the dissipation potential consists of relative entropies with respect to geometric averages, now of forward and backward rates along a transition $\nu \to \nu \pm \frac{1}{n}\delta_x$.

Moreover, note the difference in the continuity equations CE_n and that of Part II.A, where the analogue of the latter would be for edge fluxes $j(\mathrm{d}\nu, \mathrm{d}\eta)$. Indeed, we will show that one can go from formulation to the other. The benefit of writing the continuity way as done for CE_n is that $\mathsf{J}^\pm(\mathrm{d}\nu, \mathrm{d}x)$ are now rescaled to be *mass fluxes*, i.e. they signify the mass of particles created or annihilated around x when the rescaled empirical measure of the particle system is around ν. This allows us to establish convergence in the large population limit of fluxes $\mathsf{J}^{n,\pm}$ and not just the laws P^n, as discussed in our convergence result below.

Finally, the moment conditions are necessary since the underlying jump kernel is unbounded, and as we saw in Part II.A this can imply unboundedness of the Fisher information if the reference measure is not the invariant measure. Since the birth/death kernels $\chi^\pm[\nu](\mathcal{T})$ are bounded by a constant times $1 + \nu(\mathcal{T})^2$ this translates into quadratic moment bounds in terms of mass to apply the framework of Chapter 13.

Liouville equation

Let $\mathrm{Cyl}_c(\Gamma)$ be the space of compactly supported smooth cylinder functions of the form

$$F(\nu) = g\left(\langle 1, \nu\rangle, \langle f_1, \nu\rangle, \dots, \langle f_m, \nu\rangle\right), \qquad g \in C_c^\infty(\mathbb{R}^m),\ m \in \mathbb{N},$$

where $f_1, \dots, f_m \in C_b(\mathcal{T})$, and grad_Γ is the gradient defined by

$$\mathrm{grad}_\Gamma F(\nu, x) = (\nabla g)\big(\langle 1, \nu\rangle, \langle f_1, \nu\rangle, \dots, \langle f_m, \nu\rangle\big) \cdot (1, f_1(x), \dots, f_m(x))^\top. \tag{15.13}$$

We consider triples $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$, with $\mathsf{P}_t \in \mathcal{P}(\Gamma)$, $\mathsf{J}^\pm \in \mathcal{M}_{\mathrm{loc}}(\Gamma \times \mathcal{T})$, satisfying the continuity equation

$$\langle F, \partial_t \mathsf{P}_t\rangle = \langle \mathsf{grad}_\Gamma F, \mathsf{J}_t^+\rangle - \langle \mathsf{grad}_\Gamma F, \mathsf{J}_t^-\rangle, \quad \forall F \in \mathrm{Cyl}_c(\Gamma). \tag{CE_∞}$$

Let us define the Fisher information $\mathcal{D}_\infty$ as stated in Definition 18.6, free energy

$$\mathcal{F}_\infty(\mathsf{P}) := \frac{1}{2}\int_\Gamma \mathcal{E}\mathrm{nt}(\nu|\gamma)\,\mathrm{d}\mathsf{P},$$

and dissipation potential $\mathcal{R}_\infty$ and Lagrangian $\mathcal{L}_\infty$

$$\begin{aligned}
\mathcal{R}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) &:= \mathcal{E}\mathrm{nt}(\mathsf{J}^+|\Theta_\mathsf{P}^\infty) + \mathcal{E}\mathrm{nt}(\mathsf{J}^-|\Theta_\mathsf{P}^\infty), \qquad \Theta_\mathsf{P}^\infty(\mathrm{d}\nu, \mathrm{d}x) := \theta_\nu(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu),\\
\mathcal{L}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) &:= \mathcal{E}\mathrm{nt}(\mathsf{J}^+|(\mathsf{P}\chi_\nu^+)) + \mathcal{E}\mathrm{nt}(\mathsf{J}^-|(\mathsf{P}\chi_\nu^-)).
\end{aligned}$$

Moreover, let the corresponding EDP-functional I_n be given as

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}_t^+, \mathsf{J}_t^-)\,\mathrm{d}t + \mathcal{F}_\infty(\mathsf{P}_T) - \mathcal{F}_\infty(\mathsf{P}_0) + \int_0^T \mathcal{D}_\infty(\mathsf{P}_t)\,\mathrm{d}t,$$

and the rate functional

$$\mathcal{I}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{L}_\infty(\mathsf{P}_t, \mathsf{J}_t^+, \mathsf{J}_t^-)\,\mathrm{d}t.$$

Theorem 15.6 (Liouville, cf. Theorem 18.9)**.** *For any* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$ *with* $\mathcal{F}_\infty(\mathsf{P}_0) < \infty$ *and*

$$\sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^2\,\mathrm{d}\mathsf{P}_t < \infty, \tag{15.14}$$

we have the equivalence

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = \mathcal{I}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-). \tag{15.15}$$

In particular, $\mathcal{J}_\infty$ *is nonnegative for such curves. Moreover, for any* P_0 *with* $\mathcal{F}_\infty(\mathsf{P}_0) < \infty$ *and*

$$\int_\Gamma \nu(\mathcal{T})^2\,\mathrm{d}\mathsf{P}_0 < \infty, \tag{15.16}$$

a unique EDP-solution $(\hat{\mathsf{P}}, \hat{\mathsf{J}}^\pm)$ *exists, with* $\hat{\mathsf{P}}_t$ *equal to the unique solution to* (Li) *and* $\hat{\mathsf{J}}_t^\pm = \hat{\mathsf{P}}_t\chi_\nu^\pm$ *for almost every* $t \in [0, T]$.

To prove Theorem 15.6, we will show in addition that for any $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ such that $I_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) < \infty$, there exists (with a little abuse of notation) a Borel probability measure Ω over curves satisfying the mean-field continuity equation ($\mathscr{C}\mathscr{E}$) such that for all t the time-marginals $(e_t)_\#\Omega$ are equal to P_t, and

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int \mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-)\, \mathrm{d}\Omega. \tag{15.17}$$

The statement of (15.17) is the aforementioned superposition principle, which is a modified version of the superposition principle [AT14] in metric measure spaces, and the ones used in [EFLS16], [Erb16]. It allows one to essentially jump back and forth between the Liouville equation and the mean-field dynamics, and, in particular, provides us with the non-negativity of $\mathcal{J}_\infty$ and the uniqueness of EDP-solutions.

15.3 Convergence results

Our final and most important result is that the above structures converge in the sense of *EDP-convergence* (e.g. see [LMPR17, PS22]), a generalization of the evolutionary Γ-convergence approach stated by [SS04, Ser11] and expanded on in [Mie16], which implies convergence of EDP-solutions and their free energies.

We say that a sequence $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n$ converges to some $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$ if for all $t \in [0, T]$ the probability measures P^n_t converge narrowly to P_t in $\mathcal{P}(\Gamma)$, and $\mathsf{J}^{n,\pm}_t(\mathrm{d}\nu, \mathrm{d}x)\,\mathrm{d}t$ converge vaguely to $\mathsf{J}^\pm_t(\mathrm{d}\nu, \mathrm{d}x)\,\mathrm{d}t$ in $\mathcal{M}^+_{\mathrm{loc}}([0, T]\times\Gamma\times\mathcal{T})$. Again postponing technicalities, see Theorem 19.2, we have the following lower semicontinuity and compactness result:

Theorem 15.7 (cf. Theorem 19.2)**.** *The sequence of free energies $\mathcal{F}_n$ Γ-converges to $\mathcal{F}_\infty$.*

Moreover, the sequence of Fisher-information functionals and dissipation potentials are all sequentially lower semicontinuous for sequences of curves $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n$ *converging to* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$ *such that*

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^{2+\varepsilon} \mathrm{d}\mathsf{P}^n_t < \infty, \tag{15.18}$$

for some fixed $\varepsilon > 0$. In particular, if $\mathcal{F}_n(\mathsf{P}^n_0) \to \mathcal{F}_\infty(\mathsf{P}_0)$ as well, we have

$$\liminf_{n\to\infty} \mathcal{J}_n(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \geq \mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).$$

In addition, for any sequence $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n$ *such that* (15.18) *is satisfied and*

$$\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_0) < \infty,$$

$$\limsup_{n\to\infty} \mathcal{J}_n(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) < \infty,$$

there exists a subsequence converging to some $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$.

Here the $(2+\varepsilon)$-moment conditions (15.18) are necessary to control the Fisher information functionals, which consist of lower semi-continuous functionals and correction terms that are on the order of quadratic moments themselves.

Moreover, the notion of EDP-convergence or evolutionary Γ-convergence (where the Γ is not to be confused with our space of positive measures Γ) relates to the Γ-convergence of the free energies $\mathcal{F}_n$ and suitable liminf-estimates for the dissipation potentials and Fisher-information functionals (or local slopes in a metric setting).

These liminf-estimates are sufficient to obtain convergence of the solutions, an approach also taken in [EFLS16, Erb16, MM20]. Namely, by a lower semicontinuity and compactness argument, Theorem 15.7 implies the convergence of both the solutions and the free energies $\mathcal{F}_n$, if the initial data are well prepared.

Theorem 15.8 (cf. Theorem 19.6)**.** *Suppose that* $\mathsf{P}_0^n \to \mathsf{P}_0$ *with* $\mathcal{F}_n(\mathsf{P}_0^n) \to \mathcal{F}_\infty(\mathsf{P}_0)$ *as well, and*

$$\int_\Gamma \nu(\mathcal{T})^{2+\varepsilon} \mathrm{d}\mathsf{P}_0 < \infty, \tag{15.19}$$

for a fixed $\varepsilon > 0$. *Then for the sequence* $\hat{\mathsf{P}}^n$ *of solutions to* $(\mathsf{FKE_n})$, *and* $\hat{\mathsf{P}}$ *the solution to* (Li), *we have that for all* $t \in [0, T]$

$$\hat{\mathsf{P}}_t^n \to \hat{\mathsf{P}}_t \ \textit{narrowly}, \quad \textit{and} \quad \lim_{n\to\infty} \mathcal{F}_n(\hat{\mathsf{P}}_t^n) = \mathcal{F}_\infty(\hat{\mathsf{P}}_t).$$

In particular, if $\mathsf{P}_0 = \delta_{\hat{\nu}_0}$ *and* $\hat{\nu}_t$ *is the solution to the mean-field problem* (MF), *then for all* $t \in [0, T]$

$$\hat{\mathsf{P}}_t^n \to \delta_{\hat{\nu}_t} \ \textit{narrowly}, \quad \textit{and} \quad \lim_{n\to\infty} \frac{1}{n} \mathcal{E}\mathrm{nt}(\hat{\mathsf{P}}_t^n | \Pi_n) = \mathcal{E}\mathrm{nt}(\hat{\nu}_t | \gamma).$$

The second half of Theorem 15.8, on the concentration around mean-field solutions and convergence of entropies, follows directly from the definition of $\mathcal{F}_\infty$ and uniqueness.

For interacting particle systems where the number of particles is fixed at $n \in \mathbb{N}$ the narrow convergence $\hat{\mathsf{P}}_t^n \to \delta_{\hat{\nu}_t}$ is equivalent to propagation of chaos in the sense of Sznitman [Szn91], and would imply narrow convergence of the k-particle marginals at time t to $\nu_t^{\otimes k}$. However, in our setting, this implies convergence of the k-correlation functions, see [BGSRS20].

Moreover, the convergence of the free energies $\mathcal{F}_n$ implies the stronger notion of entropic propagation of chaos if the initial data is sufficiently regular.

Theorem 15.9 (cf. Theorem 19.7). *Suppose that* $\mathsf{P}_0^n \to \delta_{\hat{\nu}_0}$ *with* $C^{-1} \leq \mathrm{d}\hat{\nu}_0/\mathrm{d}\gamma \leq C$ *for some* $C > 0$. *If the initial sequence* P_0^n *satisfies the moment condition* (15.19) *and is entropically chaotic in the sense that*

$$\lim_{n\to\infty} \frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}_0^n|\Pi_{n,\hat{\nu}_0}) = 0,$$

then this is propagated along the solution, i.e.

$$\lim_{n\to\infty} \frac{1}{n}\mathcal{E}\mathrm{nt}(\hat{\mathsf{P}}_t^n|\Pi_{n,\hat{\nu}_t}) = 0, \qquad \text{for all } t \geq 0,$$

where $\Pi_{n,\nu} \in \mathcal{P}(\Gamma)$ *stems from the Poisson measure* $\pi_{n,\nu}$ *with intensity measure* ν,

$$\pi_{n,\nu} := \frac{1}{e^{n\nu(\mathcal{T})}} \sum_{N=1}^{\infty} \frac{n^N}{N!}\nu^{\otimes N}.$$

15.4 Detailed balance

In the author's work [HT22] similar results as above were presented in the reversible setting, i.e. if the detailed balance condition with respect to the invariant measure Π_n holds for the Forward Kolmogorov equation with respect , although the equivalence to the rate function was left for future research. In contrast, above we discussed in detail the irreversible setting, and where Π_n played the role of a reference measure. However, for completeness, and to use the results of Chapter 14, let us assume in addition:

Assumption 15.10.

$c(x,x) = 0$	*for all* $x \in \mathcal{T}$	*(no natural death)*
$m(y,x) = c(x,y)$	*for all* $x, y \in \mathcal{T}$	*(mean-field detailed balance)*

Here the condition on no natural death states that particles cannot delete themselves, and the symmetry condition $m(x, y) = c(x, y)$ implies a gradient-flow structure for the mean-field equation. As we will see in Chapter 17, the two conditions together imply that the jump kernel for the Forward Kolmogorov equation satisfies the detailed balance condition with respect to the measure Π_n.

Remark 15.11. Note that in the case of mean-field detailed balance, i.e. $m(x, y) = c(x, y)$ for all $x, y \in \mathcal{T}$, we have $b^+[\nu] = b^-[\nu]$ and hence $\bar{\chi}_\nu^\pm = \chi_\nu^\mp$. Similarly, if in addition the no natural death condition $c(x, x) = 0$ for all $x \in \mathcal{T}$ holds and thus Assumption 15.10 is satisfied, we have $\bar{\chi}_\nu^{n,\pm} = \chi_\nu^{n,\mp}$. □

While the detailed balance condition severely restricts the birth/death kernels available, it does alleviate the problem of moment conditions on the mass, as stated below.

Theorem 15.12. *Suppose that Assumption 15.10 holds as well. Then all listed theorems above, except for the full equivalence for the Liouville equation (see Theorem 18.10) hold even if the mass-moment conditions are not satisfied.*

15.5 Discussion and bibliographical notes

The Bolker–Pacala–Dieckmann–Law model is a staple of population dynamics and theoretical biology, with the goal of deriving macroscopic equations from microscopic models [CFM06, FKK09]. Depending on the setting, the mathematical description of the model can be made rigorous in various ways: for example via an analytical approach on configuration spaces as done in [FKK09], which in fact models infinite configurations of particles over $\mathbb{R}^d$, or via martingale techniques with $\mathcal{T}$ a closed subset of $\mathbb{R}^d$ and $\gamma = \mathscr{L}^d|_{\mathcal{T}}$ (see [FM04]). Moreover, in the latter, under the assumption of continuous, bounded, and integrable mutation/competition rates, it is also shown that under the mean-field scaling and the large-population limit $n \to \infty$ the process converges to the mean-field equation.

While different choices of scalings are possible, the mean-field equation describes the macroscopic properties of the measure-valued process when the population is large. An alternative way is to study the evolution of the moments, which form a hierarchy similar to the BBGKY-hierarchy of correlation functions, and under the so-called Vlasov scaling the first moment or correlation function converges to (MF). For the case of infinite configurations over $\mathbb{R}^d$ this has been established, see [FKK09, FKK10], and both propagation of chaos in the Vlasov limit and the sub-Poissonian property have been established as well [FKKK15].

Moreover, in recent years, there has been activity in studying the mean-field equation and the BPDL model in more general spaces, allowing for dynamics involving multiple species and combinations of discrete and continuous traits. See for example [FKK21] for an overview of existing models, where instead of $\mathbb{R}^d$ the underlying space is an arbitrary locally compact Polish space. However, convergence in the large-population limit is not considered.

In contrast, in this thesis, we lay out a procedure to establish convergence of the sequence of forward Kolmogorov equations for Polish trait spaces, under the assumption that the trait space is in addition compact. Moreover, the mutation and competition kernels we consider are merely assumed to be bounded, and are part of a larger class of measure-dependent birth and death rates, as will be seen in Assumption 19.1. Under these constraints and the regularity of the mean-field solution, we show in addition entropic propagation of chaos, which to the author's knowledge is a first for the BPDL model.

This begs the question if one can establish *quantitative* estimates for propagation of chaos for such general systems involving birth and death, which we leave to future work.

The techniques we used to establish convergence are rooted in the powerful variational tools that have been developed in the last decade for studying mean-field interacting jump processes and their limits under the assumption of detailed balance. To highlight only a few: [EFLS16] studied mean-field limits for measure-dependent jump processes; [Erb16] proved the convergence of the spatially homogeneous Kac-process to the Boltzmann equation; [Sch19] investigated the macroscopic limit of Becker-Döring models; [KJZ19] showed hydrodynamic limits for zero-range and exclusion processes; [MM20] discussed convergence and higher-order approximations for chemical reaction networks, an approach that was subsequently used in the setting of discretized reaction-diffusion equations in [MSW22].

In the author's work [HT22] these techniques were extended and applied to the BPDL model using generalized gradient flows and one-way fluxes instead of net fluxes, which was only possible under the strict assumption of detailed balance and for entropies with respect to the invariant measure, conditions that in this thesis have now been lifted.

Our formulation tracks the effective *mass fluxes* for both creation (arising from mutation) and annihilation (arising from competition) separately. The use of mass fluxes instead of usual particle fluxes ensures that in our convergence results as $n \to \infty$ we have both convergences of laws and fluxes.

Separating the effects of birth and death instead of their combined contribution allows us to incorporate more information in our variational formulation. The downside is that we are forced to work with *positive fluxes*, while the framework in the aforementioned examples involves either quadratic or generalized structures for signed net fluxes. On the other hand, this brings us closer to the variational representations stemming from large deviations, involving so-called one-way or unidirectional fluxes, see for example [MPR14, PR19, BBBO21, PS22]. Indeed, due to the equivalence of the EDP-functionals and the rate functionals, our proposed structure is not only motivated by large deviation theory but might be used as a tool to prove them, as we will discuss briefly in Appendix G.

Chapter 16

Mean-field system

In this chapter, we will discuss the variational formulation of the mean-field equation. Let us first make precise the context of Theorem 15.4, and embed it within the more general statement of Theorem 16.8 below.

Recall that the trait space $\mathcal{T}$ is a compact Polish space, and $\Gamma := \mathcal{M}^+(\mathcal{T})$ is the space of finite nonnegative measures over $\mathcal{T}$ equipped with the narrow topology. Fix a reference measure $\gamma \in \Gamma$ with $\gamma(\mathcal{T}) > 0$. The mean-field equation then reads

$$\partial_t \nu_t = \chi^+[\nu_t] - \chi^-[\nu_t]. \tag{MF}$$

with measure-dependent $\chi^\pm : \Gamma \to \Gamma$ given by

$$\chi^+[\nu](\mathrm{d}x) := b^+[\nu](x)\gamma(\mathrm{d}x), \qquad \chi^-[\nu](\mathrm{d}x) := b^-[\nu](x)\nu(\mathrm{d}x).$$

Recall that the total variation norm $\|\cdot\|_{TV}$ on $\mathcal{M}(\mathcal{T})$ is defined as

$$\|\mu\|_{TV} := \sup\left\{ \int_{\mathcal{T}} f\,\mathrm{d}\mu \,:\, f \in \mathcal{B}_b(\mathcal{T}),\ \|f\|_\infty \le 1 \right\}, \qquad \mu \in \mathcal{M}(\mathcal{T}).$$

Throughout we will use a slightly more broader set of conditions than the definition of $b^\pm$ in terms of the mutation and competition kernel m, c as stated in (15.2). Instead, throughout Part II.B, we will assume the following, with an upcoming additional restriction in Chapter 19 to establish EDP-convergence.

Assumption 16.1. *The birth/death rates $b^\pm \ge 0$ are uniformly TV-Lipschitz in ν on sets of bounded mass, i.e. for every $C > 0$ there exists a constant L_C such that*

$$|b^\pm[\nu](x) - b^\pm[\mu](x)| \le L_C \|\mu - \nu\|_{TV},$$

for all $x \in \mathcal{T}$ and $\mu, \nu \in \Gamma$ with $\max(\mu(\mathcal{T}), \nu(\mathcal{T})) \le C$.

Moreover, they have at most linear growth in terms of the mass $\nu(\mathcal{T})$, i.e. there exists a constant $M' > 0$ such that

$$b^{\pm}[\nu](x) \leq M'\nu(\mathcal{T}), \qquad \textit{for all } x \in \mathcal{T} \textit{ and } \nu \in \Gamma\,.$$

We will show in Section 16.2 that strong solutions to (MF) in either total variation or appropriate L^1 spaces exist and are unique.

Moreover, note that if the rates are given by

$$b^{+}[\nu](x) := \int_{\mathcal{T}} m(y,x)\nu(\mathrm{d}y), \qquad b^{-}[\nu](x) := \int_{\mathcal{T}} c(x,y)\nu(\mathrm{d}y)$$

with mutation/competition rates m, c satisfying Assumption 15.1, i.e. $m, c \in \mathcal{B}_b(\mathcal{T}\times\mathcal{T})$, then Assumption 16.1 applies.

In addition, we consider

$$\bar{\chi}^{+}[\nu](\mathrm{d}x) = b^{+}[\nu](x)\nu(\mathrm{d}x) \qquad \bar{\chi}^{-}[\nu](\mathrm{d}x) = b^{-}[\nu](x)\gamma(\mathrm{d}x), \tag{16.1}$$

and routinely, we will also adopt the shorthand notation $b^{\pm}_{\nu}(x)$, $\chi^{\pm}_{\nu}(\mathrm{d}x)$, $\bar{\chi}^{\pm}$, etc.

Remark 16.2. Note that under the mean-field detailed balance condition $m(x,y) = c(y,x)$ we have $\bar{\chi}^{\pm} = \chi^{\mp}$. Moreover, in that case the dynamics simplify to

$$\partial_t \nu_t(\mathrm{d}x) = b^{+}[\nu_t](x)(\gamma(\mathrm{d}x) - \nu_t(\mathrm{d}x)),$$

and it is clear that γ is a stationary measure under this evolution. □

Finally, note that under Assumption 16.1

$$\chi^{+}_{\nu}(\mathcal{T}) \leq M'\nu(\mathcal{T})\gamma(\mathcal{T}), \qquad \chi^{\pm}_{\nu}(\mathcal{T}), \bar{\chi}^{\pm}_{\nu}(\mathcal{T}) \leq M(1+\nu(\mathcal{T})^2), \tag{16.2}$$

for the constant M given by

$$M = M'(1+\gamma(\mathcal{T})) \tag{16.3}$$

which we will use repeatedly for our estimates in Section 16.1.

We will now consider curves satisfying the continuity equation

$$\partial_t \nu_t = \lambda^{+}_t - \lambda^{-}_t, \tag{$\mathscr{C}\mathscr{E}$}$$

in an appropriately weak sense.

Definition 16.3 (Mean-field continuity equation)**.** A triple $(\nu, \lambda^{+}, \lambda^{-})$ satisfies the mean-field continuity equation $\mathscr{C}\mathscr{E}$ if

1. the curve $[0,T] \ni t \mapsto \nu_t \in \Gamma$ is absolutely continuous with respect to $\|\cdot\|_{TV}$,
2. the Borel families $(\lambda_t^\pm)_{t\in[0,T]} \subset \Gamma$ satisfy $\int_0^T \|\lambda_t^\pm\|_{TV}\,\mathrm{d}t < \infty$,
3. for every $s,t \in [0,T]$ and all $f \in B_b(\mathcal{T})$

$$\int_{\mathcal{T}} f\,\mathrm{d}\nu_t - \int_{\mathcal{T}} f\,\mathrm{d}\nu_s = \int_s^t \left(\int_{\mathcal{T}} f\,\mathrm{d}\lambda_r^+ - \int_{\mathcal{T}} f\,\mathrm{d}\lambda_r^- \right) \mathrm{d}r,$$

□

We will refer to $\lambda^{\mathrm{net}} = \lambda^+ - \lambda^-$ as the *net flux*.

Remark 16.4. When seen as approximations of particle systems the birth/death fluxes $\lambda_t^\pm$ represent the *observed* amount of mass being created/annihilated around a certain point, and ν_t represents the density of the particles, while $\chi_\nu^\pm$ correspond to the *expected* birth and death fluxes of the BPDL model.

Moreover, the $\bar{\chi}^\pm$ stem from the backward kernels of the forward Kolmogorov equation, and we will refer them as *backward* expected fluxes. □

Recall that the squared Hellinger distance H^2 is given by (11.9), i.e.

$$H^2(\nu,\eta) := \frac{1}{2}\int_{\mathcal{T}} \left(\sqrt{\frac{\mathrm{d}\nu}{\mathrm{d}\sigma}} - \sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\sigma}} \right)^2 \mathrm{d}\sigma, \tag{16.4}$$

with σ a measure dominating both μ and ν. Moreover, recall the entropy function $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and its Legendre dual $\phi^* : \mathbb{R} \to \mathbb{R}$ given by

$$\phi(s) := s\log s - s + 1, \qquad \phi^*(z) := e^z - 1,$$

and the relative entropy of ν with respect to μ as

$$\mathcal{E}\mathrm{nt}(\nu|\mu) := \begin{cases} \displaystyle\int_{\mathcal{T}} \phi\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)\mathrm{d}\mu, & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \tag{16.5}$$

Definition 16.5. Let $\theta_\nu^\pm$ be the geometric average of $\chi_\nu^\pm$ and $\bar{\chi}_\nu^\pm$, i.e.

$$\theta_\nu^\pm := \sqrt{\chi_\nu^\pm \bar{\chi}_\nu^\pm}.$$

for any dominating measure σ. We define the following objects:

- The dissipation potential $\mathcal{R}_{MF} : \Gamma^3 \to [0, +\infty]$,

$$\mathcal{R}_{MF}(\nu, \lambda^+, \lambda^-) := \mathcal{E}\mathrm{nt}(\lambda^+|\theta^+_\nu) + \mathcal{E}\mathrm{nt}(\lambda^-|\theta^-_\nu),$$

  and the dual dissipation potential $\mathcal{R}^*_{MF} : \Gamma \times \mathcal{B}_b(\mathcal{T})^2 \to \mathbb{R}$,

$$\mathcal{R}^*_{MF}(\nu, w^+, w^-) := \int_{\mathcal{T}} (e^{w^+} - 1)\,\mathrm{d}\theta^+_\nu + \int_{\mathcal{T}} (e^{w^-} - 1)\,\mathrm{d}\theta^-_\nu.$$

- The free energy $\mathcal{F}_{MF} : \Gamma \to [0, +\infty]$,

$$\mathcal{F}_{MF}(\nu) := \tfrac{1}{2}\mathcal{E}\mathrm{nt}(\nu|\gamma),$$

- The Fisher information $\mathcal{D}_{MF} : \Gamma \to \mathbb{R}$,

$$\mathcal{D}_{MF}(\nu) := H^2(\chi^+_\nu, \bar{\chi}^+_\nu) + H^2(\chi^-_\nu, \bar{\chi}^-_\nu) + \frac{1}{2}\int_{\mathcal{T}} \mathrm{d}\left(\chi^+_\nu + \chi^-_\nu - \bar{\chi}^+_\nu - \bar{\chi}^-_\nu\right)$$

- The EDP-functional $\mathcal{J}_{MF} : \mathscr{C}\mathscr{E} \to [0, +\infty]$ for all curves with $\mathcal{F}_{MF}(\nu_0) < \infty$

$$\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) := \int_0^T \mathcal{R}_{MF}(\nu_t, \lambda^+_t, \lambda^-_t)\,\mathrm{d}t + \mathcal{F}(\nu_T) - \mathcal{F}(\nu_0) + \int_0^T \mathcal{D}_{MF}(\nu_t)\,\mathrm{d}t. \tag{16.6}$$

- The Lagrangian $\mathcal{L}_{MF} : \Gamma^3 \to [0, +\infty]$,

$$\mathcal{L}_{MF}(\nu, \lambda^+, \lambda^-) := \mathcal{E}\mathrm{nt}(\lambda^+|\chi^+_\nu) + \mathcal{E}\mathrm{nt}(\lambda^-|\chi^-_\nu),$$

  and its dual, the Hamiltonian $\mathcal{H}_{MF} : \Gamma \times \mathcal{B}_b(\mathcal{T})^2 \to \mathbb{R}$,

$$\mathcal{H}_{MF}(\nu, w^+, w^-) := \int_{\mathcal{T}} (e^{w^+} - 1)\,\mathrm{d}\chi^+_\nu + \int_{\mathcal{T}} (e^{w^-} - 1)\,\mathrm{d}\chi^-_\nu.$$

- The rate functional $\mathcal{I}_{MF} : \mathscr{C}\mathscr{E} \to [0, +\infty]$

$$\mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-) = \int_0^T \mathcal{L}_{MF}(\nu_t, \lambda^+_t, \lambda^-_t)\,\mathrm{d}t.$$

□

Remark 16.6. For the last term of the Fisher information we have

$$\chi_\nu^+ + \chi_\nu^- - \bar{\chi}_\nu^+ - \bar{\chi}_\nu^- = (b_\nu^+ - b_\nu^-)(\gamma - \nu),$$

and is a correction term describing the lack of mean-field detailed balance condition, since in that case γ is the stationary measure and $b_\nu^+ = b_\nu^-$.

Moreover, in the case that the mean-field detailed balance holds we have $\theta_\nu^+ = \theta_\nu^- =: \theta_\nu$, and $\mathcal{R}_{MF}^*$ and $\mathcal{D}_{MF}$ simplify to

$$\mathcal{R}_{MF}^*(\nu, w^+, w^-) := \int_{\mathcal{T}} (e^{w^+} + e^{w-} - 2)\,\mathrm{d}\theta_\nu, \; \mathcal{D}_{MF}(\nu) = 2H^2(\chi_\nu^+, \chi_\nu^-).$$

□

Remark 16.7. If $\nu \ll \gamma$ with $\mathrm{d}\nu = u\mathrm{d}\gamma$, note that $\mathrm{d}\theta_\nu^\pm = b_\nu^\pm \sqrt{u}\,\mathrm{d}\gamma$, and that the Fisher information simplifies to

$$\mathcal{D}_{MF}(\nu) = \int_{\mathcal{T}} b_\nu^+(1 - \sqrt{u})\,\mathrm{d}\gamma + \int_{\mathcal{T}} b_\nu^-(u - \sqrt{u})\,\mathrm{d}\gamma.$$

□

We are now able to fully state the equivalence between the EDP-functional and the rate functional, and the characterization of strong solutions to the mean-field equation (MF).

Theorem 16.8. *For any* $(\nu, \lambda^+, \lambda^-) \in \mathscr{CE}$ *with* $\mathcal{F}_{MF}(\nu_0) < \infty$

$$\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) = \mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-). \tag{16.7}$$

In particular, $\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) \geq 0$. *Moreover,*

$$\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) = 0 \iff \begin{cases} \nu_t \text{ is the unique strong solution to (MF)}, \\ \lambda_t^\pm = (\chi_{\nu_t}^\pm) \quad \text{for a.e. } t \in [0, T]. \end{cases}$$

Finally, if $\mathcal{F}_{MF}(\nu_0) < \infty$ *and* $\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) < \infty$ *the chain rule for* $\mathcal{F}_{MF}$ *holds: the map* $t \mapsto \mathcal{F}_{MF}(\nu_t)$ *is absolutely continuous and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_{MF}(\nu_t) = \frac{1}{2}\int_{\mathcal{T}} \log\frac{\mathrm{d}\nu_t}{\mathrm{d}\gamma}\,\mathrm{d}(\lambda_t^+ - \lambda_t^-), \qquad \text{for a.e. } t \in [0, T].$$

The proof of Theorem 16.8 is postponed to Section 16.3, where we establish the main technical ingredient, namely the chain rule for the entropy functional. Moreover, in essence the techniques used throughout this chapter are very similar to Chapter 12, but instead of jump processes with a bounded kernel κ the arguments are applied to evolutions with measure-dependent birth/death fluxes $\chi_\nu^\pm$. Because of this, we will for certain arguments only briefly sketch the proof.

16.1 A priori estimates

In this section, we will collect some elementary estimates and results that are either necessary for the well-posedness of the mean-field equation and the corresponding variational structure, or necessary to do the same for the Liouville equation in Chapter 18.

Recall from (11.5) that $\tilde{\phi}(s) := \phi(s \vee 1)$ is the monotone relaxation of ϕ.

Lemma 16.9. *Let M be the constant of* (16.2). *Then the following estimates hold:*

(i) The measures $\theta_\nu^\pm$ are finite with

$$\theta_\nu^\pm(\mathcal{T}) \le M(1 + \nu(\mathcal{T})^2) \tag{16.8}$$

(ii) For any birth/death fluxes $\lambda^\pm \in \Gamma$,

$$\tilde{\phi}\left(\frac{\lambda^\pm(\mathcal{T})}{M(1+\nu(\mathcal{T})^2)}\right) M \le \min\left(\mathcal{R}_{MF}(\nu, \lambda^+, \lambda^-),\, \mathcal{L}_{MF}(\nu, \lambda^+, \lambda^-)\right) \tag{16.9}$$

(iii) If the mean-field detailed balance condition holds, i.e. $b^+ = b^-$, then for any net flux $\lambda^{\mathrm{net}} \in \mathcal{M}(\mathcal{T})$,

$$\Psi\left(\frac{\|\lambda^{\mathrm{net}}\|_{TV}}{M(1+\nu(\mathcal{T}))}\right) M \le \mathcal{R}_{MF}(\nu, \lambda^+, \lambda^-). \tag{16.10}$$

Proof. The first statement directly follows from the fact that $\theta_\nu^\pm(\mathcal{T}) \le \sqrt{\chi_\nu^\pm(\mathcal{T})\bar{\chi}_\nu^\pm(\mathcal{T})}$. Next, without loss of generality, suppose that $\mathcal{R}_{MF}$ is finite. Set $a(\nu) := (1 + \nu(\mathcal{T})^2)^{-1}$, and note that $0 \le a(\nu) \le 1$. We then have the following chain of inequalities,

$$\begin{aligned}\int_{\mathcal{T}} \phi\left(\frac{\mathrm{d}\lambda^\pm}{\mathrm{d}\theta_\nu^\pm}\right) \mathrm{d}\theta_\nu^\pm &\ge \int_{\mathcal{T}} \tilde{\phi}\left(\frac{\mathrm{d}\lambda^\pm}{\mathrm{d}\theta_\nu^\pm}\right) \mathrm{d}\theta_\nu^\pm \\ &\ge \int_{\mathcal{T}} \tilde{\phi}\left(\frac{\mathrm{d}(a(\nu)\lambda^\pm)}{\mathrm{d}(a(\nu)\theta_\nu^\pm)}\right) \mathrm{d}(a(\nu)\theta_\nu^\pm) \\ &\ge \tilde{\phi}\left(\frac{a(\nu)\lambda^\pm(\mathcal{T})}{a(\nu)\theta_\nu^\pm(\mathcal{T})}\right) a(\nu)\theta_\nu^\pm(\mathcal{T}),\end{aligned}$$

where the last inequality follows from Jensen's inequality. By convexity of $\tilde{\phi}$ and $\tilde{\phi}(0) = 0$ the latter expression is monotone in $\theta_\nu(\mathcal{T})$, and hence by (16.8) we find

$$\tilde{\phi}\left(\frac{\lambda^\pm(\mathcal{T})}{M(1+\nu(\mathcal{T})^2)}\right) M \le \mathcal{R}_{MF}(\nu, \lambda^+, \lambda^-).$$

A similar argument follows under the assumption of finite Lagrangian $\mathcal{L}_{MF}$.

Finally, for the net flux in the case of detailed balance, it is convenient to go through the dual representation. Namely, note that

$$\mathcal{R}^*_{MF}(\nu, w, -w) = \int_{\mathcal{T}} (e^{w} - 1)\,\mathrm{d}\theta_\nu + \int_{\mathcal{T}} (e^{-w} - 1)\,\mathrm{d}\theta_\nu = \int_{\mathcal{T}} \Psi^*(w)\,\mathrm{d}\theta_\nu.$$

for any $w \in \mathcal{B}_b(\mathcal{T})$. Therefore, by duality,

$$\mathcal{R}_{MF}(\nu, \lambda^+, \lambda^-) \geq a(\nu) \int_{\mathcal{T}} w(x)\,\mathrm{d}\lambda^{\mathrm{net}} - \int_{\mathcal{T}} \Psi^*\big(a(\nu)w(x)\big)\,\mathrm{d}\theta_\nu.$$

However, by (11.8) it follows that

$$\int_{\mathcal{T}} \Psi^*\big(a(\nu)w(x)\big)\,\mathrm{d}\theta_\nu \leq \int_{\mathcal{T}} \Psi^*\big(w(x)\big)a(\nu)^2\,\mathrm{d}\theta_\nu \leq M\Psi^*(\|w\|_\infty).$$

Taking the supremum over all $w \in \mathcal{B}_b(\mathcal{T})$ find (16.10). □

Remark 16.10. Although the estimate for θ_ν can be made more precise, namely

$$\theta^\pm_\nu(\mathcal{T}) \leq M'\gamma(\mathcal{T})^{1/2}\nu(\mathcal{T})^{3/2},$$

we will not require it for our results. □

The following regularity result is a direct analogue of Corollary 12.9 adapted to the continuity equation $\partial_t \nu_t = \lambda^+_t - \lambda^-_t$ and the fact that $\chi^+_\nu, \theta^\pm_\nu \ll \gamma$ and $\chi^-_\nu, \theta^\pm_\nu \ll \nu$, and we therefore omit the proof.

Lemma 16.11. *Suppose that* $(\nu, \lambda^+, \lambda^-) \in \mathscr{CE}$ *and either* $\mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-) < \infty$ *or*

$$\int_0^T \mathcal{R}_{MF}(\nu_t, \lambda^+_t, \lambda^-_t)\,\mathrm{d}t < \infty.$$

Then $t \mapsto \nu_t$ *is a.e. differentiable in* $(\mathcal{M}(\mathcal{T}), \|\cdot\|_{TV})$ *with derivative* $\partial_t \nu_t$ *in* $\mathcal{M}(\mathcal{T})$ *and*

$$\partial_t \nu_t = \lambda^+_t - \lambda^-_t, \qquad \text{for a.e. } t \in [0, T].$$

Finally, if $\nu_0 \ll \gamma$ *then* $\nu_t \ll \gamma$ *for all* $t \in [0, T]$*, and* $\partial_t \nu_t \ll \gamma$ *for a.e.* $t \in [0, T]$.

Next, we will consider a result that is necessary for the superposition principle of Chapter 18.

Lemma 16.12. *Let* $\{f_i\}_{i\in\mathbb{N}} \subset C_b(\mathcal{T})$ *be a countable and dense set of bounded continuous functions. Suppose* $(\nu, \lambda^+, \lambda^-)$ *is such that*

(i) the curve $[0, T] \ni t \mapsto \nu_t \in \Gamma$ *is narrowly continuous*

(ii) $(\lambda_t^{\pm})_{t\in[0,T]} \subset \Gamma$ *is a Borel family with either:*

$$\int_0^T \|\lambda_t^{\pm}\|_{TV}\,\mathrm{d}t < \infty,\ \mathcal{I}_{MF}(\nu,\lambda^+,\lambda^-) < \infty,\ \textit{or}\ \int_0^T \mathcal{R}_{MF}(\nu_t,\lambda_t^+,\lambda_t^-)\,\mathrm{d}t < \infty$$

(iii) For all $i \in \mathbb{N}$ *and for all* s,t *with* $0 \le s,t \le T$

$$\int_{\mathcal{T}} f_i\,\mathrm{d}\nu_t - \int_{\mathcal{T}} f_i\,\mathrm{d}\nu_s = \int_s^t \left(\int_{\mathcal{T}} f_i\,\mathrm{d}\lambda_r^+ - \int_{\mathcal{T}} f_i\,\mathrm{d}\lambda_r^-\right)\mathrm{d}r,$$

Then $(\nu,\lambda^+,\lambda^-) \in \mathscr{CE}$, *i.e. the triple satisfies the mean-field continuity equation.*

Proof. Since ν_t is narrowly continuous its mass is uniformly bounded in time, hence let $C := \sup_{t\in[0,T]} \nu_t(\mathcal{T})$. By (16.9) and monotonicity of $\tilde{\phi}$ we have for a.e. $t \in [0,T]$,

$$\tilde{\phi}\left(\frac{\lambda_t^{\pm}(\mathcal{T})}{M(1+C^2)}\right)M \le \min\left(\mathcal{R}_{MF}(\nu_t,\lambda_t^+,\lambda_t^-), \mathcal{L}_{MF}(\nu_t,\lambda_t^+,\lambda_t^-)\right),$$

and therefore by convexity of $\tilde{\phi}$

$$\int_0^T \lambda_t^{\pm}(\mathcal{T})\,\mathrm{d}t < \infty.$$

Since the measures $\lambda_t^{\pm}(\mathrm{d}x)\,\mathrm{d}t \in \mathcal{M}^+([0,T]\times\Gamma)$ are finite, by density of f_i in $C_b(\mathcal{T})$ it is clear that for all $f \in C_b(\mathcal{T})$, and for all s,t with $0 \le s,t \le T$

$$\int_{\mathcal{T}} f\,\mathrm{d}\nu_t - \int_{\mathcal{T}} f\,\mathrm{d}\nu_s = \int_s^t \left(\int_{\mathcal{T}} f\,\mathrm{d}\lambda_r^+ - \int_{\mathcal{T}} f\,\mathrm{d}\lambda_r^-\right)\mathrm{d}r.$$

By a monotone class argument this can be extended to all $f \in \mathcal{B}_b(\mathcal{T})$ and we derive that ν_t is indeed TV-absolutely continuous and $(\nu,\lambda^+,\lambda^-) \in \mathscr{CE}$. □

16.2 Strong solutions

Strong solutions to (MF) exist and are unique, and we list the most important properties here.

Definition 16.13. A *strong* solution to (MF) is any TV-absolutely continuous and a.e. differentiable mapping $\nu : [0,T] \to (\Gamma, \|\cdot\|_{TV})$ with derivative $\partial_t \nu_t$ satisfying

$$\partial_t \nu_t(\mathrm{d}x) = \chi_{\nu_t}^+(\mathrm{d}x) - \chi_{\nu_t}^-(\mathrm{d}x) \tag{16.11}$$

□

Recall that $\chi_\nu^+(\mathrm{d}x) = b_\nu^+(x)\gamma(\mathrm{d}x)$ and $\chi_\nu^-(\mathrm{d}x) = b_\nu^-(x)\nu(\mathrm{d}x)$.

Remark 16.14. Note that if ν is a strong solution to (MF) we have automatically $(\nu, \chi_\nu^+, \chi_\nu^-) \in \mathscr{C}\mathscr{E}$, and the converse statement follows from Lemma 16.11. □

Proposition 16.15. *For any $\bar\nu \in \Gamma$ there exists a unique strong solution ν_t to* (MF) *such that $\nu_0 = \bar\nu$. Moreover, if $\bar\nu \ll \gamma$, then also $\nu_t \ll \gamma$ for all $t \in [0,T]$.*

The proof is an adaptation from [FM04, Proposition 7.2], which is stated for Lebesgue absolutely continuous measures over $\mathcal{T} = \mathbb{R}^d$. In short, the linear dependence of the birth flux on the mass of ν gives a bound on this mass uniform in time, in which case both $\chi_\nu^\pm$ are Lipschitz in ν on $(\Gamma, \|\cdot\|)$, and classical existence theory can be applied. It is similar to the proof of Proposition 12.12 and we will therefore omit some steps.

Proof. The proof is similar to that of Proposition 12.12, and we will omit certain steps. We set $\nu_t^0 := \bar\nu$ for all $t \in [0,T]$, and perform the implicit Picard iteration

$$\partial_t \nu_t^{k+1}(\mathrm{d}x) = b^+[\nu_t^k](x)\gamma(\mathrm{d}x) - b^-[\nu_t^k](x)\nu_t^{k+1}(\mathrm{d}x), \qquad \nu_0^{k+1} := \bar\nu,$$

i.e. $\nu^{k+1} = (\mathcal{G}\nu^k)$ with

$$(\mathcal{G}\nu)_t(\mathrm{d}x) := e^{-\int_0^t b^-[\nu_s](x)\,\mathrm{d}s}\left(\int_0^t b^+[\nu_s]\gamma(\mathrm{d}x)\rangle e^{\int_0^s b^-[\nu_r](x)\,\mathrm{d}r}\mathrm{d}s + \bar\nu(\mathrm{d}x)\right).$$

It is straightforward to check that for all $t \in [0,T]$

$$\sup_{k\geq 1} \nu_t^k(\mathcal{T}) \leq e^{M'\gamma(\mathcal{T})t}\bar\nu(\mathcal{T}) \leq e^{M'\gamma(\mathcal{T})T}\bar\nu(\mathcal{T}) =: C.$$

Moreover, recall that by Assumption 16.1 the rates $b^\pm[\nu](x)$ depend Lipschitz on ν on sets of bounded mass, and hence the proof continues as in the proof of Proposition 12.12.

□

Finally, for the use in entropic propagation chaos of Theorem 19.7, it is convenient to characterize the conditions for which u_t is uniformly bounded from above and below. The following statement follows directly from a Gronwall-type argument.

Lemma 16.16. *Suppose ν_0 is such that $C^{-1} \leq \mathrm{d}\nu_0/\mathrm{d}\gamma(x) < C$ for some constant $C > 0$ and all $x \in \mathcal{T}$. Then there exists a constant $C_T > 0$ such that for the corresponding solution*

$$C_T^{-1} \leq \frac{\mathrm{d}\nu_0}{\mathrm{d}\gamma}(x) < C_T, \qquad \textit{for all } x \in \mathcal{T}, \textit{ for all } t \in [0,T].$$

16.3 Variational characterization

We will now prove Theorem 16.8. As in Section 12.3, the statement will be split up into two propositions, one establishing the chain rule and equivalence $\mathcal{J}_{MF} = \mathcal{I}_{MF}$ under finite $\mathcal{J}_{MF}$, and the other showing that $\mathcal{J}_{MF} \le \mathcal{I}_{MF}$. Due to the similarities to Propositions 12.14 and 12.16, we will only give brief sketches for certain steps.

Proposition 16.17. *Let* $(\nu, \lambda^+, \lambda^-) \in \mathscr{CE}$ *such that* $\mathcal{F}_{MF}(\nu_0) < \infty$ *and*

$$\int_0^T \mathcal{R}_{MF}(\nu_t, \lambda_t^+, \lambda_t^-)\,\mathrm{d}t < \infty. \tag{16.12}$$

Then for a.e. t *we have* $\lambda_t^\pm(\{u_t(x)\}^c) = 0$*. Moreover, for the free energy the chain rule holds:* $t \mapsto \mathcal{F}_{MF}(\nu_t)$ *is absolutely continuous and a.e. differentiable with*

$$\partial_t \mathcal{F}_{MF}(\nu_t) = \frac{1}{2}\int_{\mathcal{T}} \phi'(u_t)\,\mathrm{d}(\lambda_t^+ - \lambda_t^-), \qquad \textit{for a.e. } t \in [0,T]. \tag{16.13}$$

Finally, for a.e. $t \in [0,T]$

$$\mathcal{R}_{MF}(\nu_t, \lambda_t^+, \lambda_t^-) + \mathcal{D}_{MF}(\nu_t) + \partial_t \mathcal{F}_{MF}(\nu_t) = \mathcal{L}_{MF}(\nu_t, \lambda_t^+, \lambda_t^-) \tag{16.14}$$

and, in particular, $\mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-) = \mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-)$*.*

Remark 16.18. Suppose that the chain rule for $\mathcal{F}_{MF}$ holds and $u_t(x) > 0$ for all $t \in [0,T], x \in \mathcal{T}$. The relation (16.14) can then be derived as follows. Setting $\mathrm{d}\lambda_t^\pm = g_t^\pm b_{\nu_t}^\pm \mathrm{d}\gamma$, we have

$$\begin{aligned}
\mathcal{E}\mathrm{nt}(\lambda_t^\pm|\theta^\pm) &= \int_{\mathcal{T}} \left(g_t^\pm \log \frac{g_t^\pm}{\sqrt{u_t}} - g_t^\pm + \sqrt{u_t} \right) b_{\nu_t}^\pm \,\mathrm{d}\gamma, \\
\mathcal{D}_{MF}(\nu) &= \int_{\mathcal{T}} \left(b_\nu^+(1 - \sqrt{u}) + b_\nu^-(u - \sqrt{u}) \right) \mathrm{d}\gamma, \\
\partial_t \mathcal{F}(\nu_t) &= \frac{1}{2}\int_{\mathcal{T}} (g_t^+ b_{\nu_t}^+ - g_t^- b_{\nu_t}^-) \log u \,\mathrm{d}\gamma,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{E}\mathrm{nt}(\lambda_t^+|\chi_{\nu_t}^+) &= \int_{\mathcal{T}} \left(g_t^+ \log g_t^+ - g_t^+ + 1 \right) b_{\nu_t}^+ \,\mathrm{d}\gamma, \\
\mathcal{E}\mathrm{nt}(\lambda_t^+|\chi_{\nu_t}^-) &= \int_{\mathcal{T}} \left(g_t^- \log \frac{g_t^-}{u_t} - g_t^- + u_t \right) b_{\nu_t}^- \,\mathrm{d}\gamma,
\end{aligned}$$

Putting all terms together, one obtains (16.14). □

Remark 16.19. Similar as for the symmetric/antisymmetric splitting of the rate functional for the forward Kolmogorov equation, as done in Section 10.5.1, the equivalence of Remark 16.18 can be seen via a splitting argument as well.

Namely, define the adjoint Lagrangian by switching $\chi^{\pm}$ by $\bar{\chi}^{\pm}$, i.e.

$$\mathcal{L}^{\dagger}_{MF}(\nu, \lambda^{+}, \lambda^{-}) := \mathcal{E}\mathrm{nt}(\lambda_t^{+}|\bar{\chi}^{+}_{\nu_t}) + \mathcal{E}\mathrm{nt}(\lambda_t^{-}|\bar{\chi}^{-}_{\nu_t}). \tag{16.15}$$

Now, note that for any curve $(\nu, \lambda^{+}, \lambda^{-}) \in \mathscr{C}\mathscr{E}$ the 'reversed' curve $(\nu_{T-t}, \lambda^{-}_{T-t}, \lambda^{+}_{T-t})$ is still contained in $\mathscr{C}\mathscr{E}$, and

$$\mathcal{I}^{\dagger}_{MF}(\nu, \lambda^{+}, \lambda^{-}) := \int_0^T \mathcal{L}^{\dagger}_{MF}(\nu_{T-t}, \lambda^{-}_{T-t}, \lambda^{+}_{T-t})\,\mathrm{d}t = \int_0^T \mathcal{L}^{\dagger}_{MF}(\nu_t, \lambda^{-}_t, \lambda^{+}_t)\,\mathrm{d}t.$$

Then for suitable curves one can show the decomposition

$$\frac{1}{2}\left(\mathcal{I}_{MF} + \mathcal{I}^{\dagger}_{MF}\right) = \int_0^T \left(\mathcal{R}_{MF}(\nu_t, \lambda^{+}_t, \lambda^{-}_t) + H^2(\chi^{+}_{\nu_t}, \bar{\chi}^{+}_{\nu_t}) + H^2(\chi^{+}_{\nu_t}, \bar{\chi}^{+}_{\nu_t})\right)\mathrm{d}t,$$

$$\frac{1}{2}\left(\mathcal{I}_{MF} - \mathcal{I}^{\dagger}_{MF}\right) = \mathcal{F}_{MF}(\nu_T) - \mathcal{F}_{MF}(\nu_0) + \frac{1}{2}\int_0^T \int_{\mathcal{T}} \mathrm{d}\left(\chi^{+}_{\nu_t} + \chi^{-}_{\nu_t} - \bar{\chi}^{+}_{\nu_t} - \bar{\chi}^{-}_{\nu_t}\right)\mathrm{d}t.$$

Note that the second term for the antisymmetric splitting drops out under the mean-field detailed balance condition $b^{+} = b^{-}$, in which case we obtain a gradient-flow decomposition of the rate functional.

□

Proof of Proposition 16.17. From Lemma 16.11 we have that $\nu_t \ll \gamma$ for every $t \in [0, T]$, and therefore $\mathrm{d}\theta^{\pm}_t = b^{\pm}_{\nu_t}\sqrt{u_t}\mathrm{d}\gamma$ and $\theta^{\pm}_{\nu_t}(\{u_t(x)\}^c) = 0$. By assumption of (16.12) it follows that $\mathcal{R}(\nu_t, \lambda^{+}_t, \lambda^{-}_t) < \infty$ for a.e. $t \in [0, T]$, which implies $\lambda^{\pm}_t \le \theta^{\pm}_{\nu_t}$ and thus $\lambda^{\pm}_t(\{u_t(x)\}^c) = 0$ for such t. Now, recall the regularized entropy functions ϕ_n of (11.15), and let $\mathrm{d}\nu = u\mathrm{d}\gamma$. We consider the functionals

$$S_n(\nu) := \int_{\mathcal{T}} \phi_n(u)\,\mathrm{d}\gamma.$$

Recall that as $n \to \infty$ it holds that $S_n(\nu)$ converges to $S(\nu) := \mathcal{E}\mathrm{nt}(\nu|\gamma)$ for every $\nu \in \Gamma$ and $\phi'_n(z) \to \phi'(z)$ for every $z \ge 0$, and $|\phi'_n(z)| \le \phi'(z)$ for every $n > 0$ and $z \ge 0$. Moreover, from Lemmas 11.3 and 16.11 it is clear that for every $n > 0$ we have the chain rule: for every $(\nu, \lambda^{+}, \lambda^{-}) \in \mathscr{C}\mathscr{E}$

$$S_n(\nu_t) - S_n(s) = \int_s^t \int_{\mathcal{T}} \phi'(u_r)\,\mathrm{d}(\lambda^{+}_r - \lambda^{-}_r)\,\mathrm{d}t. \tag{16.16}$$

Following the proof of Proposition 12.14, it is clear that to obtain the chain rule for $\mathcal{F}_{MF}$ it is sufficient to have the bound

$$\mathcal{R}^{*}_{MF}(\nu, \tfrac{1}{2}|\phi'_n(u)|, \tfrac{1}{2}|\phi'_n(u)|) \le 4M(1 + \nu(\mathcal{T})^2)$$

for every $u \in L^1(\gamma)$. However, this follows directly since $e^{|z|} \leq e^z + e^{-z}$ and thus

$$\int_{\mathcal{T}} (e^{\frac{1}{2}|\phi_n'(u)|} - 1)\, \mathrm{d}\theta_u^+ = \int_{\mathcal{T}} (\sqrt{u} + \sqrt{u^{-1}}) b_\nu^\pm \sqrt{u}\, \mathrm{d}\gamma \leq \chi_\nu^+(\mathcal{T}) + \bar{\chi}_\nu^+(\mathcal{T}) \leq 2M(1+\nu(\mathcal{T})^2),$$

and a similar argument holds for θ_ν^-. Finally, with the chain rule in hand, one can now show the equivalence by repeating the calculation of Remark 16.18. □

Proposition 16.20. *For any* $(\nu, \lambda^+, \lambda^-) \in \mathscr{C}\mathscr{E}$ *such that* $\mathcal{F}_{MF}(\nu_0) < \infty$,

$$\mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-) \geq \mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-).$$

Remark 16.21. Similar as for Proposition 12.16, the statement above is proven via a dual approach, and the fact that if $u > 0$

$$\begin{aligned}
\mathcal{R}_{MF}^*(\nu, w^+, w^-) &= \mathcal{H}_{MF}(\nu, w^+ + D\mathcal{F}_{MF}, w^- - D\mathcal{F}_{MF}) - \mathcal{H}_{MF}(\nu, D\mathcal{F}_{MF}, -D\mathcal{F}_{MF}),\\
\mathcal{D}_{MF}(\nu) &= \mathcal{H}_{MF}(\nu, D\mathcal{F}_{MF}, -D\mathcal{F}_{MF})\\
\mathcal{H}_{MF}(\nu, w^+, w^-) &= \mathcal{R}_{MF}(\nu, w^+ - D\mathcal{F}_{MF}, w^- + D\mathcal{F}_{MF}) - \mathcal{H}_{MF}(\nu, -D\mathcal{F}_{MF}, D\mathcal{F}_{MF})\\
\mathcal{D}_{MF}(\rho) &= \mathcal{R}_{MF}^*(\nu, -D\mathcal{F}_{MF}, D\mathcal{F}_{MF}).
\end{aligned} \tag{16.17}$$

where $D\mathcal{F}_{MF}(\nu) := \frac{1}{2}\log u$ is the variational derivative of $\mathcal{F}_{MF}$ with respect to ν. □

Proof. Using a dual approach one can repeat the steps of Proposition 12.16. All that remains are the statements

$$\mathcal{H}_{MF}(\nu, w^+ + \tfrac{1}{2}\phi_n'(u), w^- - \tfrac{1}{2}\phi_n'(u)) \leq 4e^{\|w\|_\infty} M(1 + \nu(\mathcal{T})^2), \tag{16.18}$$

$$\limsup_{n\to\infty} \mathcal{H}_{MF}(\nu, w^+ + \tfrac{1}{2}\phi_n'(u), w^- - \tfrac{1}{2}\phi_n'(u)) \leq \mathcal{R}_{MF}^*(\nu, w^+, w^-) - \mathcal{D}_{MF}(\nu), \tag{16.19}$$

for any $\nu \in \Gamma$ and $w^\pm \in B_b(\mathcal{T})$. Now, note that we have the truncation inequality $e^{[a]_n} \leq e^a + 1$ for any $a \in \bar{\mathbb{R}}$, and therefore for any $u > 0$, $z^+, z^- \in \mathbb{R}$,

$$e^{z^+ + \frac{1}{2}\phi_n'(u)} - 1 \leq e^{z^+}\left(e^{\frac{1}{2}\phi'(u)} + 1\right) = e^{z^+}(\sqrt{u} + 1) \leq \frac{e^{z^+}}{2}(u+3).$$

Similarly,

$$\left(e^{z^- - \frac{1}{2}\phi_n'(u)} - 1\right) u \leq e^{z^-}\left(e^{-\frac{1}{2}\phi'(u)} + 1\right) u = e^{z^-}(\sqrt{u} + u) \leq \frac{e^{z^-}}{2}(3u+1),$$

and note that the resulting inequalities both hold even in the case of $u = 0$. Multiplying the terms above by $b^+\mathrm{d}\gamma, b^-\mathrm{d}\gamma$ respectively and integrating over $\mathcal{T}$, the bound (16.18) follows.

It is now straightforward to verify (16.19), since for any $u \geq 0$, $z^+, z^- \in \mathbb{R}$

$$\lim_{n\to\infty}\left(e^{z^+ + \frac{1}{2}\phi_n'(u)} - 1\right) = e^{z^+}\sqrt{u} - 1 = (e^z - 1)\sqrt{u} - (u - \sqrt{u}),$$

and

$$\lim_{n\to\infty}\left(e^{z^- - \frac{1}{2}\phi_n'(u)} - 1\right)u = e^z\sqrt{u} - u = (e^z - 1)\sqrt{u} - (\sqrt{u} - 1).$$

□

Proof of Theorem 16.8. Consider any curve $(\nu, \lambda^+, \lambda^-) \in \mathscr{C}\mathscr{E}$ with $\mathcal{F}_{MF}(\nu_0) < \infty$. By Propositions 16.17 and 16.20 we have $\mathcal{I}_{MF}(\nu, \lambda^+, \lambda^-) = \mathcal{J}_{MF}(\nu, \lambda^+, \lambda^-)$ and the chain rule if either are finite.

In particular, $\mathcal{J} = 0$ if and only if $\lambda_t^\pm = \chi_{\nu_t}^\pm$ and ν_t is a strong solution to (MF), which by Proposition 16.15 exists and is unique, which concludes the proof. □

Chapter 17

Forward Kolmogorov equation

In the introduction, we discussed how the Bolker–Pacala–Dieckmann–Law-model describes a measure-valued process ν_t^n in Γ involving particles being created and annihilated, with the corresponding Forward Kolmogorov equation

$$\partial_t \mathsf{P}_t = Q_n^* \mathsf{P}_t, \tag{$\mathsf{FKE_n}$}$$

where $\mathsf{P}_t \in \mathcal{P}(\Gamma)$ for all $t \in [0, T]$ and Q_n^* is the dual of the infinitesimal generator Q_n with

$$(Q_n F)(\nu) = n \int_{\mathcal{T}} \left(F(\nu + \tfrac{1}{n}\delta_x) - F(\nu)\right) \chi_\nu^+(\mathrm{d}x) + n \int_{\mathcal{T}} \left(F(\nu - \tfrac{1}{n}\delta_x) - F(\nu)\right) \chi_\nu^-(\mathrm{d}x), \tag{17.1}$$

for all $F \in C_c(\Gamma)$. Throughout this chapter the parameter $n > 0$ will be fixed.

In the case of $\mathcal{T} = \mathbb{R}^d$ it was already shown in [FM04] that a measure-valued process with generator Q_n exists using martingale techniques, and is in fact a jump process in Γ corresponding to the jump kernel κ_n shown below. This result was generalized in various subsequent works, see for example [BBC17] for an overview of various modifications and scalings used to establish the canonical equations for adaptive dynamics.

However, for our general setting with $\mathcal{T}$ a compact Polish space, we will take ($\mathsf{FKE_n}$) simply as a starting point, and do not consider the existence or convergence of the measure-valued process ν_t^n itself—even though we will sometimes borrow the language of jump processes for illustration purposes.

Instead, we will apply the framework of Part II.A, in particular the results of Chapters 13 and 14 that deal with unbounded kernels, and adapt it to the interacting particle systems of this chapter. We will state the general version of Theorem 15.5, by constructing the backward kernel with respect to the reference measure Π_n, establishing the equivalence between the corresponding EDP-functional $\mathcal{J}_n$

and rate functional $\mathcal{I}_n$, and characterizing the solutions as their minimizers. Similar to Chapter 16 we first give an overview of the ingredients to state the main results and then leave the proofs for the existence of solutions and the variational characterization to Sections 17.3 and 17.4.

Now, recall from (15.5) the map $L_n : \coprod_{N\geq 0} \mathcal{T}^N \to \Gamma$, which is a rescaled empirical measure mapping, given as

$$L_n(x_1, \dots, x_N) := \frac{1}{n}\sum_{i=1}^{N} \delta_{x_i}, \quad N \in \mathbb{N}, \tag{17.2}$$

with the convention $L_n(\emptyset) = 0$, and let $\Gamma_n \subset \Gamma$ be the space of finite nonnegative discrete measures with common unit weight $\frac{1}{n}$, i.e.

$$\Gamma_n := L_n\left(\coprod_{N\geq 0} \mathcal{T}^N\right). \tag{17.3}$$

Note that for suitable $\mathsf{P} \in \mathcal{P}(\Gamma)$ and $F \in \mathcal{B}(\Gamma)$ the operators Q_n, Q_n^* can be represented as

$$(Q_n F)(\nu) = \int_{\Gamma_n} (F(\eta) - F(\nu))\, \kappa(\nu, \mathrm{d}\eta),$$

$$(Q_n^* \mathsf{P})(\mathrm{d}\nu) = \int_{\eta\in\Gamma_n} \mathsf{P}(\mathrm{d}\eta)\kappa_n(\eta, \mathrm{d}\nu) - \mathsf{P}(\mathrm{d}\nu) \int_{\eta\in\Gamma_n} \kappa_n(\nu, \mathrm{d}\eta),$$

where $\kappa_n(\nu, \cdot) \in \mathcal{M}^+(\Gamma_n)$ for all $\nu \in \Gamma_n$ is a jump kernel over Γ_n given by

$$\kappa_n(\nu, \mathrm{d}\eta) := n \int_{\mathcal{T}} \delta_{\nu+\frac{1}{n}\delta_x}(\mathrm{d}\eta)\, \chi_\nu^+(\mathrm{d}x) + n \int_{\mathcal{T}} \delta_{\nu-\frac{1}{n}\delta_x}(\mathrm{d}\eta)\, \chi_\nu^-(\mathrm{d}x). \tag{17.4}$$

In addition, let us introduce the backward jump kernel $\kappa_n^\dagger$ given by

$$\kappa_n^\dagger(\nu, \mathrm{d}\eta) := n \int_{\mathcal{T}} \delta_{\nu-\frac{1}{n}\delta_x}(\mathrm{d}\eta)\, \bar{\chi}_\nu^{n,+}(\mathrm{d}x) + n \int_{\mathcal{T}} \delta_{\nu+\frac{1}{n}\delta_x}(\mathrm{d}\eta)\, \bar{\chi}_\nu^{n,-}(\mathrm{d}x), \tag{17.5}$$

where

$$\bar{\chi}_\nu^{n,+}(\mathrm{d}x) = b^+[\nu - \tfrac{1}{n}\delta_x](x)\nu(\mathrm{d}x) \qquad \bar{\chi}_\nu^{n,-}(\mathrm{d}x) = b^-[\nu + \tfrac{1}{n}\delta_x](x)\gamma(\mathrm{d}x), \tag{17.6}$$

with the convention that $\bar{\chi}^{n,+}[0] = 0$.

We will consider Poisson measures $\Pi_n \in \mathcal{P}(\Gamma_n)$ induced by the reference measure γ. Namely, with the measure $\pi_n \in \mathcal{P}(\coprod_{N\geq 0} \mathcal{T}^N)$ given by

$$\pi_n := \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^{\infty} \frac{n^N}{N!} \gamma^{\otimes N}, \tag{17.7}$$

we define

$$\Pi_n := (L_n)_\# \pi_n. \tag{17.8}$$

Remark 17.1. In Lemma 17.12 it will be shown that $\kappa_n^\dagger$ is indeed the backward kernel with respect to Π in the sense of Part II.A, i.e.

$$\Pi_n(\mathrm{d}\nu)\kappa_n(\nu, \mathrm{d}\eta) = \Pi_n(\mathrm{d}\eta)\kappa_n^\dagger(\eta, \mathrm{d}\nu). \tag{17.9}$$

□

In particular, κ_n satisfies the detailed balance condition with respect to Π_n if and only if

$$b^-[\nu + \tfrac{1}{n}\delta_x](x) = b^+[\nu](x), \qquad \text{for all } \nu \in \Gamma, \tag{17.10}$$

which for example holds under Assumption 15.10.

To discuss the continuity equation and the dissipation potentials properly, we need to introduce some additional notation. We define the following creation and annihilation operators:

$$\begin{aligned} \mathsf{T}^{n,+} &: \Gamma_n \times \mathcal{T} \to \Gamma_n \times \mathcal{T}, \qquad \mathsf{T}^{n,+}(\nu, x) = (\nu + \tfrac{1}{n}\delta_x, x) =: (\mathsf{T}_x^{n,+}\nu, x), \\ \mathsf{T}^{n,-} &: \Gamma_n \times \mathcal{T} \to \Gamma_n \times \mathcal{T}, \qquad \mathsf{T}^{n,-}(\nu, x) = (\nu - \tfrac{1}{n}\delta_x, x) =: (\mathsf{T}_x^{n,-}\nu, x), \end{aligned} \tag{17.11}$$

with the convention that $\mathsf{T}^{n,-}(\nu, x) = (\nu, x)$ if $x \notin \mathrm{supp}(\nu)$. Note that $\mathsf{T}^{n,-}\circ\mathsf{T}^{n,+} = \mathsf{Id}$ always holds, and $\mathsf{T}^{n,+}\circ\mathsf{T}^{n,-}(\nu, x) = (\nu, x)$ whenever $x \in \mathrm{supp}(\nu)$.

We further define the discrete Γ_n-gradients $\overline{\nabla}^{n,\pm} : C_c(\Gamma_n) \to C_c(\Gamma_n \times \mathcal{T})$:

$$(\overline{\nabla}^{n,\pm} F)(\nu, x) := n(F(\mathsf{T}_x^{n,\pm}\nu) - F(\nu)), \tag{17.12}$$

and the corresponding Γ_n-divergence $\overline{\mathrm{div}}^{n,\pm} : \mathcal{M}^+_{\mathrm{loc}}(\Gamma_n \times \mathcal{T}) \to \mathcal{M}_{\mathrm{loc}}(\Gamma_n)$, dual to $\overline{\nabla}^{n,\pm}$, given by

$$(\overline{\mathrm{div}}^{n,\pm}\mathsf{J}) = n\left(\mathsf{p}^{\Gamma_n}_{\#}\mathsf{J} - (\mathsf{p}^{\Gamma_n}\circ\mathsf{T}^{n,\pm})_{\#}\mathsf{J}\right), \tag{17.13}$$

where $\mathsf{p}^{\Gamma_n} : \Gamma_n \times \mathcal{T} \to \Gamma_n$ denotes the projection to the first variable.

We now consider the familes of curves satisfying

$$\partial_t \mathsf{P}_t + (\overline{\mathrm{div}}^{n,+}\mathsf{J}_t^+) + (\overline{\mathrm{div}}^{n,-}\mathsf{J}_t^-) = 0 \tag{CE_n}$$

in the following appropriate distributional sense.

Definition 17.2 (Continuity equation)**.** A triple $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ satisfies the continuity equation CE_n, if

1. the curve $[0, T] \ni t \mapsto \mathsf{P}_t \in \mathcal{P}(\Gamma_n)$ is narrowly continuous,
2. the Borel family $(\mathsf{J}_t^\pm)_{t\in[0,T]} \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma_n \times \mathcal{T})$ satisfies

$$\mathrm{supp}(\mathsf{J}_t^-) \subseteq \left\{(\nu, x) \ : \ \nu(\mathcal{T}) \geq \tfrac{1}{n},\, x \in \mathrm{supp}(\nu)\right\},$$

3. $\int_0^T \int_{\Gamma_n\times\mathcal{T}} (1+\nu(\mathcal{T})^2)^{-1}\, \mathrm{d}\mathsf{J}_t^\pm \, \mathrm{d}t < \infty$,
4. for every $s,t \in [0,T]$ and all $F \in B_c(\Gamma_n)$

$$\int_{\Gamma_n} F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_{\Gamma_n} F(\nu)\, \mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma_n\times\mathcal{T}} \left((\overline{\nabla}^{n,+} F)\, \mathrm{d}\mathsf{J}_r^+ + (\overline{\nabla}^{n,-} F)\, \mathrm{d}\mathsf{J}_r^- \right) \mathrm{d}r. \tag{17.14}$$

5. for every $s,t \in [0,T]$ and all $F \in B_b(\Gamma_n)$

$$\int_{\Gamma_n} F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_{\Gamma_n} F(\nu)\, \mathrm{d}\mathsf{P}_s \le \int_s^t \int_{\Gamma_n\times\mathcal{T}} \left((\overline{\nabla}^{n,+} F)_+\, \mathrm{d}\mathsf{J}_r^+ + (\overline{\nabla}^{n,-} F)_+\, \mathrm{d}\mathsf{J}_r^- \right) \mathrm{d}r. \tag{17.15}$$

Moreover, the curve $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ satisfies CE_n^* if in addition

$$\sup_{t\in[0,T]} \int_{\Gamma_n} \nu(\mathcal{T})^2\, \mathrm{d}\mathsf{P}_t^n < \infty, \qquad \int_0^T \|\mathsf{J}_t^\pm\|_{TV}\, \mathrm{d}t < \infty. \tag{17.16}$$

□

Here CE_n will be used for the case of detailed balance, applying the framework of Chapter 14, and CE_n^* corresponds to the case of bounded fluxes of Chapter 13.

Remark 17.3. Note that from the mass-moment bound of (17.16) for CE_n^* one can derive via an approximation argument that the continuity equation holds for all bounded $F \in B_b(\Gamma)$, which implies (17.15) for such F.

However, in general, Condition 5 might not directly follow from Condition 4 without additional moment bounds, even though a posteriori 5 would imply those same necessary moment bounds. We will see this in the proof of the equivalence between the rate functional and EDP-functional in the case of detailed balance.

Moreover, because of this, we will only consider uniqueness in the class of weak solutions satisfying (17.15). □

Throughout we will only consider measures $\mathsf{J}^\pm \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma_n \times \mathcal{T})$ if

$$\mathrm{supp}(\mathsf{J}^-) \subseteq \left\{ (\nu, x) \;:\; \nu(\mathcal{T}) \ge \tfrac{1}{n},\, x \in \mathrm{supp}(\nu) \right\}$$

and

$$\int_{\Gamma_n\times\mathcal{T}} (1+\nu(\mathcal{T})^2)^{-1}\, \mathrm{d}\mathsf{J}^\pm < \infty,$$

which we call *admissible* $\mathsf{J}^\pm$. Moreover, since Γ_n is a closed subspace of the Polish space Γ, the extension of P to $\mathcal{P}(\Gamma)$ and the extension of $\mathsf{J}^\pm$ to $\mathcal{M}^+_{\mathrm{loc}}(\Gamma \times \mathcal{T})$ are well-defined. For simplicity we will refer to them as P, $\mathsf{J}^{\pm}$ as well, and drop the n-dependence in most arguments. It is also clear that for any admissible $\mathsf{J}^{\pm}$

$$(\overline{\nabla}^{n,\pm}F)(\nu,x) := n\left(F(\nu \pm \tfrac{1}{n}\delta_x) - F(\nu)\right), \qquad (\nu,x) \in \operatorname{supp}(\mathsf{J}^{\pm})$$

and, in particular, (17.14) is equivalent to

$$\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} \left(n\big(F(\nu+\tfrac{1}{n}\delta_x)-F(\nu)\big)\,\mathrm{d}\mathsf{J}_r^+ + n\big(F(\nu-\tfrac{1}{n}\delta_x)-F(\nu)\big)\,\mathrm{d}\mathsf{J}_r^-\right)\mathrm{d}r.$$

for all $F \in B_c(\Gamma)$.

Remark 17.4. Condition (2) represents the restriction that particles can be deleted, but that measure-valued process ν_t^n cannot obtain elements such as $-\frac{1}{n}\delta$, consistent with the fact that $\nu_t^n \geq 0$ for all times

Moreover, condition (3) reflects the unboundedness of the observed fluxes $\mathsf{J}^{\pm}$ if we do not assume the additional moment condition (17.16), which stems from the unboundedness of the birth/death $\chi_\nu^{\pm}$ in ν. □

Remark 17.5. Whenever $\mathsf{J}^{\pm}$ are of the form

$$\mathsf{J}_t^{\pm}(\mathrm{d}\nu,\mathrm{d}x) = \mathsf{P}_t(\mathrm{d}\nu)\lambda^{\pm}[t,\nu](\mathrm{d}x)$$

with $\lambda^{\pm}[t,\nu] \in \mathcal{M}^+(\mathcal{T})$ for all $\nu \in \Gamma$ and $t \in [0,T]$, the continuity equation (17.14) describes the forward Kolmogorov equation corresponding to an interacting birth/death process with the birth/death kernels $\lambda^{\pm}[t,\nu]$ depending on both time and the empirical measure of the particles ν. The time-dependent jump kernel is then given by

$$\kappa_{n,t}(\mathrm{d}\nu,\mathrm{d}\eta) = n\left(\int_{\mathcal{T}} \delta_{\nu+\frac{1}{n}\delta_x}(\mathrm{d}\eta)\,\lambda^+[t,\nu](\mathrm{d}x) + \int_{\mathcal{T}} \delta_{\nu-\frac{1}{n}\delta_x}(\mathrm{d}\eta)\,\lambda^-[t,\nu](\mathrm{d}x)\right).$$

□

In order to define the dissipation potentials we will introduce some additional objects, which at first sight can seem quite involved. However, it should be noted that all objects are precisely the same objects as in Part II.A, but adapted to the new continuity equation.

Let us introduce the measures $\vartheta_{\mathsf{P}}^{\pm}, \bar{\vartheta}_{\mathsf{P}}^{n,\pm} \in \mathcal{M}_{\mathrm{loc}}^+(\Gamma\times\mathcal{T})$

$$\vartheta_{\mathsf{P}}^{\pm}(\mathrm{d}\nu\,\mathrm{d}x) := \mathsf{P}(\mathrm{d}\nu)\chi_\nu^{\pm}(\mathrm{d}x), \quad \bar{\vartheta}_{\mathsf{P}}^{n,\pm}(\mathrm{d}\nu\,\mathrm{d}x) := \mathsf{P}(\mathrm{d}\nu)\bar{\chi}_\nu^{n,\pm}(\mathrm{d}x). \tag{17.17}$$

Note that for any curve $(\mathsf{P}_t)_{t\in[0,T]}$ the measures $\mathsf{J}_t^{\pm} := \vartheta_{\mathsf{P}_t}^{\pm}$ satisfy the conditions (2) and (3). Moreover, as will be shown in Lemma 17.12, we have

$$\vartheta_{\Pi_n}^{\pm} = \mathsf{T}_{\#}^{n,\mp}\bar{\vartheta}_{\Pi_n}^{n,\pm}. \tag{17.18}$$

from which (17.18) directly follows.

Definition 17.6. Let $\Theta_{\mathsf{P}}^{n,\pm} \in \mathcal{M}_{\mathrm{loc}}(\Gamma \times \mathcal{T})$ be the geometric average of $\vartheta_{\mathsf{P}}^{\pm}$ and $\mathsf{T}_{\#}^{n,\mp}\bar{\vartheta}_{\mathsf{P}}^{n,\pm}$, i.e.

$$\Theta_{\mathsf{P}}^{n,\pm}(\mathrm{d}v, \mathrm{d}x) := \sqrt{\frac{\mathrm{d}\vartheta_{\mathsf{P}}^{\pm}}{\mathrm{d}\Sigma}\frac{\mathrm{d}(\mathsf{T}_{\#}^{n,\mp}\bar{\vartheta}_{\mathsf{P}}^{n,\pm})}{\mathrm{d}\Sigma}}\,\mathrm{d}\Sigma, \tag{17.19}$$

for any dominating measure Σ. We define the following objects:

- The dissipation potential $\mathcal{R}_n : \mathcal{P}(\Gamma) \times \mathcal{M}_{\mathrm{loc}}^{+}(\Gamma \times \mathcal{T})^2 \to [0, +\infty]$ and dual dissipation potential $\mathcal{R}_n^* : \mathcal{P}(\Gamma) \times \mathcal{B}_c(\Gamma \times \mathcal{T})^2 \to \mathbb{R}$,

$$\mathcal{R}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \mathcal{E}\mathrm{nt}(\mathsf{J}^+|\Theta_{\mathsf{P}}^{n,+}) + \mathcal{E}\mathrm{nt}(\mathsf{J}^-|\Theta_{\mathsf{P}}^{n,-}),$$
$$\mathcal{R}_n^*(\mathsf{P}, \omega^+, \omega^-) := \int_{\Gamma\times\mathcal{T}} (e^{\omega^+} - 1)\,\mathrm{d}\Theta_{\mathsf{P}}^{n,+} + \int_{\Gamma\times\mathcal{T}} (e^{\omega^-} - 1)\,\mathrm{d}\Theta_{\mathsf{P}}^{n,-}.$$

- The free energy $\mathcal{F}_n : \mathcal{P}(\Gamma) \to [0, +\infty]$,

$$\mathcal{F}_n(\mathsf{P}) := \tfrac{1}{2n}\mathcal{E}\mathrm{nt}(\mathsf{P}|\Pi_n).$$

- The Fisher information $\mathcal{D}_n : \mathcal{P}(\Gamma) \to (-\infty, +\infty]$,

$$\mathcal{D}_n(\mathsf{P}) := H^2(\vartheta_{\mathsf{P}}^+, \mathsf{T}_{\#}^{n,-}\bar{\vartheta}_{\mathsf{P}}^{+,n}) + H^2(\vartheta_{\mathsf{P}}^-, \mathsf{T}_{\#}^{n,+}\bar{\vartheta}_{\mathsf{P}}^{-,n}) + \frac{1}{2}\int_{\Gamma\times\mathcal{T}} \mathrm{d}\left(\vartheta_{\mathsf{P}}^+ + \vartheta_{\mathsf{P}}^- - \mathsf{T}_{\#}^{n,-}\bar{\vartheta}_{\mathsf{P}}^{+,n} - \mathsf{T}_{\#}^{n,+}\bar{\vartheta}_{\mathsf{P}}^{-,n}\right),$$

  where the third term is defined only if finite, or zero in the case of detailed balance.

- The EDP-functional $\mathcal{J}_n$ on CE_n for all curves with $\mathcal{F}_n(\mathsf{P}_0) < \infty$,

$$\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{R}_n(\mathsf{P}_t, \mathsf{J}_t^+, \mathsf{J}_t^-)\,\mathrm{d}t + \mathcal{F}_n(\mathsf{P}_T) - \mathcal{F}_n(\mathsf{P}_0) + \int_0^T \mathcal{D}_n(\mathsf{P}_t)\,\mathrm{d}t.$$

- The Lagrangian $\mathcal{L}_n : \mathcal{P}(\Gamma) \times \mathcal{M}_{\mathrm{loc}}^{+}(\Gamma \times \mathcal{T})^2 \to [0, +\infty]$ and its dual, the Hamiltonian $\mathcal{H}_n : \mathcal{P}(\Gamma) \times \mathcal{B}_c(\Gamma \times \mathcal{T})^2 \to \mathbb{R}$,

$$\mathcal{L}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \mathcal{E}\mathrm{nt}(\mathsf{J}^+|\vartheta_{\mathsf{P}}^+) + \mathcal{E}\mathrm{nt}(\mathsf{J}^-|\vartheta_{\mathsf{P}}^-),$$
$$\mathcal{H}_n(\mathsf{P}, \omega^+, \omega^-) := \int_{\Gamma\times\mathcal{T}} (e^{\omega^+} - 1)\,\mathrm{d}\vartheta_{\mathsf{P}}^+ + \int_{\Gamma\times\mathcal{T}} (e^{\omega^-} - 1)\,\mathrm{d}\vartheta_{\mathsf{P}}^-.$$

- The rate functional $\mathcal{J}_n : \mathsf{CE}_n \to [0, +\infty]$,

$$\mathcal{I}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{L}_n(\mathsf{P}_t, \mathsf{J}_t^+, \mathsf{J}_t^-)\,\mathrm{d}t.$$

□

Remark 17.7. It can be shown that for the correction term of the Fisher information we have the equivalence

$$\frac{1}{2}\int_{\Gamma\times\mathcal{T}} \mathrm{d}\left(\vartheta^+_{\mathsf{P}} + \vartheta^-_{\mathsf{P}} - \mathsf{T}^{n,-}_{\#}\bar{\vartheta}^{+,n}_{\mathsf{P}} - \mathsf{T}^{n,+}_{\#}\bar{\vartheta}^{-,n}_{\mathsf{P}}\right) = \frac{1}{2}\int_{\Gamma\times\mathcal{T}} \mathrm{d}\left(\vartheta^+_{\mathsf{P}} + \vartheta^-_{\mathsf{P}} - \bar{\vartheta}^{+,n}_{\mathsf{P}} - \mathsf{T}^{n,+}_{\#}\bar{\vartheta}^{-,n}_{\mathsf{P}}\right)$$

and, in particular, vanishes under the detailed balance condition (17.10). □

Remark 17.8. Formally

$$\Theta^{n,\pm}_{\mathsf{P}}(v,x) = \sqrt{(\mathsf{P}(v)\chi^{\pm}_{v})(\mathsf{P}(v \pm \tfrac{1}{n}\delta_x)\bar{\chi}^{n,\pm}[v \pm \tfrac{1}{n}\delta_x]},$$

i.e. it represents the geometric mean of the expected forward and Π_n-backward fluxes going along the transition $v \leftrightarrow v \pm \frac{1}{n}\delta_x$.

In addition, due to (17.18) the measures $\Theta^{n,\pm}_{\mathsf{P}}$ simplify whenever $\mathsf{P} \ll \Pi_n$, i.e. if $\mathrm{d}\mathsf{P} = U\mathrm{d}\Pi_n$ we have

$$\Theta^{n,\pm}_{\mathsf{P}}(\mathrm{d}v,\mathrm{d}x) = \sqrt{U(v)U(v \pm \tfrac{1}{n}\delta_x)}\,\vartheta^{\pm}_{\Pi_n}(\mathrm{d}v,\mathrm{d}x).$$

□

Finally, we also introduce a version for net fluxes.

Definition 17.9. The *upward* net flux $\mathsf{J}^{\mathrm{net}}$ is defined as

$$\mathsf{J}^{\mathrm{net}} := \mathsf{J}^+ - \mathsf{T}^{n,-}_{\#}\mathsf{J}^-$$

□

Note that $\mathsf{J}^{\mathrm{net}}(v,x)$ can be interpreted as the net flux along the jump $v \leftrightarrow v+\frac{1}{n}\delta_x$. The continuity equation for the net flux reduces to, for $F \in B_c(\Gamma)$

$$\int_\Gamma F(v)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(v)\,\mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} n\big(F(v + \tfrac{1}{n}\delta_x) - F(v)\big)\,\mathrm{d}\mathsf{J}^{\mathrm{net}}_r\,\mathrm{d}r$$

Finally, in light of Remark 17.3, we consider our notions of solutions, as minimizers of the rate functional $\mathcal{I}$.

Definition 17.10. A curve $(\mathsf{P}_t)_{t\in[0,T]} \subset \mathcal{P}(\Gamma)$ is a *weak solution* to $(\mathsf{FKE_n})$ if $\operatorname{supp}\mathsf{P}_t \in \Gamma_n$ for all $t \in [0,T]$, P_t is continuous in the narrow topology and for all $s,t \in [0,T]$, and all $F \in B_c(\Gamma)$,

$$\int_\Gamma F(v)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(v)\,\mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} \left((\overline{\nabla}^{n,+}F)\,\mathrm{d}\vartheta^+_{\mathsf{P}_r} + (\overline{\nabla}^{n,-}F)\,\mathrm{d}\vartheta^-_{\mathsf{P}_r}\right)\mathrm{d}r. \tag{17.20}$$

Moreover, it is a $\mathcal{I}$-*solution* if in addition for any bounded $F \in B_b(\Gamma)$ and $t \in [0, T]$ we have the inequality

$$\int_\Gamma F(\nu)\,\mathrm{dP}_t - \int_\Gamma F(\nu)\,\mathrm{dP}_0 \le \int_0^t \int_{\Gamma\times\mathcal{T}} \left((\overline{\nabla}^{n,+} F)_+ \,\mathrm{d}\vartheta^+_{\mathsf{P}_r} + (\overline{\nabla}^{n,-} F)_+ \,\mathrm{d}\vartheta^-_{\mathsf{P}_r} \right) \mathrm{d}r. \tag{17.21}$$

□

We are now in a position to give the general version of Theorem 15.5.

Theorem 17.11. *For any* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_n$ *with* $\mathcal{F}_n(\mathsf{P}_0) < \infty$ *we have*

$$\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = \mathcal{I}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \tag{17.22}$$

In particular, $\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \ge 0$. *Moreover, if* P_0 *is such that* $\mathcal{F}_n(\mathsf{P}_0) < \infty$ *and*

$$\int_\Gamma \nu(\mathcal{T})^2 \,\mathrm{dP}_0 < \infty, \tag{17.23}$$

then

$$\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = 0 \implies \begin{cases} \mathsf{P}_t \text{ is the unique } \mathcal{I}\text{-solution to } (\mathsf{FKE_n}), \\ \mathsf{J}^\pm_t = \vartheta^\pm_{\mathsf{P}_t} \quad \text{for a.e. } t \in [0, T]. \end{cases} \tag{17.24}$$

Finally, suppose that the detailed balance condition (17.10) *holds. Then the above statements are valid without any necessary moment conditions, i.e. merely for any* P_0 *with* $\mathcal{F}_n(\mathsf{P}_0) < \infty$ *and curves* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$.

It should be noted that the proof of uniqueness of $\mathcal{I}$-solutions follows from the uniqueness of EDP-solutions if $\mathcal{F}_n$ is strictly convex, as seen in Part II.A, and the fact that $\mathcal{I}$-solutions propagate moments.

We will first establish in Section 17.1 that $\kappa^\dagger_n$ are indeed the backward kernels, and in Section 17.2 further relate fluxes to the edge fluxes of Part II.A. We will then briefly consider existence of solutions and propagation of moments in Section 17.3, before we give the of proof of Theorem 17.11 in Section 17.4.

17.1 Backward kernels

Before we can use the framework of Part II.A, let us first verify that the identity (17.18) holds and $\kappa^\dagger$ is indeed the Π_n-backward kernel of κ_n. To make it even more precise, we define for technical reasons the kernels

$$\kappa^\pm_n(\nu, \mathrm{d}\eta) := n \int_{\mathcal{T}} \delta_{\nu \pm \frac{1}{n}\delta_x}(\mathrm{d}\eta)\, \chi^{n,\pm}_\nu(\mathrm{d}x)$$

$$\kappa^{\dagger,\pm}_n(\nu, \mathrm{d}\eta) := n \int_{\mathcal{T}} \delta_{\nu \mp \frac{1}{n}\delta_x}(\mathrm{d}\eta)\, \bar{\chi}^{n,\pm}_\nu(\mathrm{d}x)$$

Note that $\kappa^\pm_n$ are mutually singular, $\kappa_n = \kappa^+_n + \kappa^-_n$, and $\kappa^\dagger_n = \kappa^{\dagger,+}_n + \kappa^{\dagger,-}_n$.

Lemma 17.12 (Backward kernel)**.** *We have*

$$\vartheta^{\pm}_{\Pi_n} = \mathsf{T}^{n,\mp}_{\#}\bar{\vartheta}^{n,\pm}_{\Pi_n}. \tag{17.25}$$

In particular,

$$\begin{aligned}\Pi_n(\mathrm{d}\nu)\kappa_n^{\pm}(\nu,\mathrm{d}\eta) &= \Pi_n(\mathrm{d}\eta)\kappa_n^{\dagger,\pm}(\eta,\mathrm{d}\nu),\\ \Pi_n(\mathrm{d}\nu)\kappa_n(\nu,\mathrm{d}\eta) &= \Pi_n(\mathrm{d}\eta)\kappa_n^{\dagger}(\eta,\mathrm{d}\nu).\end{aligned} \tag{17.26}$$

Proof. Let $\omega \in B_c(\Gamma\times\mathcal{T})$. Recall that

$$\bar{\chi}^{n,+}_\nu(\mathrm{d}x) = b^+[\nu - \tfrac{1}{n}\delta_x](x)\nu(\mathrm{d}x) \qquad \bar{\chi}^{n,-}_\nu(\mathrm{d}x) = b^-[\nu + \tfrac{1}{n}\delta_x](x)\gamma(\mathrm{d}x),$$

and note that to obtain (17.25) we need

$$\int_{\Gamma\times\mathcal{T}} \omega(\nu,x)b^+[\nu](x)\gamma(\mathrm{d}x)\Pi_n(\mathrm{d}\nu) = \int_{\Gamma\times\mathcal{T}} \omega(\mathsf{T}^{n,-}_x\nu,x)b^+[\mathsf{T}^{n,-}_x\nu](x)\nu(\mathrm{d}x)\Pi_n(\mathrm{d}\nu),$$

$$\int_{\Gamma\times\mathcal{T}} \omega(\nu,x)b^-[\nu](x)\nu(\mathrm{d}x)\Pi_n(\mathrm{d}\nu) = \int_{\Gamma\times\mathcal{T}} \omega(\mathsf{T}^{n,+}_x\nu,x)b^-[\mathsf{T}^{n,+}_x\nu](x)\gamma(\mathrm{d}x)\Pi_n(\mathrm{d}\nu),$$

We will establish the first equality, and the one for b^- follows via similar argument but in reverse. For simplicity, for any $N \geq 0$ and the collection of variables $(x_1,\dots,x_N)$ we will use the shorthand $L_n^N := L_n(x_1,\dots,x_N)$ and $\gamma^{\otimes N} := \gamma(\mathrm{d}x_1)\cdots\gamma(\mathrm{d}x_N)$. Note that by substituting for the measure Π_n we have

$$\begin{aligned}&\int_{\Gamma\times\mathcal{T}} \omega(\nu,x)b_\nu(x)\gamma(\mathrm{d}x)\Pi_n(\mathrm{d}\nu)\\ &\qquad = \sum_{N=0}\frac{n^N}{N!}\int_{\mathcal{T}\times\mathcal{T}^N}\omega\left(L_n^N,x\right)b^+[L_n^N](x)\gamma(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes N},\end{aligned}$$

$$\begin{aligned}&\int_{\Gamma\times\mathcal{T}} \omega(\mathsf{T}^{n,-}_x\nu,x)b^+[\mathsf{T}^{n,-}_x\nu](x)\nu(\mathrm{d}x)\Pi_n(\mathrm{d}\nu)\\ &\qquad = \sum_{N=0}\frac{n^N}{N!}\int_{\mathcal{T}\times\mathcal{T}^N}\omega\left(\mathsf{T}^{n,-}_x L_n^N,x\right)b^+\left[\mathsf{T}^{n,-}_x L_n^N\right](x)L_n^N(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes N}\\ &\qquad = \sum_{N=0}\frac{n^{N+1}}{(N+1)!}\int_{\mathcal{T}\times\mathcal{T}^{N+1}}\omega\left(\mathsf{T}^{n,-}_x L_n^{N+1},x\right)b^+\left[\mathsf{T}^{n,-}_x L_n^{N+1}\right](x)L_n^{N+1}(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes(N+1)},\end{aligned}$$

where the substitution $N \to N+1$ in the last equality follows from the fact that $L_n^0 = 0$ and hence the summation starts effectively from $N = 1$. Therefore it remains to show that for every $N \geq 0$,

$$\begin{aligned}(N+1)&\int_{\mathcal{T}\times\mathcal{T}^N}\omega\left(L_n^N,x\right)b^+[L_n^N](x)\gamma(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes N}\\ &= n\int_{\mathcal{T}\times\mathcal{T}^{N+1}}\omega(\mathsf{T}^{n,-}_x L_n^{N+1},x)b^+\left[\mathsf{T}^{n,-}_x L_n^{N+1}\right](x)L_n^{N+1}(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes(N+1)}.\end{aligned}$$

However, this follows by using the symmetry of $\gamma(\mathrm{d}x)\gamma^{\otimes N}$, since by relabeling the variables $(x, x_1, \dots, x_N)$ as $(x_i, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{N+1})$ we deduce

$$
\begin{aligned}
(N+1) \int_{\mathcal{T}\times\mathcal{T}^N} & \omega\left(L_n^N, x\right) b^+[L_n^N](x)\gamma(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes N} \\
&= \sum_{i=1}^{N+1} \int_{\mathcal{T}^{N+1}} \omega\left(\mathsf{T}_{x_i}^{n,-} L_n^{N+1}, x_i\right) b^+\left[\mathsf{T}_{x_i}^{n,-} L_n^{N+1}\right](x_i)\,\mathrm{d}\gamma^{\otimes(N+1)} \\
&= n \int_{\mathcal{T}\times\mathcal{T}^{N+1}} \omega\left(\mathsf{T}_x^{n,-} L_n^{N+1}, x\right) b^+\left[\mathsf{T}_x^{n,-} L_n^{N+1}\right](x) L_n^{N+1}(\mathrm{d}x)\,\mathrm{d}\gamma^{\otimes(N+1)}.
\end{aligned}
$$

Finally, note that to obtain

$$
\int_{\Gamma^2} G(\nu,\eta)\kappa_n^{\pm}(\nu,\mathrm{d}\eta)\Pi_n(\mathrm{d}\nu) = \int_{\Gamma^2} G(\nu,\eta)\kappa_n^{\dagger,\pm}(\eta,\mathrm{d}\nu)\Pi_n(\mathrm{d}\eta)
$$

it is sufficient to show that

$$
\begin{aligned}
\int_{\Gamma\times\mathcal{T}} G(\nu, \mathsf{T}_x^{n,+}\nu) b^+[\nu](x)\gamma(\mathrm{d}x)\Pi_n(\mathrm{d}\nu) &= \int_{\Gamma\times\mathcal{T}} G(\mathsf{T}_x^{n,-}\nu, \nu) b^+[\mathsf{T}_x^{n,-}\nu](x)\nu(\mathrm{d}x)\Pi_n(\mathrm{d}\nu), \\
\int_{\Gamma\times\mathcal{T}} G(\nu, \mathsf{T}_x^{n,-}\nu) b^-[\nu](x)\nu(\mathrm{d}x)\Pi_n(\mathrm{d}\nu) &= \int_{\Gamma\times\mathcal{T}} G(\mathsf{T}_x^{n,+}\nu, \nu) b^-[\mathsf{T}_x^{n,+}\nu](x)\gamma(\mathrm{d}x)\Pi_n(\mathrm{d}\nu),
\end{aligned}
$$

but this follows directly from the previous calculation by substituting $\omega(\nu, x) = G(\nu, \mathsf{T}_x^{n,\pm}\nu)$. □

17.2 Edge flux formulation

We will now establish the necessary estimates to apply Part II.A, by relating the fluxes $\mathsf{J}^{\pm} \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma\times\mathcal{T})$ to edge fluxes $j \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma^2)$ and vice versa. Let us define

$$
a(\nu) := (1 + \nu(\mathcal{T})^2)^{-1}, \qquad \xi_k(\nu) := \alpha_k(\nu(\mathcal{T})) \tag{17.27}
$$

where $\xi_k \in C_c(\Gamma)$ defined via

$$
\alpha_k(z) := \begin{cases} 1, & 0 \le z \le k, \\ 2 - \dfrac{z}{k}, & k \le z \le 2k, \\ 0, & z \ge 2k. \end{cases}
$$

Remark 17.13. Note that $0 \le \alpha_k \le 1$, $|\alpha_k'(z)| = 0$ for all $z < k$ or $z > 2k$, $|\alpha_k'(z)| \le 2/z$ for all $z \ge 0$, and that α_k converges monotonically to 1 as $k \to \infty$.

Moreover, note that the decay of $|\alpha_k'(z)| \leq 2/z$ is optimal in the sense that for any $\varepsilon > 0, L > 0$ there does not exists a sequence of compact functions $\tilde{\alpha}_k$ satisfying $0 \leq \alpha_k \leq 1$, $\alpha_k(z) \uparrow 1$, and $|\alpha_k'(z)| \leq L/z^{1+\varepsilon}$. This can be seen by the fact that

$$\lim_{a\to\infty} \int_a^\infty \frac{1}{z^{\varepsilon+1}} \, \mathrm{d}z = 0.$$

□

We can now give analogues of the conditions (10.14) and (14.6).

Lemma 17.14. *Fix any $\nu \in \Gamma$. Then we have the following estimates:*

(i) For every $p \geq 1$ there exist constants $C_{n,p}$ that are nondecreasing in n and such that

$$\int_\Gamma (\eta(\mathcal{T})^p - \nu(\mathcal{T})^p)_+ \kappa_n(\nu, \mathrm{d}\eta) \leq C_{n,p}(\nu(\mathcal{T})^p + 1). \tag{17.28}$$

(ii) Let $M_n = \max(1 + 2/n^2, 2)M$. Then

$$\max\left(a(\nu) \int_\Gamma \kappa_n(\nu, \mathrm{d}\eta),\ a(\nu) \int_\Gamma \kappa_n^\dagger(\nu, \mathrm{d}\eta), \int_\Gamma a(\nu)\kappa_n(\nu, \mathrm{d}\eta) \right) \leq 2nM_n. \tag{17.29}$$

(iii) Let k_n^* be a sufficiently large constant such that $n|\alpha_k(z \pm \frac{1}{n}) - \alpha_k(z)| \leq 3/(1+z)$ for all $z \geq 0$ and $k \geq k_n^*$. Then for every $k \geq k_n^*$,

$$\int_\Gamma |\overline{\nabla}\xi_k|^2 \kappa_n(\nu, \mathrm{d}\eta) \leq 18nM, \qquad \int_\Gamma |\overline{\nabla}\xi_k| \kappa_n(\nu, \mathrm{d}\eta) \leq 62nM(1 + \nu(\mathcal{T})). \tag{17.30}$$

Remark 17.15. The constant $M_n = \max(1 + 2/n^2, 2)M$ will pop up in various of our estimates, and stems from the observation that $M \leq M_n$ and

$$(1 + (\nu(\mathcal{T}) + \tfrac{1}{n})^2) \leq \max(1 + 2/n^2, 2)(1 + \nu(\mathcal{T})^2) \qquad \text{for all } \nu \in \Gamma,$$

due to the inequality $1 + (\frac{1}{n} + z)^2 \leq 1 + \frac{2}{n^2} + 2z^2$ for all $z \geq 0$. □

Proof. (i) Let $F_p := f_p(\nu(\mathcal{T}))$ with $f_p(z) = z^p$. Then

$$\int_\Gamma (\overline{\nabla} F_p)_+ \kappa_n(\nu, \mathrm{d}\eta) = \int_\mathcal{T} (\overline{\nabla}^{n,+} F_p)_+ b_\nu^+(\mathrm{d}x)\gamma(\mathrm{d}x) + \int_\mathcal{T} (\overline{\nabla}^{n,-} F_p)_+ b_\nu^-(x)\nu(\mathrm{d}x).$$

But $(\overline{\nabla}^{n,-} F_p)_+ \leq 0$ since f_p is nondecreasing. Moreover, $b_\upsilon^+(x)\gamma(\mathcal{T}) \leq M'\gamma(\mathcal{T})\nu(\mathcal{T})$ by assumption, and hence

$$\int_\mathcal{T} (\overline{\nabla}^{n,+} F_p)_+ b_\nu^+(\mathrm{d}x)\gamma(\mathrm{d}x) \leq p\nu(\mathcal{T})(\nu(\mathcal{T}) + \tfrac{1}{n})^{p-1} \leq C_{n,p}(1 + \nu(\mathcal{T})^p),$$

for suitable constants $C_{n,p}$ that are nondecreasing in n.

(ii) Recall that $\chi_\nu^\pm(\mathcal{T}) \le M(1+\nu(\mathcal{T})^2)$, and hence

$$\int_\Gamma \kappa_n(\nu, \mathrm{d}\eta) \le 2nM(1+\nu(\mathcal{T})^2).$$

Therefore we can write

$$\begin{aligned}\int_\Gamma (1+\eta(\mathcal{T})^2)^{-1}\kappa(\nu,\mathrm{d}\eta) &\le \frac{nM(1+\nu(\mathcal{T})^2)}{1+(\nu(\mathcal{T})+\frac{1}{n})^2} + \frac{nM(1+\nu(\mathcal{T})^2)}{1+(\nu(\mathcal{T})-\frac{1}{n})^2}1_{\nu(\mathcal{T})\ge\frac{1}{n}}\\ &\le nM + \sup_{\nu\in\Gamma}\frac{M(1+(\nu(\mathcal{T})+\frac{1}{n})^2)}{1+\nu(\mathcal{T})^2} \le 2nM_n\end{aligned}$$

The calculation for $\bar{\chi}^{n,\pm}$ and κ_n follow in a similar fashion.

(iii) Next, note that $|\overline{\nabla}^{n,\pm}\xi_k|(\nu,x) \le 3/(1+\nu(\mathcal{T}))$ for any $k \ge k_n^*$, and $(1+\nu(\mathcal{T})^2) \le (1+\nu(\mathcal{T}))^2$. Therefore

$$\begin{aligned}\int_\Gamma |\overline{\nabla}\xi_k|^2\kappa_n(\nu,\mathrm{d}\eta) &= n\int_{\mathcal{T}} |\overline{\nabla}^{n,+}\xi_k|^2\chi_\nu^+(\mathrm{d}x) + n\int_{\mathcal{T}} |\overline{\nabla}^{n,-}\xi_k|^2\chi_\nu^-(\mathrm{d}x)\\ &\le \frac{18n}{(1+\nu(\mathcal{T}))^2}M(1+\nu(\mathcal{T})^2) \le 18nM,\end{aligned}$$

and the final desired inequality follows similarly.

□

The following lemma allows us to jump back and forth between our mass fluxes $\mathsf{J}^\pm$ and the edge fluxes j.

Lemma 17.16. *Consider any* $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) \in \mathsf{CE}_n$*. Define for every* $t \in [0,T]$ *the measure* $j_t \in \mathcal{M}_{\mathrm{loc}}(\Gamma^2)$ *as*

$$j_t(\mathrm{d}\nu,\mathrm{d}\eta) := n\int_{\mathcal{T}} \delta_{\nu+\frac{1}{n}\delta_x}(\mathrm{d}\eta)\mathsf{J}_t^+(\mathrm{d}\nu,\mathrm{d}x) + n\int_{\mathcal{T}} \delta_{\nu-\frac{1}{n}\delta_x}(\mathrm{d}\eta)\mathsf{J}_t^-(\mathrm{d}\nu,\mathrm{d}x). \tag{17.31}$$

Then (P,j) *for any* $F \in B_c(\Gamma)$ *and any* $s,t \in [0,T]$ *we have*

$$\int_\Gamma F\mathrm{d}\mathsf{P}_t - \int_\Gamma F\mathrm{d}\mathsf{P}_s = \int_s^t\int_{\Gamma^2}(F(\eta)-F(\nu))\,j_r(\mathrm{d}\nu,\mathrm{d}\eta)\,\mathrm{d}r, \tag{17.32}$$

and moreover

$$\int_0^T\int_{\Gamma^2}(a(\nu)+a(\eta))\,\mathrm{d}j_t\,\mathrm{d}t < \infty. \tag{17.33}$$

Vice versa, suppose that (P,j) *is such that* (17.32) *and* (17.33) *are satisfied, and* $j_t \ll \mathsf{P}_t\kappa_n$ *with* $j_t(\mathrm{d}\nu,\mathrm{d}\eta) = g(\nu,\eta)\mathsf{P}_t(\mathrm{d}\nu)\kappa_n(\nu,\mathrm{d}\eta)$*. Define* $\mathsf{J}_t^\pm$ *by*

$$\mathsf{J}_t^\pm(\mathrm{d}\nu,\mathrm{d}x) = g(\nu,\nu\pm\tfrac{1}{n}\delta_x)\chi_\nu^\pm(\mathrm{d}x)\mathsf{P}_t(\mathrm{d}\nu). \tag{17.34}$$

Then (17.31) *holds and* $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) \in \mathsf{CE}_n$*.*

Proof. Fix a curve $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$. Then we find for every $t \in [0, T]$ that

$$\int_{\Gamma^2} a(\nu)\, \mathrm{d}j = n \int_{\Gamma\times\mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1} (\mathsf{J}_t^+(\mathrm{d}\nu, \mathrm{d}x) + \mathsf{J}_t^-(\mathrm{d}\nu, \mathrm{d}x)). \tag{17.35}$$

Moreover, proceeding as in the proof of Lemma 17.14,

$$\begin{aligned}\int_{\Gamma^2} a(\eta)\, \mathrm{d}j &\le n \int_{\Gamma\times\mathcal{T}} a(\nu + \tfrac{1}{n}\delta_x) \mathsf{J}_t^+(\mathrm{d}\nu, \mathrm{d}x) + n \int_{\Gamma\times\mathcal{T}} a(\nu - \tfrac{1}{n}\delta_x) \mathsf{J}_t^-(\mathrm{d}\nu, \mathrm{d}x) \\ &\le 2nM_n \int_{\Gamma\times\mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1} (\mathsf{J}_t^+(\mathrm{d}\nu, \mathrm{d}x) + \mathsf{J}_t^-(\mathrm{d}\nu, \mathrm{d}x)).\end{aligned}$$

Integrating over the interval $[0, T]$ we obtain (17.33). Now, to derive (17.32), choose a $F \in B_c(\Gamma)$ and note that by the continuity equation (17.14)

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} \left((\overline{\nabla}^{n,+} F)\, \mathrm{d}\mathsf{J}_r^+ + (\overline{\nabla}^{n,-} F)\, \mathrm{d}\mathsf{J}_r^- \right) \mathrm{d}r.$$

However, by definition of j_r,

$$\begin{aligned}\int_{\Gamma^2} (F(y) - F(x))\, \mathrm{d}j_r = n &\int_{\Gamma\times\mathcal{T}} (F(\nu + \tfrac{1}{n}\delta_x) - F(\nu)) \mathsf{J}_r^+(\mathrm{d}\nu, \mathrm{d}x) \\ &+ n \int_{\Gamma\times\mathcal{T}} (F(\nu - \tfrac{1}{n}\delta_x) - F(\nu)) \mathsf{J}_r^+(\mathrm{d}\nu, \mathrm{d}x),\end{aligned}$$

which provides the desired equivalence.

The reverse implication follows via a similar argument, and the fact that for any $w \in B_c(\Gamma^2)$

$$\begin{aligned}n \int_{\Gamma^2} w(\nu, \eta) \int_{\mathcal{T}} \delta_{\nu \pm \frac{1}{n}\delta_x}(\mathrm{d}\eta) \mathsf{J}_t^\pm(\mathrm{d}\nu, \mathrm{d}x) &= n \int_{\Gamma\times\mathcal{T}} w(\nu, \nu \pm \tfrac{1}{n}\delta_x) g(\nu, \nu \pm \tfrac{1}{n}\delta_x) \chi_\nu^\pm(\mathrm{d}x) \mathsf{P}_t(\mathrm{d}\nu) \\ &= \int_{\Gamma^2} w(\nu, \eta) g(\nu, \eta) \kappa_n^\pm(\nu, \mathrm{d}\eta) \mathsf{P}_t(\mathrm{d}\nu).\end{aligned}$$

□

With the transformations (17.31) and (17.34) in hand, we can now relate the building blocks of the EDP-functional and rate functional of Part II.A to $\mathcal{J}_n$ and $\mathcal{I}_n$.

Lemma 17.17. *Let* $\mathsf{P}, \mathsf{J}^\pm, j$ *satisfy the conditions of Lemma 17.16 (dropping the time-dependence). Define*

$$\mathcal{R}(\mathsf{P}, j; \Pi_n, \kappa_n) := \mathcal{E}\mathrm{nt}\left(j \middle| \sqrt{(\mathsf{P}\kappa_n) \mathsf{s}_\#(\mathsf{P}\kappa_n^\dagger)} \right), \qquad \mathcal{L}(\mathsf{P}, j; \kappa_n) := \mathcal{E}\mathrm{nt}(j | \mathsf{P}\kappa_n),$$

$$\mathcal{D}(\mathsf{P}; \Pi_n, \kappa_n) := H^2(\mathsf{P}\kappa_n, \mathsf{s}_\#(\mathsf{P}\kappa_n^\dagger)) + \tfrac{1}{2} \int_{\Gamma^2} \mathrm{d}(\mathsf{P}\kappa_n - \mathsf{P}\kappa_n^\dagger), \qquad \mathcal{F}(\mathsf{P}; \Pi_n) := \mathcal{E}\mathrm{nt}(\mathsf{P} | \Pi_n).$$

Then

$$\mathcal{R}(\mathsf{P}, j; \Pi_n, \kappa_n) = n\mathcal{R}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-), \qquad \mathcal{L}(\mathsf{P}, j; \kappa_n) = n\mathcal{L}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-),$$
$$\mathcal{D}(\mathsf{P}; \Pi_n, \kappa_n) = n\mathcal{D}_n(\mathsf{P}), \qquad \mathcal{F}(\mathsf{P}; \Pi_n) = n\mathcal{F}_n(\mathsf{P}).$$

Remark 17.18. If for suitable curves (P, j) and related $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ we define

$$\mathcal{J}(\rho, j; \Pi_n, \kappa_n) := \int_0^T \big(\mathcal{R}(\mathsf{P}_t, j_t; \Pi_n, \kappa_n) + \mathcal{D}(\mathsf{P}_t; \Pi_n, \kappa_n)\big)\, \mathrm{d}t + \mathcal{F}(\mathsf{P}_T; \Pi_n) - \mathcal{F}(\mathsf{P}_0; \Pi_n)$$
$$\mathcal{I}(\rho, j; \Pi_n, \kappa_n) := \int_0^T \mathcal{L}(\mathsf{P}_t, j_t; \kappa_n)\, \mathrm{d}t$$

then it is clear that

$$\mathcal{J}(\rho, j; \Pi_n, \kappa_n) = n\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-), \qquad \mathcal{I}(\rho, j; \kappa_n) = n\mathcal{I}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).$$

As shown in Chapter 19 we will have convergence of $\mathcal{J}_n, \mathcal{I}_n$ for suitable converging sequences of curves $(\mathsf{P}^n, \mathsf{J}^{n,\pm})$, which implies a blow up of $\mathcal{J}(\cdot; \Pi_n, \kappa_n)$ and $\mathcal{I}(\cdot; \Pi_n, \kappa_n)$ of the order n except if the limiting curve is the solution to the Liouville equation. This can be related this to large deviations of the underlying process, see Appendix G. □

Proof. We will only show one part of equivalence for $\mathcal{R}_n$, since the rest of the identities follow in a similar fashion. Suppose that $\mathcal{R}_n$ is finite and $\mathsf{J}^\pm(\mathrm{d}\nu, \mathrm{d}x) = G^+(\nu, x)\Theta_{\mathsf{P}}^{n,\pm}(\mathrm{d}\nu, \mathrm{d}x)$. Define j as in (17.31). Moreover, note that

$$\sqrt{(\mathsf{P}\kappa_n)\mathsf{s}_\#(\mathsf{P}\kappa_n^\dagger)} = \sqrt{(\mathsf{P}\kappa_n^+)\mathsf{s}_\#(\mathsf{P}\kappa_n^{\dagger,+})} + \sqrt{(\mathsf{P}\kappa_n^-)\mathsf{s}_\#(\mathsf{P}\kappa_n^{\dagger,-})},$$
$$\sqrt{(\mathsf{P}\kappa_n^+)\mathsf{s}_\#(\mathsf{P}\kappa_n^{\dagger,+})} \perp \sqrt{(\mathsf{P}\kappa_n^-)\mathsf{s}_\#(\mathsf{P}\kappa_n^{\dagger,-})}$$

where the two terms are supported on the two sets

$$\{(\nu, \eta) : \eta = \nu + \tfrac{1}{n}\delta_x, \nu \in \Gamma_n, x \in \mathcal{T}\}, \{(\nu, \eta) : \eta = \nu - \tfrac{1}{n}\delta_x, \nu \in \Gamma_n, x \in \mathrm{supp}(\nu)\}$$

respectively. Define the measures $\hat{j}^\pm$, with $\hat{j}^+ \perp \hat{j}^+$ and $\hat{j} := \hat{j}^+ + \hat{j}^-$, by

$$\frac{\mathrm{d}\hat{j}^\pm}{\mathrm{d}\sqrt{(\mathsf{P}\kappa_n^\pm)\mathsf{s}_\#(\mathsf{P}\kappa_n^{\dagger,\pm})}}(\nu, \nu \pm \tfrac{1}{n}\delta_x) = G^\pm(\nu, x).$$

It is straightforward to verify that in fact $\hat{j} = j$, since for any integrable $w \in B(\Gamma^2)$

$$\int_{\Gamma^2} w(\nu, \eta)\, \mathrm{d}j = n \int_{\Gamma\times\mathcal{T}} w(\nu, \nu + \tfrac{1}{n}\delta_x)\mathsf{J}_t^+(\mathrm{d}\nu, \mathrm{d}x) + n \int_{\Gamma\times\mathcal{T}} w(\nu, \nu - \tfrac{1}{n}\delta_x)\mathsf{J}_t^-(\mathrm{d}\nu, \mathrm{d}x),$$

where

$$\int_{\Gamma\times\mathcal{T}} w(\nu,\nu\pm\tfrac{1}{n}\delta_x)\mathsf{J}_t^{\pm}(\mathrm{d}\nu,\mathrm{d}x)=\int_{\Gamma\times\mathcal{T}} w(\nu,\nu\pm\tfrac{1}{n}\delta_x)\boldsymbol{G}^{+}(\nu,x)\Theta_{\mathsf{P}}^{n,\pm}(\mathrm{d}\nu,\mathrm{d}x),$$

and the fact that for any integrable $w'\in B(\Gamma^2)$

$$\begin{aligned}\int_{\Gamma^2} w'(\nu,\eta)\mathrm{d}\sqrt{(\mathsf{P}\kappa_n^{\pm})\mathsf{s}_{\#}(\mathsf{P}\kappa_n^{\dagger,\pm})}&=n\int_{\Gamma\times\mathcal{T}} w'(\nu,\nu\pm\tfrac{1}{n}\delta_x)\,\mathrm{d}\sqrt{(\mathsf{P}\chi_\nu^{\pm})(\mathsf{T}_{\#}^{n,\mp}(\mathsf{P}\bar{\chi}_\nu^{n,+}))}\\&=n\int_{\Gamma\times\mathcal{T}} w'(\nu,\nu\pm\tfrac{1}{n}\delta_x)\,\mathrm{d}\Theta_{\mathsf{P}}^{n,\pm}(\mathrm{d}\nu,\mathrm{d}x).\end{aligned}$$

Therefore we can write

$$\mathcal{E}\mathrm{nt}\left(j\middle|\sqrt{(\mathsf{P}\kappa_n)\mathsf{s}_{\#}(\mathsf{P}\kappa_n^{\dagger})}\right)=\mathcal{E}\mathrm{nt}\left(j^{+}\middle|\sqrt{(\mathsf{P}\kappa_n^{+})\mathsf{s}_{\#}(\mathsf{P}\kappa_n^{\dagger,+})}\right)+\mathcal{E}\mathrm{nt}\left(j^{-}\middle|\sqrt{(\mathsf{P}\kappa_n^{-})\mathsf{s}_{\#}(\mathsf{P}\kappa_n^{\dagger,-})}\right),$$

$$\begin{aligned}\mathcal{E}\mathrm{nt}\left(j^{\pm}\middle|\sqrt{(\mathsf{P}\kappa_n^{\pm})\mathsf{s}_{\#}(\mathsf{P}\kappa_n^{\dagger,\pm})}\right)&=\int_{\Gamma\times\mathcal{T}}\phi\left(\boldsymbol{G}^{\pm}(\nu,x)\right)n\Theta_{\mathsf{P}}^{n,\pm}(\mathrm{d}\nu,\mathrm{d}x)\\&=n\mathcal{E}\mathrm{nt}(\mathsf{J}^{\pm}|\Theta_{\mathsf{P}}^{n,\pm}),\end{aligned}\tag{17.36}$$

from which we conclude

$$\mathcal{E}\mathrm{nt}\left(j\middle|\sqrt{(\mathsf{P}\kappa_n)\mathsf{s}_{\#}(\mathsf{P}\kappa_n^{\dagger})}\right)=n\mathcal{R}_n(\mathsf{P},\mathsf{J}^{+},\mathsf{J}^{-}).$$

□

Finally, to obtain uniform estimates for the fluxes as $n\to\infty$, we provide the following bounds.

Lemma 17.19. *For any* $\mathsf{P}\in\mathcal{P}(\Gamma_n)$

$$\begin{aligned}\max\left(\int_{\Gamma\times\mathcal{T}}(1+\nu(\mathcal{T})^2)\,\mathrm{d}\vartheta_{\mathsf{P}}^{\pm},\int_{\Gamma\times\mathcal{T}}(1+\nu(\mathcal{T})^2)\,\mathrm{d}(\mathsf{T}_{\#}^{n,\mp}\bar{\vartheta}_{\mathsf{P}}^{n,\pm})\right)&\le M_n,\\ \max\left(\int_{\Gamma\times\mathcal{T}}\mathrm{d}\vartheta_{\mathsf{P}}^{\pm},\int_{\Gamma\times\mathcal{T}}\mathrm{d}\bar{\vartheta}_{\mathsf{P}}^{n,\pm}\right)\le M_n\int_{\Gamma}(1+\nu(\mathcal{T})^2)\,\mathrm{d}\mathsf{P}.&\end{aligned}\tag{17.37}$$

Moreover, for any $\mathsf{P}\in\mathcal{P}(\Gamma_n)$ *and admissible* $\mathsf{J}^{\pm}\in\mathcal{M}_{\mathrm{loc}}(\Gamma\times\mathcal{T})^{+}$

$$\tilde{\phi}\left(\frac{1}{M_n}\int_{\Gamma\times\mathcal{T}}(1+\nu(\mathcal{T})^2)^{-1}\,\mathsf{J}^{\pm}(\mathrm{d}\nu,\mathrm{d}x)\right)M_n\le\mathcal{R}_n(\mathsf{P},\mathsf{J}^{+},\mathsf{J}^{-})\tag{17.38}$$

and

$$\tilde{\phi}\left(\frac{1}{AM_n}\int_{\Gamma\times\mathcal{T}}\mathsf{J}^{\pm}(\mathrm{d}\nu,\mathrm{d}x)\right)AM_n\le\mathcal{R}_n(\mathsf{P},\mathsf{J}^{+},\mathsf{J}^{-}),\tag{17.39}$$

where

$$A=\int_{\Gamma}(1+\nu(\mathcal{T})^2)\,\mathrm{d}\mathsf{P}.$$

Remark 17.20. It can be shown, using the weighted total variation norm $\|\cdot\|_{TV,a}$ defined in (14.15), that

$$\|\mathsf{P}_s - \mathsf{P}_t\|_{TV,a} \le 4n \max\left\{1 + \tfrac{2}{n^2}, 2\right\} \int_s^t \int_{\Gamma\times\mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1} \mathrm{d}(\mathsf{J}_r^+ + \mathsf{J}_r^-)\,\mathrm{d}r. \quad (17.40)$$

However, note that this estimate blows up as $n \to \infty$. For the proof of EDP-convergence we instead use a weaker metric, the transportation-like metric W defined by (18.12), which does behave uniform-in-n for a sequence of curves with finite $\limsup_{n\to\infty} \mathcal{J}_n$. □

Proof. Note that we have the inequalities

$$\chi^{\pm}(\mathcal{T}) \le M(1 + \nu(\mathcal{T})^2), \quad \bar{\chi}^{n,\pm}(\mathcal{T}) \le M_n(1 + \nu(\mathcal{T})^2)$$

and for any fixed $x^* \in \mathcal{T}$

$$\bar{\chi}^{n,\pm}(\mathcal{T}) \le M_n \left(1 + (\mathsf{T}_{x^*}^{n,\mp}(\nu))(\mathcal{T})^2\right),$$

from which together we can obtain (17.37), and observe that the same bounds hold for $\Theta_{\mathsf{P}}^{n,\pm}$. The desired assertions now follow after applying Jensen's inequality. □

17.3 Solutions

Recall the notion of weak solutions to ($\mathsf{FKE_n}$) in the sense of Definition 17.10, i.e. curves $(\mathsf{P}_t)_{t\in[0,T]} \subset \mathcal{P}(\Gamma)$ such that $\operatorname{supp}\mathsf{P}_t \in \Gamma_n$ for all $t \in [0,T]$, P_t continuous in the narrow topology and for all $s,t \in [0,T]$, and all $F \in B_c(\Gamma)$,

$$\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} \left((\overline{\nabla}^{n,+} F)\,\mathrm{d}\vartheta_{\mathsf{P}_r}^+ + (\overline{\nabla}^{n,-} F)\,\mathrm{d}\vartheta_{\mathsf{P}_r}^-\right) \mathrm{d}r. \quad (17.41)$$

Moreover, $\mathcal{I}$-solutions are weak solutions for which we have the following inequality for bounded $F \in B_b(\Gamma)$ and $s,t \in [0,T]$,

$$\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s \le \int_s^t \int_{\Gamma\times\mathcal{T}} \left((\overline{\nabla}^{n,+} F)_+\,\mathrm{d}\vartheta_{\mathsf{P}_r}^+ + (\overline{\nabla}^{n,-} F)_+\,\mathrm{d}\vartheta_{\mathsf{P}_r}^-\right) \mathrm{d}r. \quad (17.42)$$

Remark 17.21. Recall that $\int (1+\nu(\mathcal{T})^2)^{-1}\,\mathrm{d}\vartheta_{\mathsf{P}_t}^{\pm} \le M_n$ independently of P_t. Hence it is easy to check that (P) is a $\mathcal{I}$-solution if and only if $(\mathsf{P}, \vartheta_{\mathsf{P}}^+, \vartheta_{\mathsf{P}}^-) \in \mathsf{CE}_n$. □

Remark 17.22. As noted in Remark 17.3, the restriction (17.42) is necessary to ensure propagation of moments for solutions. This is because it allows us to substitute the functions $F_m(\nu) = \min(\nu(\mathcal{T})^p, m)$ and use the fact that $\overline{\nabla}^{n,-} F_m \le 0$. Now, note that by an approximation argument the continuity equation holds with respect to bounded functions if we assume that (17.41) holds and

$$\int_0^T \vartheta^{\pm}_{\mathsf{P}_t}(\Gamma \times \mathcal{T})\,\mathrm{d}t < \infty,$$

in which case (17.42) follows.

However, without this assumption one cannot directly obtain (17.42) from (17.41). For example, if one used the approximation $F_k = \xi_k F$, than it is straightforward to verify that the bound

$$\limsup_k \int_{\Gamma\times\mathcal{T}} \int_0^T \left(|\overline{\nabla}^{n,+} \xi_k|\,\mathrm{d}\vartheta^+_{\mathsf{P}_r} + |\overline{\nabla}^{n,-} \xi_k|\,\mathrm{d}\vartheta^-_{\mathsf{P}_r} \right) \mathrm{d}r < \infty, \tag{17.43}$$

is sufficient, which by Lemma 17.14 is valid if P_t has uniformly bounded first moment.

It should be noted that if P_t corresponds to a explicit stochastic process the estimate likely holds, and moreover one might show even stronger propagation of moments and uniqueness directly in the sense that

$$\mathbb{E}\left[\sup_{t\in[0,T]} F(\nu^n_t(\mathcal{T}))\right] < \infty,$$

as done for example in [FM04] for the case of $\mathcal{T} = \mathbb{R}^d$ using a standard argument involving stopping times and Gronwall's inequality. □

Recall the constants $C_{n,p}$ of Lemma 17.14. We then have the following existence and propagation of moments result.

Proposition 17.23. *Suppose that* $\bar{\mathsf{P}}$ *is such that*

$$\int_\Gamma \nu(\mathcal{T})^p\,\bar{\mathsf{P}}(\mathrm{d}\nu) < \infty$$

for some $p \geq 1$. *Then for any* $\mathcal{I}$*-solution* P_t *with* $\mathsf{P}_0 = \bar{\mathsf{P}}$ *we have for every* $t \in [0,T]$

$$\int_\Gamma \nu(\mathcal{T})^p\,\mathsf{P}_t(\mathrm{d}\nu) \leq \left(\int_\Gamma \nu(\mathcal{T})^p\,\mathsf{P}_0(\mathrm{d}\nu) + tC_{n,p}\right) e^{tC_{n,p}}.$$

Suppose in addition that $\mathcal{E}\mathrm{nt}(\bar{\mathsf{P}}|\Pi_n) < \infty$ *and the moment estimate holds for* $p \geq 2$. *Then a* $\mathcal{I}$*-solution* P_t *exists,* $(\mathsf{P}, \vartheta^+_{\mathsf{P}}, \vartheta^-_{\mathsf{P}}) \in \mathsf{CE}^*_n$ *and the continuity equation holds with respect to bounded functions* $\mathcal{F} \in B_b(\Gamma)$.

Finally, suppose instead that the detailed balance condition holds and $\mathcal{E}\mathrm{nt}(\bar{\mathsf{P}}|\Pi_n) < \infty$. *Then a* $\mathcal{I}$*-solution* P_t *exists and*

$$\sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})\mathsf{P}_t(\mathrm{d}\nu) < \infty.$$

Remark 17.24. We will show in Section 19.1 that finite $\mathcal{F}_n(\mathsf{P})$ in fact implies that

$$\int_\Gamma \nu(\mathcal{T})\,\mathsf{P} < \infty,$$

which justifies a moment estimate on the initial data for the detailed balance case. □

Proof. The moment bounds for any $\mathcal{I}$-solution follow from subsequently Lemma 17.14, the inequality (17.42), and a Gronwall-type argument for the functions $F_m(\nu) := \min(\nu(\mathcal{T})^2, m)$.

Next, set $\mathsf{f} = (1+\nu(\mathcal{T})^p)$ for any $p \geq 2$. By Lemma 17.14 the function f satisfies Assumption 13.3 with the modification of Remark 13.8, and hence by Proposition 13.4 a curve P_t exists that is a solution in the sense of Definition 12.11, which implies that P_t is a $\mathcal{I}$-solution and $(\mathsf{P}, \vartheta^+_{\mathsf{P}}, \vartheta^-_{\mathsf{P}}) \in \mathsf{CE}^*_n$.

Finally, consider the case of detailed balance. Existence of a curve P_t satisfying (17.41) follows from Proposition 14.12. Moreover, following the proof of Propositions 13.4 and 14.12, it is clear that for the sequence of approximating regularized solutions P^ε_t we have the estimate

$$\int_\Gamma \nu(\mathcal{T})\,\mathsf{P}^\varepsilon_t(\mathrm{d}\nu) \leq \left(\int_\Gamma \nu(\mathcal{T})\,\mathsf{P}_0(\mathrm{d}\nu) + tC_{n,1}\right) e^{tC_{n,1}},$$

which survives in the limit $\varepsilon \to 0$, and hence (17.42) is valid as discussed in Remark 17.22, which implies P_t is a $\mathcal{I}$-solution. □

17.4 Variational characterization

We will now give a proof of the main result of this chapter, applying the framework of Chapters 13 and 14.

Proof of Theorem 17.11. Let us first consider the setting of Chapter 13. Take any $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_n$ with $\mathcal{F}_n(\mathsf{P}_0) < \infty$. Then the moment bounds of (17.16) and Condition 3 of the continuity equation imply that in fact

$$\sup_{t\in[0,T]} \int_{\Gamma^2} (1 + \nu(\mathcal{T})^2)\,\kappa^\pm_n(\nu, \mathrm{d}\eta)\mathsf{P}^n_t\kappa^\pm_n < \infty \tag{17.44}$$

$$\int_0^T \mathsf{J}^\pm_t(\Gamma \times \mathcal{T})\,\mathrm{d}t < \infty, \tag{17.45}$$

and by an approximation argument one can show that the continuity equation will hold with respect to bounded functions $\mathcal{F} \in B_b(\Gamma)$.

Now, by Lemmas 17.16 and 17.17 we can construct a curve (P, j) satisfying the continuity equation of Definition 13.1, and such that the corresponding EDP and rate functional are equal to $n\mathcal{J}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ and $n\mathcal{I}_n(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ respectively. Therefore, we can deduce from Theorem 13.2 that $\mathcal{J}_n = \mathcal{I}_n$ and its null-minimizers are the unique solutions satisfying the bound (17.16). If now in addition we have the bound

$$\int_\Gamma \nu(\mathcal{T})^2\, \bar{\mathsf{P}}(\mathrm{d}\nu) < \infty,$$

we can conclude by Proposition 17.23 that any $\mathcal{I}$-solution satisfies (17.16), and that these solutions exist.

Next, we consider the case of detailed balance, for which we employ the functions $a(\nu)$ and $\xi_k(\nu)$ with the properties listed in 17.14. Note that by Theorem 14.3 and Remark 14.5 we have the inequality $\mathcal{J}_n \geq \mathcal{I}_n$ for all curves $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$, and the converse statement for all curves such that

$$\limsup_{k\to\infty} \sup_{t\in[0,T]} \int_\Gamma |\overline{\nabla}\xi_k| \kappa_n(\nu, \mathrm{d}\eta) \mathsf{P}_t(\mathrm{d}\nu) < \infty,$$

which by Lemma 17.14 applies to curves with uniformly bounded first moment. Since any $\mathcal{I}$-solution satisfies this property if $\mathcal{F}_n(\mathsf{P}_0) < \infty$ by Proposition 17.23, and such solutions exist, it remains to show that the moment bounds apply for any curve $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$ with finite $\mathcal{I}_n$.

To do so, set $F_m(\nu) = \min(\nu(\mathcal{T}), m)$. Then by (17.15) and duality we have for every $t \in [0, T]$

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_0 \leq \int_0^t \Big(\mathcal{L}_n(\mathsf{P}_r, \mathsf{J}_r^+, \mathsf{J}_r^-) + \mathcal{H}_n(\mathsf{P}_r, (\overline{\nabla}^{n,+} F_m)_+, (\overline{\nabla}^{n,-} F_m)_+) \Big)\, \mathrm{d}r,$$

where, recall,

$$\mathcal{H}_n^*(\mathsf{P}, \omega^+, \omega^-) := \int_{\Gamma\times\mathcal{T}} (e^{\omega^+(\nu,x)} - 1)\, \chi_\nu^+(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu) + \int_{\Gamma\times\mathcal{T}} (e^{\omega^-} - 1)\, \mathrm{d}\chi_\nu^-(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu).$$

Note that $(\overline{\nabla}^{n,-} F_m) \leq 0$, and $e^{(\overline{\nabla}^{n,+} F_m)} - 1 \leq (e^1 - 1)$. Moreover, $\chi_\nu^+(\mathcal{T}) \leq M'\gamma(\mathcal{T})\nu(\mathcal{T})$, and therefore we estimate

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_0 \leq \int_0^t \left(\mathcal{L}_n(\mathsf{P}_r, \mathsf{J}_r^+, \mathsf{J}_r^-) + M'\gamma(\mathcal{T}) \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_r \right) \mathrm{d}r,$$

from which the desired bound follows after a Gronwall-type argument. □

Chapter 18

Liouville equation

In this chapter, we will consider the variational formulation for our proposed limit of the forward Kolmogorov equation $\mathsf{FKE_n}$, namely the *Liouville equation*

$$\partial_t \mathsf{P}_t + \mathrm{div}_\Gamma \left(\mathsf{P}_t \big(\chi^+ - \chi^- \big) \right) = 0. \tag{Li}$$

It can be interpreted as a transport equation lifted from the mean-field dynamics, in the sense that it describes the evolution of the law of a deterministic process satisfying the mean-field equation but with possibly random initial conditions. We will consider the same ingredients as in previous sections, namely a nonnegative EDP-functional consisting of an action term, a difference of free energies, and a corresponding Fisher information term. The main technical tool that we use is a new superposition principle, which allows us to prove the chain rule via the results on mean-field curves of Chapter 16.

Solutions to (Li) are defined as appropriate weak solutions to

$$\partial_t \mathsf{P}_t = Q_\infty^* \, \mathsf{P}_t,$$

where $\mathsf{P}_t \in \mathcal{P}(\Gamma)$ for all $t \in [0, T]$ and the operator Q_∞^* is the dual of Q_∞ given by

$$(Q_\infty F)(\nu) = \int_{\mathcal{T}} (\mathrm{grad}_\Gamma F)(\nu, x) V[\nu](\mathrm{d}x),$$
$$V[\nu] := \chi_\nu^+ - \chi_\nu^-,$$

for all $F \in \mathrm{Cyl}_c(\Gamma)$. Here $\mathrm{Cyl}_c(\Gamma)$ is the space of all compactly supported smooth cylinder functions, i.e. those of the form

$$F(\nu) = g\left(\langle 1, \nu \rangle, \langle f_1, \nu \rangle, \ldots, \langle f_m, \nu \rangle \right),$$

where $g \in C_c^\infty(\mathbb{R}^m)$ with $m \in \mathbb{N}$, and $f_1, \dots, f_m \in C_b(\mathcal{T})$, and grad_Γ is the gradient defined by

$$\mathrm{grad}_\Gamma F(\nu, x) = (\nabla g)\left(\langle 1, \nu\rangle, \langle f_1, \nu\rangle, \dots, \langle f_m, \nu\rangle\right) \cdot (1, f_1(x), \dots, f_m(x))^\top.$$

Moreover, let $\mathrm{Cyl}_b(\Gamma)$ be the space of bounded cylinder functions.

In light of Definition 17.10 we consider the following type of solutions, which will be natural minimizers of the rate functional.

Definition 18.1. A curve $(\mathsf{P}_t)_{t\in[0,T]}$ is a *weak solution* to (Li) if P_t is continuous in the narrow topology and for all $s, t \in [0, T]$, and all $F \in \mathrm{Cyl}_c(\Gamma)$,

$$\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} (\mathrm{grad}_\Gamma F)(\nu, x) V[\nu](\mathrm{d}x)\mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}r. \quad (18.1)$$

Moreover, it is a $\mathcal{I}$*-solution* if in addition for any bounded cylinder function $\mathcal{F} \in \mathrm{Cyl}_b(\Gamma)$

$$\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s \le \int_s^t \int_{\Gamma\times\mathcal{T}} \left((\mathrm{grad}_\Gamma F)_+\,\mathrm{d}\vartheta^+_{\mathsf{P}_r} - (\mathrm{grad}_\Gamma F)_-\,\mathrm{d}\vartheta^-_{\mathsf{P}_r}\right)\mathrm{d}r. \quad (18.2)$$

□

Remark 18.2. Similar as in Chapter 17 the (18.2) condition is necessary to propagate moment estimates in the detailed balance case, and it clear from an approximation argument as in Remark 17.22 that the inequality holds if P_t is a weak solution and has first mass-moment uniformly bounded in time. □

Remark 18.3. Note that (Li) is the transport equation associated to the measure-valued vector field $V[\nu]$. Now let the flow $G : [0, T]\times\Gamma \to \Gamma$ be the unique strong solution to the mean-field equation, i.e. with

$$\partial_t G_t[\nu] = V[G_t[\nu]]. \quad (18.3)$$

As will be shown in Section 18.2, $\mathsf{P}_t := (G_t)_\#\bar{\mathsf{P}}$ is a $\mathcal{I}$-solution to (Li) for any initial datum $\bar{\mathsf{P}} \in \mathcal{P}(\Gamma)$. In particular, if ν_t is a solution to (MF) than $\mathsf{P}_t := \delta_{\nu_t}$ is a $\mathcal{I}$-solution to (Li). □

We will now consider arbitrary curves satisfying

$$\partial_t \mathsf{P}_t + \mathrm{div}_\Gamma(\mathsf{J}^+_t - \mathsf{J}^-_t) = 0, \quad (\mathsf{CE}_\infty)$$

in the following appropriate distributional sense.

Definition 18.4 (Continuity equation)**.** A triple $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ satisfies the continuity equation CE_∞ if

1. the curve $[0,T] \ni t \mapsto \mathsf{P}_t \in \mathcal{P}(\Gamma)$ is narrowly continuous,
2. the Borel family $(\mathsf{J}_t^{\pm})_{t\in[0,T]} \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma \times \mathcal{T})$ satisfies

$$\int_0^T \int_{\Gamma\times\mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1}\, \mathrm{d}\mathsf{J}_t^{\pm}\, \mathrm{d}t < \infty,$$

3. for every $s,t \in [0,T]$ and all $F \in \mathrm{Cyl}_c(\Gamma)$

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma\times\mathcal{T}} \mathrm{grad}_\Gamma F\, (\mathrm{d}\mathsf{J}_r^+ - \mathrm{d}\mathsf{J}_r^-)\, \mathrm{d}r.$$

4. for every $s,t \in [0,T]$ and all $F \in \mathrm{Cyl}_b(\Gamma)$

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_s \le \int_s^t \int_{\Gamma\times\mathcal{T}} \big((\mathrm{grad}_\Gamma F)_+\, \mathrm{d}\mathsf{J}_r^+ - (\mathrm{grad}_\Gamma F)_-\, \mathrm{d}\mathsf{J}_r^-\big)\, \mathrm{d}r. \tag{18.4}$$

Moreover, a triple $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ satisfies the continuity equation CE^*_∞, if in addition

$$\sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^2\, \mathsf{P}_t(\mathrm{d}\nu) < \infty, \qquad \int_0^T \int_{\Gamma\times\mathcal{T}} \|\mathsf{J}_t^{\pm}\|_{TV}\, \mathrm{d}t < \infty. \tag{18.5}$$

□

Similar as in Chapter 17 the continuity equation CE^*_∞ corresponds to the case of bounded fluxes and uniform moment estimates, and the weaker formulation of CE_∞ is used for the case of mean-field detailed balance.

Remark 18.5. The inequality (18.4) can be used to show propagation of first moments under finite $\mathcal{I}$. However, unfortunately, this is not enough to show equivalence of the rate functional and the EDP-functional. We will see in Section 18.3 that (18.4) is valid if a certain superposition principle holds. □

Let us introduce the EDP-functional. Recall from Chapter 17 the notation $\vartheta^{\pm}_{\mathsf{P}}(\mathrm{d}\nu, \mathrm{d}x) := \chi^{\pm}_\nu(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu)$, and let us set

$$\bar\vartheta^{\pm}_{\mathsf{P}}(\mathrm{d}\nu\, \mathrm{d}x) := \bar\chi^{\pm}_\nu(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu). \tag{18.6}$$

Definition 18.6. Let $\Theta^{\infty,\pm}_{\mathsf{P}} \in \mathcal{M}_{\mathrm{loc}}(\Gamma \times \mathcal{T})$ be the geometric averages of $\vartheta^{\pm}_{\mathsf{P}}$ and $\bar\vartheta^{\pm}_{\mathsf{P}}$, i.e.

$$\Theta^{\infty,\pm}_{\mathsf{P}}(\mathrm{d}\nu, \mathrm{d}x) := \sqrt{\frac{\mathrm{d}\vartheta^{\pm}_{\mathsf{P}}}{\mathrm{d}\Sigma}\frac{\mathrm{d}\bar\vartheta^{\mp}_{\mathsf{P}}}{\mathrm{d}\Sigma}}\, \mathrm{d}\Sigma,$$

for any dominating measure Σ. We define the following objects:

- The dissipation potential $\mathcal{R}_\infty : \mathcal{P}(\Gamma) \times \mathcal{M}^+_{\text{loc}}(\Gamma \times \mathcal{T})^2 \to [0, +\infty]$,

$$\mathcal{R}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \mathcal{E}\text{nt}(\mathsf{J}^+|\Theta^{\infty,+}_{\mathsf{P}}) + \mathcal{E}\text{nt}(\mathsf{J}^-|\Theta^{\infty,-}_{\mathsf{P}}).$$

and the dual dissipation potential $\mathcal{R}^*_\infty : \mathcal{P}(\Gamma) \times \mathcal{B}_c(\Gamma \times \mathcal{T})^2 \to \mathbb{R}$,

$$\mathcal{R}^*_\infty(\mathsf{P}, \omega^+, \omega^-) := \int_{\Gamma\times\mathcal{T}} (e^{\omega^+} - 1)\,\mathrm{d}\Theta^{\infty,+}_{\mathsf{P}} + \int_{\Gamma\times\mathcal{T}} (e^{\omega^-} - 1)\,\mathrm{d}\Theta^{\infty,-}_{\mathsf{P}}.$$

- The free energy $\mathcal{F}_\infty : \mathcal{P}(\Gamma) \to [0, +\infty]$,

$$\mathcal{F}_\infty(\mathsf{P}) := \int_\Gamma \mathcal{F}_{MF}(\nu)\,\mathsf{P}(\mathrm{d}\nu).$$

- The Fisher information $\mathcal{D}_\infty : \mathcal{P}(\Gamma) \to [0, +\infty]$,

$$\mathcal{D}_\infty(\mathsf{P}) := \int_\Gamma \mathcal{D}_{MF}(\nu)\,\mathsf{P}(\mathrm{d}\nu).$$

- The EDP-functional $\mathcal{J}_\infty$ on CE_∞ for all curves with $\mathcal{F}_\infty(\mathsf{P}_0) < \infty$,

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t + \mathcal{F}_\infty(\mathsf{P}_T) - \mathcal{F}_\infty(\mathsf{P}_0) + \int_0^T \mathcal{D}_\infty(\mathsf{P}_t)\,\mathrm{d}t.$$

- The Lagrangian $\mathcal{L}_\infty : \mathcal{P}(\Gamma) \times \mathcal{M}^+_{\text{loc}}(\Gamma \times \mathcal{T})^2 \to [0, +\infty]$,

$$\mathcal{L}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \mathcal{E}\text{nt}(\mathsf{J}^+|\vartheta^+_{\mathsf{P}}) + \mathcal{E}\text{nt}(\mathsf{J}^-|\vartheta^+_{\mathsf{P}}).$$

and its dual, the Hamiltonian $\mathcal{H}_\infty : \mathcal{P}(\Gamma) \times \mathcal{B}_c(\Gamma \times \mathcal{T})^2 \to \mathbb{R}$,

$$\mathcal{H}_\infty(\mathsf{P}, \omega^+, \omega^-) := \int_{\Gamma\times\mathcal{T}} (e^{\omega^+} - 1)\,\mathrm{d}\vartheta^+_{\mathsf{P}} + \int_{\Gamma\times\mathcal{T}} (e^{\omega^-} - 1)\,\mathrm{d}\vartheta^-_{\mathsf{P}}.$$

- The rate functional $\mathcal{I}_\infty : \mathsf{CE}_\infty \to [0, +\infty]$,

$$\mathcal{I}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) := \int_0^T \mathcal{L}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t.$$

□

Remark 18.7. Recall from Chapter 16 that $\mathcal{F}_{MF}(\nu) := \frac{1}{2}\mathcal{E}\text{nt}(\nu|\gamma)$ and

$$\mathcal{D}_{MF}(\nu) := H^2(\chi^+_\nu, \bar{\chi}^+_\nu) + H^2(\chi^-_\nu, \bar{\chi}^-_\nu) + \frac{1}{2}\int_{\mathcal{T}} \mathrm{d}\left(\chi^+_\nu + \chi^-_\nu - \bar{\chi}^+_\nu - \bar{\chi}^-_\nu\right).$$

It is clear that this allows us to write

$$\mathcal{D}(\mathsf{P}) = H^2(\vartheta^+_{\mathsf{P}}, \bar{\vartheta}^+_{\mathsf{P}}) + H^2(\vartheta^-_{\mathsf{P}}, \bar{\vartheta}^-_{\mathsf{P}}) + \frac{1}{2}\int_{\Gamma\times\mathcal{T}} \mathrm{d}\left(\vartheta^+_{\mathsf{P}} + \vartheta^-_{\mathsf{P}} - \bar{\vartheta}^+_{\mathsf{P}} - \bar{\vartheta}^-_{\mathsf{P}}\right).$$

□

Remark 18.8. Note that $\Theta^{\infty,\pm}_{\mathsf{P}}(\mathrm{d}\nu,\mathrm{d}x) = \mathsf{P}(\mathrm{d}\nu)\theta^{\pm}_{\nu}(\mathrm{d}x)$. Moreover, if $\mathcal{E}\mathrm{nt}(\mathsf{J}^{\pm}_t|\Theta^{\pm,\infty}_{\mathsf{P}_t})$ is finite, we can set

$$\lambda^{\pm}_t[\nu](\mathrm{d}x) := \frac{\mathrm{d}\mathsf{J}^{\pm}_t}{\mathrm{d}\Theta^{\infty,\pm}_{\mathsf{P}_t}}(\nu,x)\,\theta^{\pm}_{\nu}(\mathrm{d}x),$$

and it is straightforward to verify that we have the disintegration

$$\mathsf{J}^{\pm}(\mathrm{d}\nu,\mathrm{d}x) = \lambda^{\pm}_t[\nu](\mathrm{d}x)\mathsf{P}_t(\mathrm{d}\nu),$$

and the equivalence

$$\mathcal{E}\mathrm{nt}(\mathsf{J}^{\pm}_t|\Theta^{\infty}_{\mathsf{P}_t}) = \int_{\Gamma}\mathcal{E}\mathrm{nt}(\lambda^{\pm}_t[\nu]|\theta^{\pm}_{\nu})\,\mathrm{d}\mathsf{P}_t. \tag{18.7}$$

Together with the definitions of $\mathcal{F}_\infty$ and $\mathcal{D}_\infty$ this implies that if $\mathcal{J}_\infty(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)$ is finite then the $\lambda^{\pm}_t[\nu]$ are well-defined for a.e. $t\in[0,T]$, $\mathcal{J}_\infty$ is equal to

$$\int_0^T\int_{\Gamma}\left(\mathcal{R}_{MF}(\nu_t,\lambda^{+}_t[\nu],\lambda^{-}_t[\nu]) + \mathcal{D}_{MF}(\nu)\right)\mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}t + \int_{\Gamma}\mathcal{F}_{MF}(\nu)\mathsf{P}_T(\mathrm{d}\nu) - \int_{\Gamma}\mathcal{F}_{MF}(\nu)\mathsf{P}_0(\mathrm{d}\nu),$$

and

$$\mathcal{I}_\infty(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) = \int_0^T\int_{\Gamma}\mathcal{L}_{MF}(\nu_t,\lambda^{+}_t[\nu],\lambda^{-}_t[\nu])\mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}t$$

Throughout the rest of this chapter we will simply write $\lambda^{\pm}_{t,\nu} := \lambda^{\pm}_t[\nu]$, or λ_ν for the curve $(\lambda_{t,\nu})_{t\in[0,T]}$. □

We will show the following two equivalences, which subsumes Theorem (15.6).

Theorem 18.9. *For any $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\in\mathsf{CE}^*_\infty$ the superposition principle holds, i.e. there exists a Borel probability measure Q over $C([0,T];\Gamma)$ such that*

1. *for the time-evaluations e_t we have $(e_t)_\#Q = \mathsf{P}_t$ for all $t\in[0,T]$,*
2. *the measure Q is concentrated on the family of curves $\nu\in AC([0,T];(\Gamma,\|\cdot\|_{TV}))$ such that $(\nu,\lambda^+_\nu,\lambda^-_\nu)\in\mathscr{C}\mathscr{E}$, where $\lambda^{\pm}_\nu$ is defined via the disintegration*

$$\mathsf{J}^{\pm}_t(\mathrm{d}\nu,\mathrm{d}x) = \lambda^{\pm}_{t,\nu}(\mathrm{d}x)\mathsf{P}_t(\mathrm{d}\nu) \qquad \textit{for a.e. } t\in[0,T],$$

Moreover, if for such a curve $\mathcal{F}_\infty(\mathsf{P}_0)<\infty$, we have

$$\mathcal{J}_\infty(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) = \mathcal{I}_\infty(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) = \int\mathcal{J}_{MF}\left(\nu,\lambda^+_\nu,\lambda^-_\nu\right)\mathrm{d}Q. \tag{18.8}$$

In particular, $\mathcal{J}_\infty \geq 0$. Moreover, if P_0 is such that $\mathcal{F}_n(\mathsf{P}_0) < \infty$ and

$$\int_\Gamma \nu(\mathcal{T})^2 \, \mathrm{d}\mathsf{P}_0 < \infty, \tag{18.9}$$

then

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = 0 \implies \begin{cases} \mathsf{P}_t \text{ is the unique } \mathcal{I}\text{-solution to } (\mathsf{Li}), \\ \mathsf{J}_t^\pm = \vartheta^\pm_{\mathsf{P}_t} \quad \text{for a.e. } t \in [0,T]. \end{cases} \tag{18.10}$$

Theorem 18.10. *Suppose that the mean-field detailed balance condition $b^+_\nu = b^-_\nu$ holds. Then for any $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$ we have that $\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) < \infty$ if and only if the superposition principle holds and*

$$\int \mathcal{J}_{MF}\left(\nu, \lambda^+_\nu, \lambda^-_\nu\right) \mathrm{d}Q < \infty,$$

in which case (18.8) *is valid.*

In particular, $\mathcal{J}_\infty \geq 0$. Moreover, if P_0 is such that $\mathcal{F}_n(\mathsf{P}_0) < \infty$ and then

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = 0 \implies \begin{cases} \mathsf{P}_t \text{ is the unique } \mathcal{I}\text{-solution to } (\mathsf{Li}), \\ \mathsf{J}_t^\pm = \vartheta^\pm_{\mathsf{P}_t} \quad \text{for a.e. } t \in [0,T]. \end{cases} \tag{18.11}$$

Remark 18.11. Note that we do not have the full equivalence (18.8) for all curves $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$, but *only* if the EDP-functional $\mathcal{J}_\infty$ is finite. This is because at the moment we do not have a superposition principle under finite $\mathcal{I}_\infty$, and whether this actually is the case is a interesting subject for future research. □

In both cases, the uniqueness of $\mathcal{I}$-solutions P_t under suitable moment conditions for the initial data is a consequence of the fact that they propagate moments. This allows the superposition to be applied, in which case it can be shown that $\mathsf{P}_t = (G_t)_\# \mathsf{P}_0$, where $G_t : \Gamma \to \Gamma$ maps $\bar{\nu}$ to the unique mean-field solution ν_t at time t, see Remark 18.3. It is determined by

$$\partial_t G_t[\nu] = V[G_t[\nu]].$$

Moreover, we have the following consequence.

Corollary 18.12. *Suppose $\mathsf{P}_0 = \delta_{\nu_0}$ with $\mathcal{F}_{MF}(\nu_0) < \infty$. Then*

$$\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = 0 \iff \begin{cases} \mathsf{P}_t = \delta_{\nu_t}, \quad \nu_t \text{ is the unique strong solution to } (\mathsf{MF}) \\ \mathsf{J}_t^\pm = \vartheta^\pm_{\mathsf{P}} \quad \text{for a.e. } t \in [0,T] \end{cases}$$

18.1 A priori estimates

Due to the representation (18.7) of the dissipation potential in terms of mean-field objects, we can directly derive the following estimates from Lemma 16.9.

Corollary 18.13. *Let* $\mathsf{P} \in \mathcal{P}(\Gamma)$, $\mathsf{J}^\pm \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma \times \mathcal{T})$ *be such that* $\mathcal{R}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) < \infty$, *and set*

$$\lambda^{\mathrm{net}}_\nu := \lambda^+_\nu - \lambda^-_\nu.$$

Then the following estimates hold:

$$\int_\Gamma M\tilde{\phi}\left(\frac{\lambda^\pm_\nu(\mathcal{T})}{M(1+\nu(\mathcal{T})^2)}\right)\mathsf{P}(\mathrm{d}\nu) \le \min\left(\mathcal{R}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-), \mathcal{L}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)\right)$$

and, if the mean-field detailed balance condition $b^+_\nu = b^-_\nu$ *applies,*

$$\int_\Gamma M\Psi\left(\frac{\|\lambda^{\mathrm{net}}_\nu\|_{TV}}{M(1+\nu(\mathcal{T}))}\right)\mathsf{P}(\mathrm{d}\nu) \le \mathcal{R}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).$$

Finally, we consider the time-regularity for arbitrary curves, with respect a suitable metric.

Definition 18.14. We define the following metric:

$$W(\mathsf{P}^1, \mathsf{P}^2) := \sup_{F \in \mathbb{F}}\left\{\int_\Gamma F\,\mathrm{d}(\mathsf{P}^1 - \mathsf{P}^2)\right\}, \qquad \mathsf{P}^1, \mathsf{P}^2 \in \mathcal{P}(\Gamma), \tag{18.12}$$

where

$$\mathbb{F} := \left\{F \in \mathrm{Cyl}_c(\Gamma) \;:\; \left\|(1+\nu(\mathcal{T})^2)\mathrm{grad}_\Gamma F\right\|_\infty \le 1\right\}.$$

□

Note that W is narrowly lower semicontinuous. Moreover, for any $F \in \mathrm{Cyl}_c(\Gamma)$ automatically $\|(1+\nu(\mathcal{T})^2)\mathrm{grad}_\Gamma F\|_\infty < \infty$, and hence by a density argument it is straightforward to verify that convergence in W implies vague convergence on Γ, and therefore narrow convergence on narrowly pre-compact subsets.

Remark 18.15. Formally, one can represent W as a transport distance, in the sense that

$$W(\mathsf{P}^1, \mathsf{P}^2) = W_{d_\Gamma}(\mathsf{P}^1, \mathsf{P}^2),$$

where W_{d_Γ} is the 1-Wasserstein metric on $\mathcal{P}(\Gamma)$ induced by the metric d_Γ over Γ given by

$$d_\Gamma(\nu^1, \nu^2) := \inf_{(\nu_t)_{t\in[0,1]}}\left\{\int_0^1 \frac{|\dot{\nu}_t|_{TV}}{1+\nu_t(\mathcal{T})^2}\,\mathrm{d}t \;:\; \nu_0 = \nu^0,\ \nu_1 = \nu^2\right\}.$$

However, we do not require such representations in this current work. □

Lemma 18.16. *For any* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$ *we have*

$$W(\mathsf{P}_s, \mathsf{P}_t) \leq 2 \int_s^t \int_{\Gamma \times \mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1} \mathrm{d}(\mathsf{J}_r^+ + \mathsf{J}_r^-)\, \mathrm{d}r, \qquad \textit{for all } s, t \in [0, T].$$

Proof. This follows directly from the continuity equation, since for any $F \in \mathbb{F}$, $s, t \in [0, T]$:

$$\begin{aligned}\left| \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_s \right| &\leq \int_s^t \int_{\Gamma \times \mathcal{T}} \left| (1 + \nu(\mathcal{T})^2) \mathrm{grad}_\Gamma F \right| (1 + \nu(\mathcal{T})^2)^{-1}\, \mathrm{d}(\mathsf{J}_r^+ + \mathrm{d}\mathsf{J}_r^-)\, \mathrm{d}r \\ &\leq \int_s^t \int_{\Gamma \times \mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1}\, \mathrm{d}(\mathsf{J}_r^+ + \mathrm{d}\mathsf{J}_r^-)\, \mathrm{d}r.\end{aligned}$$

Taking the supremum over all $F \in \mathbb{F}$ we obtain the desired statement. □

18.2 Solutions

Here we briefly consider existence and representations for solutions to the Liouville equation. Recall from Definition 18.1 that weak solutions to (Li) are curves P_t that weak solutions are curves such that for $s, t \in [0, T]$ and $F \in \mathrm{Cyl}_c$

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_s = \int_s^t \int_{\Gamma \times \mathcal{T}} (\mathrm{grad}_\Gamma F)(\nu, x) V[\nu](\mathrm{d}x) \mathsf{P}_t(\mathrm{d}\nu)\, \mathrm{d}r. \tag{18.13}$$

and for $\mathcal{I}$-solutions we have in addition for every $F \in \mathrm{Cyl}_b(\Gamma)$

$$\int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\, \mathrm{d}\mathsf{P}_s \leq \int_s^t \int_{\Gamma \times \mathcal{T}} \left((\mathrm{grad}_\Gamma F)_+\, \mathrm{d}\vartheta^+_{\mathsf{P}_r} - (\mathrm{grad}_\Gamma F)_-\, \mathrm{d}\vartheta^-_{\mathsf{P}_r} \right) \mathrm{d}r. \tag{18.14}$$

Proposition 18.17. *For any* $\bar{\mathsf{P}}_t \in \mathcal{P}(\Gamma)$ *there exists a* $\mathcal{I}$*-solution* P_t *to* (Li) *with initial datum* $\bar{\mathsf{P}}$.

Moreover, for any $\mathcal{I}$*-solution* P_t *we have propagation of moments, i.e. for any* $p \geq 1$ *there exists a constant* C_p *such that*

$$\int_\Gamma \nu(\mathcal{T})^p \mathsf{P}_t(\mathrm{d}\nu) \leq e^{tC_p} \int_\Gamma \nu(\mathcal{T})^p\, \mathsf{P}_0(\mathrm{d}\nu)$$

Proof. Recall the flow $G : [0, T] \times \Gamma \to \Gamma$ determined by

$$\partial_t G_t[\nu] = V[G_t[\nu]],$$

Set $\mathsf{P}_t := (G_t)_\# \bar{\mathsf{P}}$. We will show that P_t is weak solution in the sense of (18.1). Namely, consider any $F \in \mathrm{Cyl}_c(\Gamma)$. Due the strong regularity of solutions to the mean-field equation it is straightforward to show that for all $s, t \in [0, T]$ we have the chain rule

$$F \circ G_t(\nu) - F(\nu) = \int_s^t \int_{\mathcal{T}} (\mathrm{grad}_\Gamma F)(G_r \circ \nu, x)\, \mathrm{d}V[G_r \circ \nu]\, \mathrm{d}r, \tag{18.15}$$

and hence

$$\begin{aligned}\int_\Gamma F \mathrm{d}\mathsf{P}_t - \int_\Gamma F \mathrm{d}\mathsf{P}_s &= \int_\Gamma \left(\int_s^t \int_{\mathcal{T}} (\mathrm{grad}_\Gamma F)(G_r \circ \nu, x)\, V[G_r[\nu]](\mathrm{d}x)\, \mathrm{d}r \right) \bar{\mathsf{P}}(\mathrm{d}\nu) \\ &= \int_s^t \int_{\Gamma \times \mathcal{T}} (\mathrm{grad}_\Gamma F)(\nu, x) V[\nu](\mathrm{d}x) \mathsf{P}_r(\mathrm{d}\nu)\, \mathrm{d}r,\end{aligned}$$

and thus P_t is indeed a weak solution. Moreover, to obtain (18.14), note that the chain rule holds (18.15) for any $F \in \mathrm{Cyl}_b(\Gamma)$, and therefore for such F

$$\begin{aligned}\int_\Gamma F \mathrm{d}\mathsf{P}_t - \int_\Gamma F \mathrm{d}\mathsf{P}_s &\le \int_\Gamma \int_s^t \left(\int_{\mathcal{T}} (\mathrm{grad}_\Gamma F)(G_r \circ \nu, x)\, V[G_r[\nu]](\mathrm{d}x) \right)_+ \mathrm{d}r \bar{\mathsf{P}}(\mathrm{d}\nu) \\ &\le \int_s^t \int_{\Gamma \times \mathcal{T}} \left((\mathrm{grad}_\Gamma F)_+ \,\mathrm{d}\vartheta^+_{\mathsf{P}_r} - (\mathrm{grad}_\Gamma F)_- \,\mathrm{d}\vartheta^-_{\mathsf{P}_r} \right) \mathrm{d}r.\end{aligned}$$

Finally, to prove propagation of moments, consider for every m the function $f_m(z) = \min(z, m)$ and a sequence of uniformly bounded, smooth, and monotone approximations $f_{m,\varepsilon}(z)$ such that as $\varepsilon \to 0$ we have $f_{m,\varepsilon}(z) \to f_m(z)$ for all $z \ge 0$ and $f'_{m,\varepsilon}(z) \uparrow f'_m(z)$ for all points of differentiability. Note that for any $t \in [0, T]$,

$$\begin{aligned}\int_\Gamma F_{m,\varepsilon} \mathrm{d}\mathsf{P}_t - \int_\Gamma F_{m,\varepsilon} \mathrm{d}\mathsf{P}_s &\le \int_s^t \int_{\Gamma \times \mathcal{T}} f'_{m,\varepsilon}(\nu(\mathcal{T}))\, \mathrm{d}\vartheta^+_{\mathsf{P}_r}\, \mathrm{d}r \\ &\le M' \gamma(\mathcal{T}) \int_s^t \int_{\Gamma \times \mathcal{T}} f'_{m,\varepsilon}(\nu(\mathcal{T})) \nu(\mathcal{T})\, \mathrm{d}\mathsf{P}_r\, \mathrm{d}r \\ &\le C_p \int_s^t \int_{\Gamma \times \mathcal{T}} f_m(\nu(\mathcal{T}))\, \mathrm{d}\mathsf{P}_r\, \mathrm{d}r,\end{aligned}$$

for a suitable constant C_p. Letting $\varepsilon \to 0$ we can now apply Gronwall's inequality, similar as in Section 17.3. □

18.3 Superposition principle

One of our main tools in proving the chain rule, uniqueness of solutions, and the variational representation of Theorem 18.9 is the superposition principle. It guarantees that we can represent the action as an expectation of the mean-field action under some measure over curves in $\mathscr{C}\mathscr{E}$, and allows us to use the theory on mean-field dynamics of Chapter 16. In this section, we will make this notion precise.

Theorem 18.18. *Let either:* $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\in\mathsf{CE}^*_\infty$*; or* $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)=(\mathsf{P},\vartheta^+_{\mathsf{P}},\vartheta^-_{\mathsf{P}})$ *with* P_t *a weak solution with uniformly bounded first moments; or* $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\in\mathsf{CE}_\infty$ *be such that the mean-field detailed balance condition holds and*

$$\int_0^T \mathcal{R}_\infty(\mathsf{P}_t,\mathsf{J}^+_t,\mathsf{J}^-_t)\,\mathrm{d}t<\infty.$$

Then there exists a Borel probability measure $Q\in\mathcal{P}(C([0,T];\Gamma))$ *satisfying* $(e_t)_\# Q=\mathsf{P}_t$ *for all* $t\in[0,T]$*, and concentrated on curves* $\nu\in AC([0,T];(\Gamma,\|\cdot\|_{TV}))$*, for which* $(\nu,\lambda^+_\nu,\lambda^-_\nu)\in\mathscr{CE}$*. Moreover,*

$$\int_0^T \mathcal{R}_\infty(\mathsf{P}_t,\mathsf{J}^+_t,\mathsf{J}^-_t)\,\mathrm{d}t=\int_{C([0,T];\Gamma)}\left(\int_0^T \mathcal{R}_{MF}\left(\nu_t,\lambda^+_{t,\nu_t},\lambda^-_{t,\nu_t}\right)\,\mathrm{d}t\right)Q(\mathrm{d}\nu). \tag{18.16}$$

with a similar statement for $\mathcal{I}_{MF}$*.*

Conversely, suppose there exists a Borel probability measure $Q\in\mathcal{P}(C([0,T];\Gamma))$ *concentrated on curves* $\nu\in AC([0,T];(\Gamma,\|\cdot\|_{TV}))$ *and a Borel family* $\{\lambda^\pm_{t,\nu}\}$*, for which* $(\nu,\lambda^+_\nu,\lambda^-_\nu)\in\mathscr{CE}$*.*

If either

$$\int_{C([0,T];\Gamma)}\left(\int_0^T \mathcal{R}_{MF}\left(\nu_t,\lambda^+_{t,\nu_t},\lambda^-_{t,\nu_t}\right)\,\mathrm{d}t\right)Q(\mathrm{d}\nu)<\infty,$$

or the same estimate for $\mathcal{I}_{MF}$ *holds, then* $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\in\mathsf{CE}_\infty$ *for* $\mathsf{P}_t:=(e_t)_\# Q$*,* $\mathsf{J}^\pm_t:=\mathsf{P}_t\lambda^\pm_{t,\nu}$*, and* (18.16) *holds as well.*

If instead

$$\int_{C([0,T];\Gamma)}\left(\int_0^T \|\lambda^\pm_{t,\nu_t}\|_{TV}\,\mathrm{d}t\right)Q(\mathrm{d}\nu)<\infty,\qquad \sup_{t\in[0,T]}\int_\Gamma \nu(\mathcal{T})^2\mathsf{P}_t(\mathrm{d}\nu)<\infty,$$

then $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\in\mathsf{CE}^*_\infty$*, and* (18.16) *holds as well.*

Remark 18.19. The inequality 18.4 is not used in obtaining a superposition principle in the case of detailed balance and bounded action $\int\mathcal{R}_\infty$, and in fact follows a posteriori from the superposition itself. □

Remark 18.20. As noted in Remark 18.11, it is clear that there is one scenario missing: the case of mean-field detailed balance and curves $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\in\mathsf{CE}_\infty$ such that $\mathcal{I}_n(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)<\infty$. This is because in that setting we cannot get strong enough bounds to prove the superposition principle, even though *a posteriori* the necessary bounds hold if the superposition applies and $\mathcal{F}_\infty(\mathsf{P})<\infty$, by the equivalence $\mathcal{J}_{MF}=\mathcal{I}_{MF}$. □

The inspiration for using a superposition principle stems from similar approaches in [EFLS16], [Erb16], where it is applied to transport equations lifted from the Boltzmann-equation or mean-field jump dynamics respectively, and the main ingredient is the abstract superposition principle over $\mathbb{R}^{\mathbb{N}}$ of [AT14]. However, these results are not directly applicable to our setting, since the mass of $\nu_t(\mathcal{T})$ for a mean-field curve is not fixed, and $V[\nu](\mathcal{T})$ is finite but unbounded over Γ unless we assume $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_\infty$. We remedy this by combining two known superposition principles: on the one hand, the abstract superposition principle over $\mathbb{R}^{\mathbb{N}}$ of [AT14] and, on the other hand, one for finite-dimensional vector fields with linear growth, found in [AC08]. Our result is stated in Theorem E.1.

Proof. Consider any $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_\infty$ and for a.e. $t \in [0, T]$ set $\lambda^{\mathrm{net}}_{t,\nu} := \lambda^+_{t,\nu} - \lambda^-_{t,\nu}$. Then

$$\int_0^T \int_\Gamma \|\lambda^{\mathrm{net}}_{t,\nu}\|_{TV} \mathsf{P}_t(\mathrm{d}\nu)\, \mathrm{d}t < \infty.$$

Similarly, if P_t a weak solution with uniformly bounded first moments then $\lambda^{\mathrm{net}}_{t,\nu} = \chi^+_\nu - \chi^-_\nu$, we have for any $t \in [0, T]$

$$\frac{\|\chi^+_\nu - \chi^-_\nu\|_{TV}}{1 + \nu(\mathcal{T})} \le 2M \frac{1 + \nu(\mathcal{T})^2}{1 + \nu(\mathcal{T})} \le 2M(1 + \nu(\mathcal{T}))$$

and therefore

$$\int_0^T \int_\Gamma \frac{\|\lambda^{\mathrm{net}}_{t,\nu}\|_{TV}}{1 + \nu(\mathcal{T})} \mathsf{P}_t(\mathrm{d}\nu)\, \mathrm{d}t < \infty. \tag{18.17}$$

Finally, suppose that the mean-field detailed balance condition holds and consider a $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$. By Corollary 18.13,

$$\int_\Gamma M\Psi\left(\frac{\|\lambda^{\mathrm{net}}_{t,\nu}\|_{TV}}{M(1 + \nu(\mathcal{T}))}\right) \mathsf{P}_t(\mathrm{d}\nu) \le \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, J^-_t),$$

which by Jensen's inequality again implies (18.17). In all three cases the inequality (18.17) is satisfied, which allows us to use Theorem E.1.

To do so, take a countable and dense set $f_1, f_2, \ldots \in C_b(\mathcal{T})$, with $f_1 = 1$, $\|f_i\|_\infty \le 1$, $i \ge 2$, and define $\mathbb{T} : \Gamma \to \mathbb{R}^{\mathbb{N}}$

$$\mathbb{T}(\nu) := \left(\int_{\mathcal{T}} f_1 \,\mathrm{d}\nu, \int_{\mathcal{T}} f_2 \,\mathrm{d}\nu \ldots\right).$$

Note that $\mathbb{T}(\nu)$ is injective, continuous when Γ is equipped with the narrow topology and $\mathbb{R}^{\mathbb{N}}$ with product topology, and is an isometry between $(\Gamma, \|\cdot\|_{TV})$ and $(\mathbb{T}(\Gamma), |\cdot$

$|_\infty)$, where $|\cdot|_\infty$ is the uniform norm over $\mathbb{R}^{\mathbb{N}}$. We set $\sigma_t := \mathbb{T}_\# \mathsf{P}_t \in \mathcal{P}(\mathbb{R}^{\mathbb{N}})$, and for a.e. $t \in [0,T]$ define the vector field $\mathbf{W}_t : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ via its components

$$W_i(t,z) := \int_X f_i(x)\, \lambda^{\text{net}}_{t,\mathbb{T}^{-1}(z)}(\mathrm{d}x).$$

Note that the support of $\mathbf{W}_t$ is in $\mathbb{T}(\Gamma)$, that $|\mathbf{W}_t(z)|_\infty \le \|\lambda^{\text{net}}_{t,\mathbb{T}^{-1}(z)}\|_{TV}$ and $(\mathbb{T}(\nu))_1 = \nu(\mathcal{T})$. Therefore, by (18.17) we have the estimate

$$\int_0^T \int_{\mathbb{R}^{\mathbb{N}}} \frac{|\mathbf{W}_t(z)|_\infty}{1+|z_1|} \sigma(\mathrm{d}z)\,\mathrm{d}t < \infty.$$

Moreover, $(\sigma, \mathbf{W})$ satisfy the continuity equation, in the sense that for all $g \in \mathrm{Cyl}_c(\mathbb{R}^{\mathbb{N}})$, we have

$$\int_{\mathbb{R}^{\mathbb{N}}} g\,\mathrm{d}\sigma_t - \int_{\mathbb{R}^{\mathbb{N}}} g\,\mathrm{d}\sigma_s = \int_s^t \int_{\mathbb{R}^{\mathbb{N}}} (\mathbf{W}_r, \nabla g)\,\mathrm{d}\sigma_r\,\mathrm{d}r \qquad \text{for every } s,t \in [0,T].$$

Indeed, take any $g \in \mathrm{Cyl}_c(\mathbb{R}^{\mathbb{N}})$ and define $F := g \circ \mathbb{T}$, i.e.

$$F(\nu) = g\left(\langle f_1, \nu\rangle, \dots, \langle f_m, \nu\rangle\right),$$

for some $m \in \mathbb{N}$. Note that $F \in \mathrm{Cyl}_c(\Gamma)$, and therefore since $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$,

$$\begin{aligned}
\int_{\mathbb{R}^{\mathbb{N}}} g(z)\,\sigma_t(\mathrm{d}z) - \int_{\mathbb{R}^{\mathbb{N}}} g(z)\,\sigma_s(\mathrm{d}z) &= \int_\Gamma F\mathrm{d}\mathsf{P}_t - \int_\Gamma F\mathrm{d}\mathsf{P}_s \\
&= \int_s^t \int_{\Gamma\times\mathcal{T}} (\mathrm{grad}_\Gamma F)(\nu,x)(\mathsf{J}^+_r - \mathsf{J}^-_r)(\mathrm{d}\nu,\mathrm{d}x)\,\mathrm{d}r \\
&= \int_s^t \int_\Gamma \sum_i (\partial_i g)(\mathbb{T}(\nu)) \left(\int_{\mathcal{T}} f_i(x)\lambda^{\text{net}}_{r,\nu}(\mathrm{d}x)\right) \mathsf{P}_r(\mathrm{d}\nu)\,\mathrm{d}r \\
&= \int_s^t \int_{\mathbb{R}^{\mathbb{N}}} \nabla g(z)\cdot \mathbf{W}_r(z)\,\sigma_r(\mathrm{d}z)\,\mathrm{d}r.
\end{aligned}$$

Thus, we are now in a position to apply Theorem E.1, and obtain a Borel probability measure Ω over $C([0,T];\mathbb{R}^{\mathbb{N}})$ satisfying $(e_t)_\#\Omega = \sigma_t$ for all $t \in [0,T]$, and which is concentrated on the family of curves $z \in AC([0,T];\mathbb{R}^{\mathbb{N}})$ that are solutions to the ODE

$$\dot z_t = \mathbf{W}_t(z_t) \qquad \text{for almost every } t \in [0,T].$$

Note that since $\mathrm{supp}(\sigma) \subseteq \mathbb{T}(\Gamma)$, we have $\mathrm{supp}(\Omega) \subseteq AC([0,T];\mathbb{T}(\Gamma))$. Now let $\tilde{\mathbb{T}} : C([0,T];\Gamma) \to C([0,T];\mathbb{R}^{\mathbb{N}})$ be defined via $(\tilde{\mathbb{T}}(\nu))_t := \mathbb{T}(\nu_t)$. Similar as for $\mathbb{T}$, $\tilde{\mathbb{T}}$ is injective and an isometry when seen as a map $\tilde{\mathbb{T}} : AC([0,T];(\Gamma, \|\cdot\|_{TV})) \to AC([0,T];(\mathbb{R}^{\mathbb{N}}, |\cdot|_\infty))$. Therefore, it is clear the measure $Q := \tilde{\mathbb{T}}^{-1}_\#\Omega \in$

$\mathcal{P}(C([0,T];\Gamma))$ is well-defined, satisfies $\mathsf{P}_t = (e_t)_\# Q$ and is concentrated on the family of curves $\nu \in AC([0,T];(\Gamma, \|\cdot\|_{TV}))$, for which

$$\int_{\mathcal{T}} f_i \,\mathrm{d}\nu_t - \int_{\mathcal{T}} f_i \,\mathrm{d}\nu_s = \int_s^t f_i \,\mathrm{d}(\lambda^+_{r,\nu} - \lambda^-_{r,\nu})\,\mathrm{d}r \qquad \text{for all } s,t \in [0,T],\ i \in \mathbb{N}.$$

Moreover,

$$\begin{aligned}\int_{C([0,T];\Gamma)} \left(\int_0^T \mathcal{R}_{MF}\left(\nu_t, \lambda^+_{t,\nu_t}, \lambda^-_{t,\nu_t}\right) \mathrm{d}t\right) Q(\mathrm{d}\nu) &= \int_0^T \int_\Gamma \mathcal{R}_{MF}(\nu, \lambda^+_{t,\nu}, \lambda^-_{t,\nu})\, \mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}t \\ &= \int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t,\end{aligned}$$

by Fubini-Tonelli and the fact that $\mathcal{R}_{MF} \geq 0$, with a similar statement for $\mathcal{L}_\infty$ and $\mathcal{L}_{MF}$. Note that either $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_\infty$, in which case

$$\begin{aligned}\int_{C([0,T];\Gamma)} \left(\int_0^T \|\lambda^\pm_{t,\nu_t}\|_{TV}\,\mathrm{d}t\right) Q(\mathrm{d}\nu) &= \int_0^T \int_\Gamma \|\lambda^\pm_{t,\nu}\|_{TV}\, \mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}t \\ &= \int_0^T \|\mathsf{J}^\pm\|_{TV}\ \mathrm{d}t < \infty,\end{aligned}$$

or $\mathcal{J}_\infty = 0 < \infty$, or

$$\int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t < \infty.$$

In all three cases we conclude by Lemma 16.12 that $(\nu, \lambda^+_\nu, \lambda^+_\nu) \in \mathscr{CE}$ Q-almost everywhere.

The reverse statement can be derived straightforwardly, similarly to Proposition 18.17, and we omit the proof. □

18.4 Variational characterization

Having all the ingredients at hand, we can now prove the variational characterization for the Liouville equation, namely Theorems 18.9 and 18.10.

Proof of Theorem 18.9. Suppose $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_\infty$. Then by the superposition principle of Theorem 18.18 we obtain a Borel probability measure Q over $C([0,T];\Gamma)$ satisfying $(e_t)_\# Q = \mathsf{P}_t$ for all $t \in [0,T]$ and concentrated on the family of curves $\nu \in AC([0,T];(\Gamma, \|\cdot\|_{TV}))$ for which $(\nu, \lambda^+_\nu, \lambda^-_\nu) \in \mathscr{CE}$. Moreover,

$$\int_{C([0,T];\Gamma)} \left(\int_0^T \mathcal{R}_{MF}\left(\nu_t, \lambda^+_{t,\nu}, \lambda^-_{t,\nu}\right) \mathrm{d}t\right) Q(\mathrm{d}\nu) = \int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t < \infty,$$

and a similar statement for $\mathcal{L}_{MF}, \mathcal{L}_\infty$.

Now assume in addition that $\mathcal{F}_\infty(\mathsf{P}_0) < \infty$, in which case for Q-a.e. curve $\mathcal{F}_{MF}(\nu_0) < \infty$. Moreover, since both $\mathcal{F}_\infty$ and $\mathcal{D}_\infty$ are simply their mean-field counterparts integrated by P, we find

$$\begin{aligned}
&\int_{C([0,T];\Gamma)} \mathcal{J}_{MF}\left(\nu, \lambda^+_\nu, \lambda^-_\nu\right) Q(\mathrm{d}\nu)\\
&= \int_{C([0,T];\Gamma)} \left(\int_0^T \mathcal{R}_{MF}\left(\nu_t, \lambda^+_{t,\nu}, \lambda^-_{t,\nu}\right)\mathrm{d}t + \mathcal{F}_{MF}(\nu_T) - \mathcal{F}_{MF}(\nu_0) + \int_0^T \mathcal{D}_{MF}(\nu_t)\,\mathrm{d}t\right) Q(\mathrm{d}\nu)\\
&= \int_{C([0,T];\Gamma)} \left(\int_0^T \mathcal{R}_{MF}\left(\nu_t, \lambda^+_{t,\nu}, \lambda^-_{t,\nu}\right)\mathrm{d}t\right) Q(\mathrm{d}\nu) + \mathcal{F}_\infty(\mathsf{P}_T) - \mathcal{F}_\infty(\mathsf{P}_0) + \int_0^T \mathcal{D}_\infty(\mathsf{P}_t)\,\mathrm{d}t\\
&= \mathcal{I}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).
\end{aligned} \tag{18.18}$$

Here the second equality follows from Fubini-Tonelli, the fact that $\mathcal{R}_{MF}, H^2, \mathcal{F}_{MF} \geq 0$, $\mathcal{F}_\infty(\mathsf{P}_0) < \infty$ and the bound

$$\sup_{t\in[0,T]} \nu(\mathcal{T})^2 \mathrm{d}\mathsf{P}_t < \infty \tag{18.19}$$

since the latter ensures finiteness of the correction term for $\mathcal{D}_\infty(\mathsf{P}_t)$. Moreover, by the equivalence of $\mathcal{I}_{MF}$ and $\mathcal{J}_{MF}$ it is clear that

$$\mathcal{I}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = \int_{C([0,T];\Gamma)} \mathcal{I}_{MF}\left(\nu, \lambda^+_\nu, \lambda^-_\nu\right) Q(\mathrm{d}\nu) = \mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).$$

In particular, $\mathcal{J}_\infty \geq 0$ and equal to zero if only if $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = (\mathsf{P}, \vartheta^+_{\mathsf{P}}, \vartheta^-_{\mathsf{P}}) \in \mathsf{CE}^*_\infty$, i.e. if $\mathsf{J}^\pm_t = \vartheta^\pm_{\mathsf{P}_t}$ for a.e. $t \in [0,T]$ and P_t is a $\mathcal{I}$-solution with uniformly bounded second moment estimates (18.19).

In either case we have for such a solution the unique characterization $\mathsf{P}_t = (G_t)_\#\mathsf{P}_0$, where $G_t : \Gamma \to \Gamma$ defined by (18.3) maps any $\bar{\nu}$ to the unique solution to (MF) for initial datum $\nu_0 = \bar{\nu}$. This follows from the fact that $\mathcal{J}_\infty = 0$ implies that Q is concentrated on the unique strong solutions of the mean-field equation, in which case Q is characterized by

$$Q = \tilde{G}_\#\mathsf{P}_0,$$

where $\tilde{G} : \Gamma \to C([0,T],\Gamma)$ is defined via $(\tilde{G}(\nu_0))_t := G_t(\nu_0)$.

We conclude the proof by noting that $\mathcal{I}$-solutions propagate moments, as shown in Proposition 18.17. □

Proof of Theorem 18.10. Suppose that the mean-field detailed balance condition holds. Let $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$ be such that $\mathcal{F}_\infty(\mathsf{P}_0) < \infty$ and $\mathcal{I}_\infty < \infty$. Since $\mathcal{F}_\infty$ is nonnegative we have in particular that

$$\int_0^T \mathcal{R}_\infty(\mathsf{P}_t, J_t^+, J_t^-)\,\mathrm{d}t < \infty, \quad \mathcal{F}_\infty(\mathsf{P}_T) < \infty, \quad \int_0^T \mathcal{D}_\infty(\mathsf{P}_t)\,\mathrm{d}t < \infty.$$

Hence, as in the proof of Theorem 18.9, we can again apply the superposition principle of Theorem 18.18, and obtain the corresponding Borel probability measure Q over $C([0,T];\Gamma)$. Moreover, since now $\mathcal{D}_{MF} \geq 0$, we still have the equality

$$\begin{aligned}\mathcal{I}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) &= \int_{C([0,T];\Gamma)} \mathcal{J}_{MF}\left(\nu, \lambda_\nu^+, \lambda_\nu^-\right) Q(\mathrm{d}\nu) \\ &= \int_{C([0,T];\Gamma)} \mathcal{I}_{MF}\left(\nu, \lambda_\nu^+, \lambda_\nu^-\right) Q(\mathrm{d}\nu) \\ &= \mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).\end{aligned}$$

In particular, $\mathcal{J}_\infty \geq 0$ and $\mathcal{J} = 0$ implies that $\mathsf{J}_t^\pm = \vartheta_{\mathsf{P}_t}^\pm$ for a.e. $t \in [0,T]$ and $\mathsf{P}_t = (G_t)_\# \mathsf{P}_0$.

Finally, by the same reasoning, $\mathcal{J}_\infty = 0$ if $\mathsf{J}_t^\pm = \vartheta_{\mathsf{P}_t}^\pm$ for a.e. $t \in [0,T]$, P_t is a $\mathcal{I}$-solution and the superposition principle holds. Since it can be shown that finite $\mathcal{F}_\infty$ implies finite first moment, the result now follows from Theorem 18.9 and again the propagation of moments for $\mathcal{I}$-solutions as stated in Proposition 18.17. □

Chapter 19

EDP convergence

In the previous chapters, we have established variational formulations for the solution to the forward Kolmogorov equation of the interacting particle system, for the solutions to the mean-field equation, and the corresponding Liouville equation. Moreover, for the latter, we have shown how the corresponding EDP-functional can be represented as the expectation over a functional of mean-field paths.

We are now in a position to rigorously discuss the convergence of the forward Kolmogorov equation to the Liouville equation, in terms of EDP-convergence of their variational structures.

To do so, we need slightly more restrictive conditions on $b^{\pm}$ than the one stated in Assumption 16.1. Instead, will work the following set of assumptions.

Assumption 19.1. *The measure-dependent birth and death rates $b^{\pm}$ are of the form*

$$b^{\pm}[\nu](x) = \psi^{\pm}\left(x, \nu, \int_{\mathcal{T}} m^{\pm}(x, y)\, \mathrm{d}\nu(y)\right). \tag{19.1}$$

Where $\psi^{\pm} \in C(\mathcal{T} \times \Gamma \times \mathbb{R}_{\geq 0})$ are nonnegative continuous functions such that

1. *$\psi^{\pm}$ are uniformly Lipschitz in ν, z on bounded sets, i.e. for every $K > 0$ there exists a constant L'_K such that*

$$|\psi^{\pm}(x, \mu, z) - \psi^{\pm}(x, \nu, z')| \leq L'_K \left(\|\mu - \nu\|_{TV} + |z - z'|\right), \tag{19.2}$$

for all $x \in \mathcal{T}$, $\mu, \nu \in \Gamma$ and $z \in \mathbb{R}_{\geq 0}$ such that $\max(\mu(\mathcal{T}), \nu(\mathcal{T}), z) \leq K$.

2. *$\psi^{\pm}$ has linear growth, i.e. there exists a constant $\tilde{M}$ such that*

$$\psi^{\pm}(x, \nu, z) \leq \tilde{M}(\nu(\mathcal{T}) + |z|) \tag{19.3}$$

3. *The kernels $m^{\pm}$ are nonnegative and bounded, i.e. $m^{\pm} \in B_b^{+}(\mathcal{T}^2)$.*

It is clear that the above conditions imply Assumption 16.1 for suitable constants. Moreover, note the similarity to the assumptions of Part I.B, in the sense that $b^{\pm}$ is allowed to be non-continuous but only via integral functionals in the form of $\int m^{\pm} \mathrm{d}\nu$. In fact, to lift from continuous kernels $m^{\pm}$ to bounded measurable ones we use the fact that the rescaled entropies $\mathcal{F}_n$ are bounded, and employ a similar large deviation estimate as in Chapter 4 of Part I, as will be shown in Appendix F. As will be clear from the proof one can modify the assumptions to incorporate k-particle and n-dependent interactions as desired, although we will not treat that generalization in this thesis.

Now, let us denote a sequence of curves $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n$ *converging* to a curve $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$, denoted by $\lim_{n\to\infty}(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) = (\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$, if:

- $\mathsf{P}^n_t \to \mathsf{P}_t$ narrowly for all $t \in [0, T]$,
- $\mathsf{J}^{n,\pm}_t(\mathrm{d}\nu, \mathrm{d}x)\,\mathrm{d}t \to \mathsf{J}^{\pm}_t(\mathrm{d}\nu, \mathrm{d}x)\,\mathrm{d}t$ vaguely on $\mathcal{M}^+_{\mathrm{loc}}(\Gamma \times \mathcal{T} \times [0, T])$.

We then have the following convergence result.

Theorem 19.2. *Suppose that a sequence* $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}^*_n$*,* $n \geq 0$*, is such that*

$$\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_0) < \infty, \qquad \limsup_{n\to\infty} \mathcal{J}_n(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) < \infty,$$

and

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^{2+\varepsilon} \mathrm{d}\mathsf{P}^n_t < \infty, \tag{19.4}$$

for some fixed $\varepsilon > 0$*. Then there exists a (not relabelled) subsequence* $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-})$ *and a* $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}^*_\infty$ *such that*

$$\lim_{n\to\infty} (\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) = (\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-),$$

Moreover, for any converging sequence satisfying the bounds above, we have the liminf-estimates

$$\begin{gathered}
\liminf_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_t) \geq \mathcal{F}_\infty(\mathsf{P}_t), \qquad \textit{for all } t \in [0, T], \\
\liminf_{n\to\infty} \int_0^T \mathcal{R}_n(\mathsf{P}^n_t, \mathsf{J}^{n,+}_t, \mathsf{J}^{n,-}_t)\,\mathrm{d}t \geq \int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t, \\
\liminf_{n\to\infty} \int_0^T \mathcal{D}_n(\mathsf{P}^n_t)\,\mathrm{d}t \geq \int_0^T \mathcal{D}_\infty(\mathsf{P}^n_t)\,\mathrm{d}t.
\end{gathered} \tag{19.5}$$

Finally, suppose that the detailed balance condition (17.10) *holds. Then all the above statements are valid without the moment assumption of* (19.4)*, and with replacing* CE^*_n*,*CE^*_∞ *by their less restrictive counterparts* $\mathsf{CE}_n, \mathsf{CE}_\infty$*.*

In all cases, the superposition principle of Theorem 18.18 holds for $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-)$*.*

Remark 19.3. In fact, the compactness result is slightly stronger. As shown in the proof of Theorem 19.2 the measures $\mathsf{J}_r^{n,\pm}(\mathrm{d}v, \mathrm{d}x)\,\mathrm{d}r$ converge vaguely on $\mathcal{M}^+_{\mathrm{loc}}(\Gamma \times \mathcal{T} \times [s,t])$ for any $s,t \in [0,T]$. In particular, the lower semicontinuity results for $\mathcal{R}$ is valid on any interval $[s,t]$, and moreover, the lower semicontinuity for $\mathcal{D}$ holds for every $t \in [0,T]$. □

Remark 19.4. Recall that the Fisher-information functionals $\mathcal{D}_n(\mathsf{P})$ consist of Hellinger distances, which are lower semicontinuous in suitable topologies, and linear correction terms that are on the order of

$$\int_\Gamma v(\mathcal{T})^2 \mathrm{d}\mathsf{P}_t.$$

It is to control these terms that the moment bounds (19.4) are necessary, instead of merely $\varepsilon = 0$ to obtain the chain rule and equivalence for $\mathcal{J}_n$. □

Note that if in addition the initial data is well-prepared, in the sense that

$$\lim_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_0) = \mathcal{F}_\infty(\mathsf{P}_0),$$

then for any converging subsequence satisfying the bounds of Theorem 19.2 , we clearly have the liminf-estimate

$$\liminf_{n\to\infty} \mathcal{J}_n(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \ge \mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-). \tag{19.6}$$

Even more so, in the case of detailed balance the inequality (19.6) always holds, even without the moment assumptions (19.4), and we obtain evolutionary Γ-convergence of $\mathcal{J}_n$ to $\mathcal{J}_\infty$.

Remark 19.5. Evolutionary Γ-convergence is merely another name for appropriate liminf-estimates, and, confusingly, is not equivalent to Γ-convergence of $\mathcal{J}_n$ to $\mathcal{J}_\infty$ over the space of curves. In order to obtain the latter one would need suitable limsup-estimates, which are related to questions on large deviations, as we briefly touch on in Appendix G. □

Now, recall by Theorem 17.11 that unique $\mathcal{I}$-solutions to the forward Kolmogorov equations ($\mathsf{FKE_n}$) exist and propagate moments by Proposition (17.23). Similarly, $\mathcal{I}$-solutions to the Liouville equation (Li) are unique by Theorem 18.9. Therefore, modifying classical arguments from [SS04, Ser11], we can directly conclude convergence for the sequence of solutions.

Theorem 19.6. *Consider a converging sequence $\mathcal{P}(\Gamma_n) \ni \bar{\mathsf{P}}^n \to \bar{\mathsf{P}} \in \mathcal{P}(\Gamma)$ such that*

$$\lim_{n\to\infty} \mathcal{F}_n(\bar{\mathsf{P}}^n) = \mathcal{F}_\infty(\bar{\mathsf{P}}), \tag{19.7}$$

and

$$\limsup_{n\to\infty} \int_\Gamma \nu(\mathcal{T})^{2+\varepsilon} \mathrm{d}\bar{\mathsf{P}} < \infty, \tag{19.8}$$

for some fixed $\varepsilon > 0$*. For each* $n \geq 0$ *let* P^n_t *be the unique* $\mathcal{I}$*-solution to* ($\mathsf{FKE_n}$) *with initial datum* $\bar{\mathsf{P}}^n$*. Then there exists a unique* $\mathcal{I}$*-solution* P_t *to* (Li) *with initial datum* $\bar{\mathsf{P}}$*. Moreover, we have the convergence*

$$\lim_{n\to\infty} (\mathsf{P}^n, \vartheta^+_{\mathsf{P}^n}, \vartheta^-_{\mathsf{P}^n}) = (\mathsf{P}, \vartheta^+_{\mathsf{P}}, \vartheta^-_{\mathsf{P}})$$
$$\lim_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_t) = \mathcal{F}_\infty(\mathsf{P}_t), \qquad \textit{for all } t \in [0,T].$$

Moreover, suppose that the detailed balance condition (17.10) *holds. Then the above statements are valid without the moment assumption 19.8.*

Proof. Recall that $\mathcal{J}_n(\mathsf{P}^n, \vartheta^+_{\mathsf{P}^n}, \vartheta^-_{\mathsf{P}^n}) = 0$ for all $n \geq 0$. Moreover, setting $p = 2 + \varepsilon$ and applying Proposition (17.23), the p-moments propagate as

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^p \, \mathsf{P}^n_t(\mathrm{d}\nu) \leq \left(\limsup_{n\to\infty} \left(\int_\Gamma \nu(\mathcal{T})^p \, \mathsf{P}^n_0(\mathrm{d}\nu) \right) + TC_{1,p} \right) e^{TC_{1,p}}.$$

Therefore, by (19.7), (19.8), and Theorem 19.2 we have for any subsequence indexed by n' converging to a $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty$ that (19.6) holds, and hence

$$0 = \liminf_{n'\to\infty} \mathcal{J}_n(\mathsf{P}^{n'}, \vartheta^+_{\mathsf{P}^{n'}}, \vartheta^-_{\mathsf{P}^{n'}}) \geq \mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-).$$

Thus $\mathcal{J}_\infty(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) = 0$, which implies that P_t is the unique gradient-flow solution to (Li) and $\mathsf{J}^\pm_t = \vartheta^\pm_{\mathsf{P}_t}$ for a.e. $t \in [0,T]$. The convergence of P^n_t now follows from a compactness argument, and by lower semicontinuity we conclude that for every $t \in [0,T]$ we have the liminf-estimate $\liminf_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_t) \geq \mathcal{F}_\infty(\mathsf{P}_t)$, and

$$\begin{aligned}\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_t) &\leq \limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_0) - \liminf_{n\to\infty} \int_0^t \left(\mathcal{R}_n(\mathsf{P}^n, \vartheta^+_{\mathsf{P}^n}, \vartheta^-_{\mathsf{P}^n}) + \mathcal{D}_n(\mathsf{P}^n_t) \right) \mathrm{d}t \\ &= \mathcal{F}_\infty(\mathsf{P}_0) - \int_0^t \left(\mathcal{R}_\infty(\mathsf{P}, \vartheta^+_{\mathsf{P}}, \vartheta^-_{\mathsf{P}}) + \mathcal{D}_\infty(\mathsf{P}) \right) \mathrm{d}t = \mathcal{F}_\infty(\mathsf{P}_t).\end{aligned}$$

The case of detailed balance follows via a similar argument. □

Now suppose that in addition the initial sequence of measures $\bar{\mathsf{P}}^n$ is chaotic, in the sense that

$$\bar{\mathsf{P}}^n \to \delta_{\bar{\nu}} \quad \text{narrowly for some } \bar{\nu} \in \Gamma.$$

Then as a consequence of Theorem 19.6 we have propagation of chaos, namely

$$\bar{\mathsf{P}}^n \to \delta_{\bar{\nu}_t} \quad \text{narrowly for all } t \in [0,T],$$

where ν_t is the unique solution to the mean-field equation (16.11) with initial datum $\bar{\nu}$. As mentioned in the Introduction, while for interacting particle systems with the number of particles fixed at $n \in \mathbb{N}$ this would imply narrow convergence of the k-marginals at time t to $\nu_t^{\otimes k}$ (e.g. see [Szn91]), in our setting this implies convergence of the k-correlation functions [BGSRS20].

Moreover, note that we have a stronger notion of convergence, since the free energies $\mathcal{F}_n$ converge as well. Under appropriate conditions on the initial datum $\bar{\nu}$, this guarantees a version of propagation of entropic chaoticity. Namely, for any $\nu \in \Gamma$ we define the Poisson measures

$$\Pi_{n,\nu} := (L_n)_\# \pi_{n,\nu}, \qquad \text{where} \qquad \pi_{n,\nu} := \frac{1}{e^{n\nu(\mathcal{T})}} \sum_{N=0}^{\infty} \frac{n^N}{N!} \nu^{\otimes N}.$$

It is straightforward to check that $\Pi_{n,\nu} \to \delta_\nu$ narrowly for any $\nu \in \Gamma$. We then have the following result.

Theorem 19.7 (Propagation of chaos)**.** *Consider the setting of Theorem 19.6 and assume additionally that $\bar{\mathsf{P}} = \delta_{\bar{\nu}}$ for some $\bar{\nu} \in \Gamma$ with $\mathcal{E}\mathrm{nt}(\bar{\nu}|\gamma) < \infty$. Let ν_t be the unique solution to* (MF) *with initial datum $\bar{\nu}$. Then for all $t \in [0, T]$,*

$$\mathsf{P}_t^n \to \delta_{\nu_t} \quad \textit{narrowly}, \qquad \textit{and} \qquad \lim_{n\to\infty} \mathcal{E}\mathrm{nt}(\mathsf{P}_t^n|\Pi_n) = \mathcal{E}\mathrm{nt}(\nu_t|\gamma).$$

If additionally there exists a constant $C > 1$ such that $C^{-1} \leq \mathrm{d}\bar{\nu}/\mathrm{d}\gamma \leq C$ then

$$\lim_{n\to\infty} \mathcal{E}\mathrm{nt}(\mathsf{P}_t^n|\Pi_{n,\nu_t}) = 0, \qquad \textit{for all } t \in [0, T].$$

Theorems 19.2 and 19.7 are proved in Section 19.3. However, first we show Γ-convergence of the free energies in Section 19.1, and establish the necessary estimates in Section 19.2.

19.1 $\mathit{\Gamma}$-convergence of $\mathcal{F}_n$

While only the liminf-estimates for the free energy $\mathcal{F}_n$ are necessary for the proof of Theorem 19.2 and the convergence of solutions, we provide here the full $\mathit{\Gamma}$-convergence result. We rely strongly on the characterization of [Mar12], which connects a large deviation principle with rate function I to the fact that

$$\mathit{\Gamma}\text{-}\lim_{n\to\infty} \frac{1}{n} \mathcal{E}\mathrm{nt}(\mathsf{P}|\Pi^n) = \int_\Gamma I(\nu)\, \mathsf{P}(\mathrm{d}\nu),$$

and provides useful sufficient conditions for both.

Recall in our setting that

$$\mathcal{F}_n(\mathsf{P}) = \frac{1}{2n}\mathcal{E}\mathrm{nt}(\mathsf{P}|\Pi_n), \qquad \mathcal{F}_\infty(\mathsf{P}) = \frac{1}{2}\int_\Gamma \mathcal{E}\mathrm{nt}(\nu|\gamma)\,\mathsf{P}(\mathrm{d}\nu).$$

We then have the following result, which we prove after Lemma 19.10 below.

Theorem 19.8. *The family $\{\mathcal{F}_n\}_{n>0}$ is equicoercive and Γ-converges to $\mathcal{F}$ in the sense that*

- *for any converging sequence $\mathsf{P}^n \to \mathsf{P} \in \mathcal{P}(\Gamma)$:*

$$\mathcal{F}_\infty(\mathsf{P}) \le \liminf_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n),$$

- *for any $\mathsf{P} \in \mathcal{P}(\Gamma)$ with $\mathcal{F}_\infty(\mathsf{P}) < \infty$ there exists a sequence $\mathsf{P}^n \in \Gamma$ converging to P such that*

$$\lim_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n) = \mathcal{F}_\infty(\mathsf{P}).$$

Remark 19.9. The equicoercivity directly follows from the fact that first moments are bounded by $\mathcal{F}_n$. To be precise, we have the estimate

$$\gamma(\mathcal{T})\phi\left(\gamma(\mathcal{T})^{-1}\int_\Gamma \nu(\mathcal{T})\mathsf{P}(\mathrm{d}\nu)\right) \le 2\mathcal{F}_n(\mathsf{P}), \tag{19.9}$$

for any $n > 0$ and $\mathsf{P} \in \mathcal{P}(\Gamma)$. □

By the results of [Mar12, Theorems 3.4, 3.5] it is sufficient to merely show the corresponding bounds or limits for any P of the form $\mathsf{P} = \delta_\nu$ for some $\nu \in \Gamma$. Because of this reduction, we can make use of the so-called cumulant generating functionals G_n given by

$$G_n(f) := \frac{1}{n}\log\int_\Gamma e^{n\langle f,\nu\rangle}\,\Pi_n(\mathrm{d}\nu),$$

for any $f \in \mathcal{B}_b(\Gamma)$, and their limit counterpart

$$G(f) := \int_{\mathcal{T}} (e^f - 1)\,\mathrm{d}\gamma.$$

Note that by duality of the entropy, we have for all $n > 0$ the inequality

$$\int_\Gamma \langle f,\nu\rangle\,\mathrm{d}\mathsf{P} \le \frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}|\Pi_n) + G_n(f), \tag{19.10}$$

and for the Legendre-dual of G we have

$$G^*(\nu) := \sup_{f\in C_b(\mathcal{T})}\big\{\langle f,\nu\rangle - G(f)\big\} = \mathcal{E}\mathrm{nt}(\gamma|\nu).$$

We will first simplify G_n and show that in fact it is equal to G, which is a property of the Poisson distribution and the fact that for a fixed number of particles the particle positions are i.i.d. variables.

Lemma 19.10. *Let $f \in \mathcal{B}_b(\mathcal{T})$. Then for each $n > 0$*

$$G_n(f) = G(f).$$

Throughout we will use the convention

$$\int_{\mathcal{T}^0} \gamma^{\otimes 0} = 1.$$

Proof. Using the representation for the Poisson measure Π_n we have

$$\begin{aligned}\int_\Gamma e^{n\langle f,\nu\rangle}\Pi_n(\mathrm{d}\nu) &= \frac{1}{e^{n\gamma(\mathcal{T})}}\sum_{N=0}^\infty \frac{n^N}{N!}\int_{\mathcal{T}^N} e^{\sum_{i=1}^N f(x_i)}\mathrm{d}\gamma^{\otimes N} \\ &= \frac{1}{e^{n\gamma(\mathcal{T})}}\sum_{N=0}^\infty \frac{n^N\left(\int_\mathcal{T} e^f\,\mathrm{d}\gamma\right)^N}{N!} = \frac{e^{n\int_\mathcal{T} e^f\,\mathrm{d}\gamma}}{e^{n\gamma(\mathcal{T})}},\end{aligned}$$

and after taking logarithms and dividing by n we obtain the desired statement. □

Next, we establish convergence for suitable linear functionals of ν.

Lemma 19.11. *Suppose that the sequence P^n converges narrowly and*

$$\limsup_{n\to\infty} \mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) < \infty.$$

Then for any $f \in B_b(\mathcal{T})$ it holds that

$$\lim_{n\to\infty}\int_\Gamma \langle f,\nu\rangle\,\mathrm{d}\mathsf{P}^n = \int_\Gamma \langle f,\nu\rangle\,\mathrm{d}\mathsf{P}. \tag{19.11}$$

In Appendix F, we will even prove convergence for quadratic functionals if the mass of $\nu(\mathcal{T})$ is appropriately controlled.

Remark 19.12. In the proof below we use the finiteness of exponential moments for Π_n and a large deviation estimate for the Poisson distribution, but it should be noted that one can also go the route of establishing super-exponential moments for Π_n, which establishes super-linear moments for P^n. □

Proof. First, let us consider $f \in C_b(\mathcal{T})$, and introduce the functions $F(\nu) := \langle f, \nu\rangle$ and its truncation $F_L(\nu) := \alpha_L(\nu(\mathcal{T}))\langle f, \nu\rangle$, where $\alpha_L(z) := \bar{\alpha}(z-L)$ with $\bar{\alpha} \in C_b(\mathbb{R})$ a continuous non-increasing function such that $0 \le \bar{\alpha}(z) \le 1$ for all z, $\bar{\alpha}(z) = 1$ for $z \le 0$, and $\bar{\alpha}(z) = 0$ for all $z \ge 1$.

Note that $F_L(\nu) \uparrow F(\nu)$ as $L \to \infty$ and that F_L is continuous and bounded for every $L \ge 0$. Hence,

$$\lim_{n\to\infty} \int_\Gamma F_L \, \mathsf{P}^n(\mathrm{d}\nu) = \int_\Gamma F_L \, \mathrm{d}\mathsf{P}, \qquad \lim_{L\to\infty} \int_\Gamma F_L \, \mathrm{d}\mathsf{P} = \int_\Gamma \langle f, \nu\rangle \, \mathrm{d}\mathsf{P}.$$

We will show that

$$\limsup_{L\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \int_\Gamma e^{n\beta |F_L - F|} \, \mathrm{d}\Pi_n = 0, \qquad \text{for all } \beta \ge 0. \tag{19.12}$$

From this we can obtain (19.11) since by duality,

$$\int_\Gamma |F_L - F|(\nu) \, \mathrm{d}\mathsf{P}^n \le \frac{1}{\beta}\left(\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) + \frac{1}{n}\log \int_\Gamma e^{n\beta|F_L-F|}\,\mathrm{d}\Pi_n\right), \qquad \text{for every } \beta, L \ge 0.$$

Taking subsequent limits in n, L and β to infinity, we deduce

$$\limsup_{L\to\infty} \limsup_{n\to\infty} \int |F_L - F| \, \mathrm{d}\mathsf{P}^n = 0,$$

thus proving the desired equality (19.11).

Now, to establish (19.12), first note that $|F_L - F|(\nu) \le |\alpha_L(N/n) - 1|\langle |f|, \nu\rangle$, and therefore

$$\begin{aligned}
\int_\Gamma e^{n\beta|F_L-F|}\,\mathrm{d}\Pi_n &\le \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^\infty \frac{n^N}{N!} \int_{\mathcal{T}^N} e^{\beta|\alpha_L(N/n)-1|\sum_{i=1}^N |f|(x_i)} \mathrm{d}\gamma^{\otimes N} \\
&= \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^\infty \frac{n^N}{N!} \left(\int_{\mathcal{T}} e^{\beta|\alpha_L(N/n)-1||f|(x)}\gamma(\mathrm{d}x)\right)^N \\
&\le \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^\infty \frac{n^N}{N!}\gamma(\mathcal{T})^N + \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=\lfloor nL\rfloor}^\infty \frac{n^N}{N!} \left(\int_{\mathcal{T}} e^{\beta\|f\|_\infty}\gamma(\mathrm{d}x)\right)^N \\
&= 1 + \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=\lfloor nL\rfloor}^\infty \frac{n^N}{N!} C_\beta^N,
\end{aligned}$$

with $C_\beta := e^{\beta\|f\|_\infty}\gamma(\mathcal{T})$. Suppose X_n is a Poisson variable with mean nC_β. Then the second term in the previous estimate can be expressed as

$$\frac{1}{e^{nC_\beta}} \sum_{N=\lfloor nL\rfloor}^\infty \frac{n^N}{N!} C_\beta^N = \mathrm{Prob}\left(\tfrac{1}{n}X_n \ge \tfrac{1}{n}\lfloor nL\rfloor\right).$$

It is clear that $\frac{1}{n}X_n \to C_\beta$ almost surely as $n \to \infty$. Moreover, by elementary large deviation results, e.g. as in Cramer's theorem [DZ10, Theorem 2.2.3], it satisfies a large deviation principle with rate n and rate function $I_\beta(z) := z\log(z/C_\beta)-z+C_\beta$, which implies

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Prob}\left(\tfrac{1}{n}X_n \geq a\right) \leq -\inf_{z\geq a} I_\beta(z).$$

Note that $\inf_{z\geq a} I_\beta(z) = I_\beta(a)$ for $a \geq C_\beta$, and hence for $L \geq C_\beta$ we obtain the bound

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \int_\Gamma e^{n\beta|F_L-F|}\,\mathrm{d}\Pi_n &\leq \limsup_{n\to\infty} \frac{1}{n} \log \max\left\{1, \frac{e^{nC_\beta}}{e^{n\gamma(\mathcal{T})}} e^{-nC_\beta} \sum_{N=\lfloor nL\rfloor}^{\infty} \frac{n^N}{N!} C_\beta^N\right\} \\ &\leq \max\left\{0, (C_\beta - \gamma(\mathcal{T})) - I_\beta(L)\right\}.\end{aligned}$$

Letting $L \to \infty$, we deduce

$$\limsup_{L\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \int_\Gamma e^{n\beta|F_L-F|}\,\mathrm{d}\Pi_n \leq \max\left\{0, (C_\beta - \gamma(\mathcal{T})) - \liminf_{L\to\infty} I_\beta(L)\right\} = 0.$$

Finally, let us now consider $f \in \mathcal{B}_b$. Using a similar density approach as above it is sufficient to show that there exists a sequence of bounded continuous functions f_k, such that

$$\limsup_{k\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \int_\Gamma e^{n\beta\langle|f-f_k|,\nu\rangle}\,\mathrm{d}\Pi_n = 0, \qquad \text{for all } \beta > 0,$$

but, by Lemma 19.10, we have

$$\lim_{n\to\infty} \frac{1}{n} \log \int_\Gamma e^{n\beta\langle|f-f_k|,\nu\rangle}\,\mathrm{d}\Pi_n = \int_{\mathcal{T}} \left(e^{\beta|f-f_k|} - 1\right) \mathrm{d}\gamma.$$

Similar as for density statements in $L^p(\gamma)$ one can find a sequence such that the above integrals vanish as $k \to \infty$, see for example Theorem C.5 of Part I. □

Proof of Theorem 19.8. First, we will show that the family $\{\mathcal{F}_n\}_{n>0}$ is equicoercive, by establishing a first moment bound for P in terms of the mass $\nu(\mathcal{T})$. Namely, setting $f \equiv 1$ in (19.10) we have for any $\mathsf{P} \in \mathcal{P}(\Gamma)$, $n > 0$, $\beta > 0$, the inequality

$$\beta \int_\Gamma \nu(\mathcal{T})\,\mathrm{d}\mathsf{P} \leq \frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}|\Pi_n) + G_n(\beta) = 2\mathcal{F}_n(\mathsf{P}) + (e^\beta - 1)\gamma(\mathcal{T}),$$

where the final term is bounded from above independently of P. Optimizing over β gives us the bound (19.9).

Next, for the limit inferior, consider a narrowly converging sequence $\mathsf{P}^n \to \mathsf{P} = \delta_{\bar\nu}$ for some $\bar\nu \in \Gamma$. Fix any $f \in C_b(\mathcal{T})$. By duality, we have for every n,

$$\frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) \geq \int_\Gamma \langle f, \nu\rangle \, \mathrm{d}\mathsf{P}^n - \frac{1}{n}\log\int_\Gamma e^{n\langle f,\nu\rangle}\, \mathrm{d}\Pi_n,$$

and due to Lemmas 19.10 and 19.11 and,

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) &\geq \liminf_{n\to\infty} \int_\Gamma \langle f, \nu\rangle \, \mathrm{d}\mathsf{P}^n - \frac{1}{n}\log\int_\Gamma e^{n\langle f,\nu\rangle}\, \mathrm{d}\Pi_n \\ &= \langle f, \bar\nu\rangle - G(f).\end{aligned}$$

Taking the supremum over all $f \in C_b(\mathcal{T})$ we find

$$\mathcal{F}_\infty(\delta_{\bar\nu}) = \frac{1}{2}\mathcal{E}\mathrm{nt}(\bar\nu|\gamma) \leq \liminf_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n).$$

Finally, consider an arbitrary $\bar\nu \in \Gamma$ with $\mathcal{E}\mathrm{nt}(\bar\nu|\gamma) < \infty$ and set $\mathsf{P} = \delta_{\bar\nu}$. We will construct a sequence of measures P^n that locally consists of Poisson measures induced by $\bar\nu$. Namely, set

$$\Pi_{n,\bar\nu} := (L_n)_\# \pi_{n,\bar\nu}, \qquad \text{with} \qquad \pi_{n,\bar\nu} := \frac{1}{e^{n\bar\nu(\mathcal{T})}}\sum_{N=0}^\infty \frac{n^N}{N!}\bar\nu^{\otimes N},$$

and consider the sequence $\mathsf{P}^n := \Pi_{n,\bar\nu}$. It is straightforward to verify that indeed $\mathsf{P}^n \to \delta_{\bar\nu}$. Moreover, note that although L_n is not bijective, we do have the equality

$$\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) = \mathcal{E}\mathrm{nt}(\pi_{n,\bar\nu}|\pi_n),$$

due to the symmetry of the N-particle distributions $\bar\nu^{\otimes N}$, $\gamma^{\otimes N}$. Therefore, we derive

$$\begin{aligned}\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) &= \mathcal{E}\mathrm{nt}(\pi_{n,\bar\nu}|\pi_n) \\ &= \frac{1}{e^{n\bar\nu(\mathcal{T})}}\sum_{N=0}^\infty \frac{n^N}{N!}\int_{\mathcal{T}^N} \log\left(\frac{e^{n\gamma(\mathcal{T})}}{e^{n\bar\nu(\mathcal{T})}}\frac{\mathrm{d}\bar\nu^{\otimes N}}{\mathrm{d}\gamma^{\otimes N}}\right)\mathrm{d}\bar\nu^{\otimes N} \\ &= \frac{1}{e^{n\bar\nu(\mathcal{T})}}\sum_{N=1}^\infty \frac{n^N}{N!} N\bar\nu(\mathcal{T})^{N-1}\int_{\mathcal{T}} \log\left(\frac{\mathrm{d}\bar\nu}{\mathrm{d}\gamma}\right)\mathrm{d}\bar\nu \\ &\quad + \frac{1}{e^{n\bar\nu(\mathcal{T})}}\sum_{N=0}^\infty \frac{n^N}{N!}\left(\bar\nu(\mathcal{T})^N \log\frac{e^{n\gamma(\mathcal{T})}}{e^{n\bar\nu(\mathcal{T})}}\right) \\ &= n\int_{\mathcal{T}} \log\left(\frac{\mathrm{d}\bar\nu}{\mathrm{d}\gamma}\right)\mathrm{d}\bar\nu + n\gamma(\mathcal{T}) - n\bar\nu(\mathcal{T}).\end{aligned}$$

Rescaling, we obtain

$$\frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) = \int_{\mathcal{T}} \log\left(\frac{\mathrm{d}\bar{\nu}}{\mathrm{d}\gamma}\right)\mathrm{d}\bar{\nu} - \bar{\nu}(\mathcal{T}) + \gamma(\mathcal{T}) = \mathcal{E}\mathrm{nt}(\bar{\nu}|\gamma),$$

therewith concluding the proof. □

19.2 Uniform estimates

In Section 17.2 we provided uniform-in-n estimates for the flux. Therefore, from Lemma 17.19 and the fact that $M_n \leq 2M$ for sufficiently large n, we directly have the following.

Corollary 19.13. *Consider a sequence* $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n$ *such that*

$$\limsup_{n\to\infty} \int_0^T \mathcal{R}_n(\mathsf{P}^n_t, \mathsf{J}^{n,+}_t, \mathsf{J}^{n,-}_t) < \infty.$$

Then

$$\limsup_{n\to\infty} \int_0^T 2M\tilde{\phi}\left(\frac{1}{2M}\int_{\Gamma\times\mathcal{T}} (1+\nu(\mathcal{T})^2)^{-1}\,\mathsf{J}^{n,\pm}_t(\mathrm{d}\nu,\mathrm{d}x)\right)\mathrm{d}t < \infty.$$

If in addition

$$A := \sup_{t\in[0,T]} \int_\Gamma (1+\nu(\mathcal{T})^2)\mathsf{P}_t(\mathrm{d}\nu) < \infty,$$

then it holds that

$$\limsup_{n\to\infty} \int_0^T 2AM\tilde{\phi}\left(\frac{1}{2AM}\|\mathsf{J}^{n,\pm}_t\|_{TV}\right)\mathrm{d}t < \infty,$$

The weighted total variation metric $d_{TV,w}$ that was introduced is not appropriate for taking limits, and instead, we take the weaker metric defined in (18.12):

$$W(\mathsf{P}^1,\mathsf{P}^2) := \sup_{F\in\mathbb{F}}\left\{\int_\Gamma F\,\mathrm{d}(\mathsf{P}^1-\mathsf{P}^2)\right\},$$

where

$$\mathbb{F} := \left\{F \in \mathrm{Cyl}_c(\Gamma) \;:\; \left\|(1+\nu(\mathcal{T})^2)\,\mathrm{grad}_\Gamma F\right\|_\infty \leq 1\right\}.$$

Recall that W is narrowly lower semicontinuous and implies narrow convergence on narrowly pre-compact subsets. We now have the follow equicontinuity result.

Lemma 19.14. *Consider a sequence* $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n$ *such that*

$$\limsup_{n\to\infty} \int_0^T \mathcal{R}_n(\mathsf{P}^n_t, \mathsf{J}^{n,+}_t, \mathsf{J}^{n,-}_t)\,\mathrm{d}t < \infty.$$

Then

$$\limsup_{n\to\infty} \int_0^T \tilde{\phi}\left(\frac{|\dot{\mathsf{P}}^n_t|_W}{8M}\right) < \infty,$$

where $|\dot{\mathsf{P}}_t|_W$ *is the* W*-metric speed.*

Remark 19.15. Note that in the case of $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}^*_n$ with uniformly bounded second moment we can actually use a stronger distance, but with our eye on the case of detailed balance as well we employ W instead. □

Proof. The proof is similar to Lemmas 14.9 and 18.16, now for the distance W instead of the weighted total variation metric $d_{TV,w}$. Namely, fix $n > 0$ and consider a curve $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_n$. Then we have for any $s, t \in [0, T]$ and any $F \in C_c(\Gamma)$,

$$\left|\int_\Gamma F \mathrm{d}(\mathsf{P}_t - \mathsf{P}_s)\right| \le \int_s^t \int_{\Gamma\times\mathcal{T}} |\overline{\nabla}^{n,+} F(\nu, x)|\,\mathrm{d}\mathsf{J}^+_r\,\mathrm{d}r + \int_s^t \int_{\Gamma\times\mathcal{T}} |\overline{\nabla}^{n,-} F(\nu, x)|\,\mathrm{d}\mathsf{J}^-_r\,\mathrm{d}r.$$

Substituting any $F \in \mathbb{F}$ it is straightforward to verify that

$$\begin{aligned}
|\overline{\nabla}^{n,+} F(\nu, x)| &= n|F(\nu + \tfrac{1}{n}\delta_x) - F(\nu)| \le (1 + \nu(\mathcal{T})^2)^{-1},\\
|\overline{\nabla}^{n,-} F(\nu, x)| &= n|F(\nu) - F(\nu - \tfrac{1}{n}\delta_x)| \le (1 + (\nu(\mathcal{T}) - \tfrac{1}{n})^2)^{-1} \le 2(1 + \nu(\mathcal{T})^2)^{-1},
\end{aligned}$$

where the last line holds for all ν with $\nu(\mathcal{T}) \ge \frac{1}{n}$ and sufficiently large n, and therefore

$$\left|\int_\Gamma F \mathrm{d}(\mathsf{P}_t - \mathsf{P}_s)\right| \le 2\int_s^t \int_{\Gamma\times\mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1}\,\mathrm{d}(\mathsf{J}^+_r + \mathsf{J}^-_r)\,\mathrm{d}r.$$

Taking the supremum over $F \in \mathbb{F}$, we find that $(\mathsf{P}_t)_{t\in[0,T]}$ is absolutely continuous w.r.t. W with

$$|\dot{\mathsf{P}}_t|_W \le 2\int_{\Gamma\times\mathcal{T}} (1 + \nu(\mathcal{T})^2)^{-1}\,\mathrm{d}(\mathsf{J}^+_t + \mathsf{J}^-_t) \qquad \text{for a.e. } t \in [0, T],$$

where $|\dot{\mathsf{P}}^n_t|_W$ is the W-metric speed. Applying the estimates in Lemma 17.19 concludes the proof. □

19.3 Proof of main results

We finally conclude with the proof of the main results.

Proof of Theorem 19.2. We will first establish the liminf-estimates. Namely, consider a sequence $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}_n^*$ that converges to the curve $(\mathsf{P}, \mathsf{J}^+, \mathsf{J}^-) \in \mathsf{CE}_\infty^*$. In particular, $\mathsf{P}_t^n \to \mathsf{P}_t$ for all $t \in [0, T]$, and hence by Theorem 19.8 on the Γ-convergence of $\mathcal{F}_n$ we immediately obtain

$$\liminf_{n\to\infty} \mathcal{F}_n(\mathsf{P}_t^n) \geq \mathcal{F}_\infty(\mathsf{P}_t), \qquad \text{for all } t \in [0, T].$$

Now suppose that

$$\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}_0^n) < \infty, \qquad \limsup_{n\to\infty} \mathcal{J}_n(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) < \infty,$$

and that we have the moment bound

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \int_\Gamma \nu(\mathcal{T})^{2+\varepsilon} \mathrm{d}\mathsf{P}_t^n < \infty. \tag{19.13}$$

In particular, using the non-negativity of $\mathcal{R}$ and the Hellinger distance, and the uniform bounds on the correction terms of the Fisher information, one can establish

$$\limsup_{n\to\infty} \int_0^T \mathcal{R}_n(\mathsf{P}_t^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-})\, \mathrm{d}t < \infty, \quad \limsup_{n\to\infty} \int_0^T \mathcal{D}_n(\mathsf{P}_t^n)\, \mathrm{d}t < \infty. \tag{19.14}$$

Due to the chain rule and the assumption on $\mathcal{F}_n(\mathsf{P}_0^n)$, we obtain

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \mathcal{F}_n(\mathsf{P}_t^n) < \infty. \tag{19.15}$$

The latter guarantees, by Theorem F.1, that for every $t \in [0, T]$ we have the vague convergence

$$\lim_{n\to\infty} \vartheta^\pm_{\mathsf{P}_t^n} = \vartheta^\pm_{\mathsf{P}_t}, \qquad \lim_{n\to\infty} \mathsf{T}_\#^{n,\mp} \bar{\vartheta}^{n,\pm}_{\mathsf{P}_t^n} = \bar{\vartheta}^\pm_{\mathsf{P}_t}, \qquad \lim_{n\to\infty} \vartheta^{n,\pm}_{\mathsf{P}_t^n} = \bar{\vartheta}^\pm_{\mathsf{P}_t}.$$

In fact, since $\chi_\nu^\pm(\mathcal{T})$, $\bar{\chi}^{n,\pm}(\mathcal{T})$ grow as $\nu(\mathcal{T})^2$ as shown in Corollary 19.13 , one can show that due to $(2 + \varepsilon)$-moment estimates of (19.13) the vague convergence is lifted to narrow convergence. Now recall that from Lemma 11.2 and the definition of the Fisher information that

$$\mathcal{E}\mathrm{nt}\left(\mathsf{J}_t^{n,\pm} | \Theta_\mathsf{P}^{n,+}\right) = \int_{\Gamma\times\mathcal{T}} \Upsilon\left(\frac{\mathrm{d}\mathsf{J}_t^{n,\pm}}{\mathrm{d}\Sigma}, \frac{\mathrm{d}\vartheta^\pm_{\mathsf{P}_t}}{\mathrm{d}\Sigma}, \frac{\mathrm{d}(\mathsf{T}_\#^{n,\mp}\bar{\vartheta}^{\pm,n}_{\mathsf{P}_t})}{\mathrm{d}\Sigma}\right)\mathrm{d}\Sigma,$$

$$\mathcal{D}_n(\mathsf{P}_t) = H^2(\vartheta^+_\mathsf{P}, \mathsf{T}_\#^{n,-}\bar{\vartheta}^{+,n}_\mathsf{P}) + H^2(\vartheta^-_\mathsf{P}, \mathsf{T}_\#^{n,+}\bar{\vartheta}^{-,n}_\mathsf{P}) + \frac{1}{2}\int_{\Gamma\times\mathcal{T}} \mathrm{d}\left(\vartheta^+_\mathsf{P} + \vartheta^-_\mathsf{P} - \bar{\vartheta}^{+,n}_\mathsf{P} - \bar{\vartheta}^{-,n}_\mathsf{P}\right),$$

for any dominating measure Σ, and similarly,

$$\mathcal{E}\mathrm{nt}\left(\mathsf{J}_t^{\pm}|\Theta_{\mathsf{P}}^{+}\right) = \int_{\Gamma\times\mathcal{T}} \Upsilon\left(\frac{\mathrm{d}\vartheta^{\pm}_{\mathsf{P}_t}}{\mathrm{d}\Sigma}, \frac{\mathrm{d}\vartheta^{\pm}_{\mathsf{P}_t}}{\mathrm{d}\Sigma}, \frac{\mathrm{d}\vartheta^{\mp}_{\mathsf{P}_t}}{\mathrm{d}\Sigma}\right)\mathrm{d}\Sigma,$$

$$\mathcal{D}(\mathsf{P}_t) = H^2(\vartheta^+_{\mathsf{P}}, \bar{\vartheta}^+_{\mathsf{P}}) + H^2(\vartheta^-_{\mathsf{P}}, \bar{\vartheta}^-_{\mathsf{P}}) + \frac{1}{2}\int_{\Gamma\times\mathcal{T}} \mathrm{d}\left(\vartheta^+_{\mathsf{P}} + \vartheta^-_{\mathsf{P}} - \bar{\vartheta}^+_{\mathsf{P}} - \bar{\vartheta}^-_{\mathsf{P}}\right).$$

By the moment estimates, narrow convexity and lower semicontinuity of Υ and H we conclude by standard semicontinuity results (e.g. see [But98, Theorem 3.4.3]),

$$\liminf_{n\to\infty} \int_0^T \mathcal{R}_n(\mathsf{P}^n_t, \mathsf{J}^{n,+}_t, \mathsf{J}^{n,-}_t)\,\mathrm{d}t \geq \int_0^T \mathcal{R}_\infty(\mathsf{P}_t, \mathsf{J}^+_t, \mathsf{J}^-_t)\,\mathrm{d}t, \quad \liminf_{n\to\infty} \mathcal{D}_n(\mathsf{P}^n_t) \geq \mathcal{D}_n(\mathsf{P}_t),$$

with the latter inequality holding for each $t \in [0, T]$, and from which (19.5) directly follows after applying Fatou's lemma.

Next, we consider the question of compactness. As in the previous part, let us consider a sequence $(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) \in \mathsf{CE}^*_n$ with

$$\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_0) < \infty, \qquad \limsup_{n\to\infty} \mathcal{J}_n(\mathsf{P}^n, \mathsf{J}^{n,+}, \mathsf{J}^{n,-}) < \infty,$$

and the moment bounds (19.13), which imply that the estimates (19.14) and (19.15) still hold. The bound on the free energy ensures by Theorem 19.8 that $\{\mathsf{P}^n_t\}_{t\in[0,T],n\geq n^*}$ for n^* large enough is pre-compact. Moreover, due to the bound on the action $\mathcal{R}_n$, we have by the results of Corollary (19.13) and Lemma (19.14) that

$$\limsup_{n\to\infty} \int_0^T \tilde{\phi}\left(\frac{1}{2M}\int_{\Gamma\times\mathcal{T}} (1 + v(\mathcal{T})^2)^{-1}\, \mathsf{J}^{n,\pm}_t(\mathrm{d}v, \mathrm{d}x)\right)\mathrm{d}t < \infty, \tag{19.16}$$

$$\limsup_{n\to\infty} \int_0^T \tilde{\phi}\left(\frac{|\dot{\mathsf{P}}^n_t|_W}{8M}\right)\mathrm{d}t < \infty, \tag{19.17}$$

where $|\dot{\mathsf{P}}^n_t|_W$ is again the W-metric speed. From (19.16), we then conclude from the non-decreasing, convex and super-linear at infinity property of $\tilde{\phi}$ that, up to choosing a subsequence n', there exists a family $\{\mathsf{J}^{\pm}_t\}_{t\in[0,T]} \in \mathcal{M}^+_{\mathrm{loc}}(\Gamma\times\mathcal{T})$ such that for all s, t the sequence of measures $\mathsf{J}^{n',\pm}_r(\mathrm{d}v, \mathrm{d}x)\,\mathrm{d}r$ converges to $\mathsf{J}^{\pm}_r(\mathrm{d}v, \mathrm{d}x)\,\mathrm{d}r$ in $\mathcal{M}_{\mathrm{loc}}(\Gamma\times\mathcal{T}\times[s,t])$, and

$$\int_0^T \int_{\Gamma\times\mathcal{T}} \tilde{\phi}\left(\frac{1}{2M}\int_{\Gamma\times\mathcal{T}} (1 + v(\mathcal{T})^2)^{-1}\, \mathsf{J}^{\pm}_t(\mathrm{d}v, \mathrm{d}x)\right)\mathrm{d}t < \infty.$$

Similarly, since the metric W is narrowly lower semicontinuous and induces narrow convergence on narrowly pre-compact subsets, we find by an Arzelà-Ascoli argument and the estimate (19.17) that, up to choosing a subsequence n'', there exists a narrowly continuous curve $(\mathsf{P}_t)_{t\in[0,T]}$ such that $\mathsf{P}_t^{n''}$ converges to P_t for all $t \in [0,T]$.

All that remains is showing that $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) \in \mathsf{CE}_\infty$. Therefore, fix any $s,t \in [0,T]$ and $F \in \mathrm{Cyl}_c(\Gamma)$. It is straightforward to verify that there exist constants K_F and C_F such that the following Taylor approximation holds:

$$\left|\mathrm{grad}_\Gamma(\nu,x) \mp n\left(F(\nu\pm\tfrac{1}{n}\delta_x) - F(\nu)\right)\right| \le \tfrac{C_F}{n} 1_{\nu(\mathcal{T})\le K_F}(\nu,x), \qquad \text{for all } \nu \in \Gamma,\ x \in \mathcal{T}.$$

Thus, we can take the limit in the continuity equation CE_n, to conclude that

$$\begin{aligned}
\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s &= \lim_{n\to\infty}\int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_t^{n''} - \int_\Gamma F(\nu)\,\mathrm{d}\mathsf{P}_s^{n''} \\
&= \lim_{n\to\infty}\int_s^t \left(\int_{\Gamma\times\mathcal{T}} (\overline{\nabla}^{n'',+}F)\,\mathrm{d}\mathsf{J}_r^{n'',+} + (\overline{\nabla}^{n'',-}F)\,\mathrm{d}\mathsf{J}_r^{n'',-}\right)\mathrm{d}r \\
&= \int_s^t \left(\int_{\Gamma\times\mathcal{T}} (\mathrm{grad}_\Gamma F)\,\mathrm{d}\mathsf{J}_r^+ - (\mathrm{grad}_\Gamma F)\,\mathrm{d}\mathsf{J}_r^-\right)\mathrm{d}r.
\end{aligned}$$

Due to the moment estimates and Corollary 19.13 one can now show that in fact $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) \in \mathsf{CE}_\infty^*$.

Now, consider the case of detailed balance. Note that due to the nonnegativity of the Fisher information $\mathcal{D}_n$ and $\mathcal{D}$ one can do away with the $(2+\varepsilon)$-moment bounds, and repeat the above steps, taking into consideration Theorems 17.11 and 18.10. To conclude that $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) \in \mathsf{CE}_\infty$ one requires the inequality (18.4), which will follow if the superposition principle holds.

However, note that either $(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-) \in \mathsf{CE}_\infty^*$, or in the case of detailed balance:

$$\int_0^T \mathcal{R}_\infty(\mathsf{P}_t,\mathsf{J}_t^+,\mathsf{J}_t^-)\,\mathrm{d}t < \infty,$$

and in both cases the superposition principle follows from Theorem 18.18 with the modification of Remark 18.19. □

Proof of Theorem 19.7. Suppose that $\bar{\mathsf{P}}^n \to \bar{\mathsf{P}} = \delta_{\bar\nu}$ with

$$\lim_{n\to\infty}\mathcal{F}_n(\bar{\mathsf{P}}^n) = \frac{1}{2}\mathcal{E}\mathrm{nt}(\nu|\gamma).$$

For each $n \in \mathbb{N}$ let P_t^n be the unique gradient-flow solution to (FKE_n) with initial datum $\bar{\mathsf{P}}^n$. Moreover, let ν_t be the unique solution to (16.11) with initial datum $\bar\nu$, and set $\mathsf{P}_t := \delta_{\nu_t}$, which is the unique gradient-flow solution to the Liouville equation (Li) with initial datum $\bar{\mathsf{P}}$. Then by Theorem 19.6 we have for every $t \in [0, T]$ that $\mathsf{P}^n_t \to \mathsf{P}_t$, and

$$\lim_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n_t) = \mathcal{F}_\infty(\mathsf{P}_t) = \frac{1}{2}\mathcal{E}\mathrm{nt}(\nu_t|\gamma).$$

Next, suppose that in addition there exists a constant $C > 1$ such that $C^{-1} \leq \mathrm{d}\bar{\nu}/\mathrm{d}\gamma \leq C$. By Lemma 16.16 there exists a $C' < \infty$ with

$$\sup_{t\in[0,T]} \left\|\log u_t\right\|_\infty < C', \qquad u_t := \mathrm{d}\nu_t/\mathrm{d}\gamma.$$

Now fix any $t \in [0, T]$, and recall that

$$\Pi_{n,\nu_t} := (L_n)_\# \pi_{n,\nu_t}, \qquad \pi_{n,\nu_t} = \frac{1}{e^{n\bar{\nu}(\mathcal{T})}} \sum_{N=0}^{\infty} \frac{n^N}{N!} \nu_t^{\otimes N}.$$

It is straightforward to check that $\Pi_n \ll \Pi_{n,\nu_t} \ll \Pi_n$ and hence for any $\Gamma_n \ni \nu = L_n(x_1, \dots, x_N)$,

$$\log\left(\frac{\mathrm{d}\Pi_{n,\nu_t}}{\mathrm{d}\Pi_n}\right)(\nu) = \log\left(\frac{e^{n\gamma(\mathcal{T})}}{e^{n\nu_t(\mathcal{T})}} \frac{\mathrm{d}\nu_t^{\otimes N}}{\mathrm{d}\gamma^{\otimes N}}\right) = \log\left(\frac{e^{n\gamma(\mathcal{T})}}{e^{n\nu_t(\mathcal{T})}}\right) + \sum_{i=1}^{N} \log u_t(x_i),$$

with all terms finite, and $|\sum \log u_t(x_i)| \leq NC'$. Therefore, by Lemma 19.11 we derive

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \int_\Gamma \log\left(\frac{\mathrm{d}\Pi_{n,\nu_t}}{\mathrm{d}\Pi_n}\right) \mathrm{d}\mathsf{P}^n_t &= \lim_{n\to\infty} \frac{1}{n} \log\left(\frac{e^{n\gamma(\mathcal{T})}}{e^{n\nu_t(\mathcal{T})}}\right) + \lim_{n\to\infty} \int_\Gamma \langle \log u_t, \nu \rangle \,\mathrm{d}\mathsf{P}^n_t \\
&= \gamma(\mathcal{T}) - \nu_t(\mathcal{T}) + \langle \log u_t, \nu_t \rangle \\
&= \mathcal{E}\mathrm{nt}(\nu_t|\gamma).
\end{aligned}$$

Subsequently, we can compute as follows:

$$\begin{aligned}
\lim_{n\to\infty} \mathcal{E}\mathrm{nt}(\mathsf{P}^n_0|\Pi_n) &= \frac{1}{n} \int_\Gamma \phi\left(\frac{\mathrm{d}\mathsf{P}^n_0}{\mathrm{d}\Pi_n}\right) \mathrm{d}\Pi_n \\
&= \lim_{n\to\infty} \frac{1}{n} \int_\Gamma \left(\log\left(\frac{\mathrm{d}\mathsf{P}^n_0}{\mathrm{d}\Pi_{n,\nu_0}}\right) + \log\left(\frac{\mathrm{d}\Pi_{n,\nu_0}}{\mathrm{d}\Pi_n}\right)\right) \mathrm{d}\mathsf{P}^n_0 \\
&= \mathcal{E}\mathrm{nt}(\nu_0|\gamma),
\end{aligned}$$

and hence the initial data are well-prepared. We can now conclude for all $t \in [0,T]$

$$\begin{aligned}
\lim_{n\to\infty} \mathcal{E}\mathrm{nt}(\mathsf{P}^n_t|\Pi_{n,\nu_t}) &= \frac{1}{n}\int_\Gamma \phi\left(\frac{\mathrm{d}\mathsf{P}^n_t}{\mathrm{d}\Pi_{n,\nu_t}}\right)\mathrm{d}\Pi_{n,\nu_t} \\
&= \lim_{n\to\infty}\frac{1}{n}\int_\Gamma \left(\log\left(\frac{\mathrm{d}\mathsf{P}^n_t}{\mathrm{d}\Pi_n}\right) + \log\left(\frac{\mathrm{d}\Pi_n}{\mathrm{d}\Pi_{n,\nu_t}}\right)\right)\mathrm{d}\mathsf{P}^n_t \\
&= \mathcal{E}\mathrm{nt}(\nu_t|\gamma) - \mathcal{E}\mathrm{nt}(\nu_t|\gamma) = 0,
\end{aligned}$$

thus establishing the entropic propagation of chaos result. □

Appendix E

Superposition principle in $\mathbb{R}^{\mathbb{N}}$

Here we present a superposition principle for continuity equations over $\mathbb{R}^{\mathbb{N}}$ with an additional weighted integrability condition on the associated vector fields.

Following [AT14, Section 7], we equip $\mathbb{R}^{\mathbb{N}}$ with the product topology, and $\pi_n := (p_1, \dots, p_n)$ the canonical projections. The space $AC_w([0,T];\mathbb{R}^{\mathbb{N}})$ consists of curves η such that $p_i \circ \eta \in AC[0,T]$ for all $i \in \mathbb{N}$. Note that both $\mathbb{R}^{\mathbb{N}}$ and $C([0,T];\mathbb{R}^{\mathbb{N}})$ are Polish spaces. Moreover, let $|\cdot|_\infty$ be the uniform norm on $\mathbb{R}^{\mathbb{N}}$.

Smooth n-cylindrical functions with compact support $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ are given in the form of

$$f(x) = g(\pi_n(x)) = g(p_1(x), \dots, p_n(x)), \qquad x \in \mathbb{R}^{\mathbb{N}},$$

with $g \in C_c^\infty(\mathbb{R}^n \to \mathbb{R})$, and define their gradient by

$$\nabla f(x) := \left(\frac{\partial g}{\partial z_1}(\pi_n(x)), \dots, \frac{\partial g}{\partial z_n}(\pi_n(x)), 0, 0, \dots \right).$$

We set $\mathrm{Cyl}_c(\mathbb{R}^{\mathbb{N}})$ as the union over $n \in \mathbb{N}$ of all smooth n-cylindrical functions with compact support.

In the following, we consider pairs $(\nu, \boldsymbol{c})$, where $(\nu_t)_{t\in[0,T]} \subset \mathcal{P}(\mathbb{R}^{\mathbb{N}})$ is a weakly continuous family of probability measures and $\boldsymbol{c} : [0,T] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ is a Borel vector field satisfying

$$\int_{\mathbb{R}^{\mathbb{N}}} f \,\mathrm{d}\nu_t - \int_{\mathbb{R}^{\mathbb{N}}} f \,\mathrm{d}\nu_s = \int_s^t \int_{\mathbb{R}^{\mathbb{N}}} (\boldsymbol{c}_r, \nabla f) \,\mathrm{d}\nu_r \,\mathrm{d}r \qquad \text{for all } f \in \mathrm{Cyl}_c(\mathbb{R}^{\mathbb{N}}),$$

and all $0 \le s \le t \le T$.

We then have the following result.

Theorem E.1. *Let $(\nu, \boldsymbol{c})$ be as above. Furthermore, suppose that*

$$\int_0^T \int_{\mathbb{R}^{\mathbb{N}}} \frac{|\boldsymbol{c}_t|_\infty}{1+|x_1|} \, \mathrm{d}\nu_t \, \mathrm{d}t < \infty.$$

Then there exists a Borel probability measure λ over $\boldsymbol{C}([0,T];\mathbb{R}^{\mathbb{N}})$ satisfying $(e_t)_\# \lambda = \nu_t$ for all $t \in [0,T]$, and is concentrated on the family of curves $\gamma \in AC([0,T];\mathbb{R}^{\mathbb{N}})$ that satisfy

$$\dot{\gamma} = \boldsymbol{c}_t(\gamma), \qquad \text{for almost every } t \in [0,T].$$

The proof of Theorem E.1 combines a slight adaptation of the proof for the superposition principle in $\mathbb{R}^{\mathbb{N}}$ found in [AT14, Theorem 7.1], developed for use in metric measure spaces, with a finite-dimensional result for vector fields over $\mathbb{R}^n$ found in [AC08, Theorem 4.4]. Due to the strong similarities with the proof found in [AT14], we merely give a brief sketch.

Proof. By tightness of ν_0, we can choose a sequence of coercive functionals Φ_i such that

$$\int \Phi_i(p_i(x)) \, \mathrm{d}\nu_0 \le 2^{-i}, \qquad \text{for all } i \in \mathbb{N},$$

and by $\nu_t(\mathrm{d}x)\mathrm{d}t$-integrability of $\frac{|\boldsymbol{c}_t|_\infty}{1+|x_1|}$ we can find a superlinear, convex and monotone function $\bar{\Phi}$ such that

$$\int_0^T \int_{\mathbb{R}^{\mathbb{N}}} \bar{\Phi}\left(\frac{|\boldsymbol{c}_t|_\infty}{1+|x_1|} \right) \mathrm{d}\nu_t \, \mathrm{d}t < \infty.$$

We now consider the functional $\mathcal{A}(\eta) : \boldsymbol{C}([0,T];\mathbb{R}^{\mathbb{N}}) \to [0,+\infty]$ given by

$$\mathcal{A}(\eta) := \begin{cases} \displaystyle\sum_{i=1}^\infty \left(\Phi_i(p_i \circ \eta(0)) + \int_0^T \bar{\Phi}\left(\frac{|\dot{\eta}(t)|_\infty}{1+|p_1 \circ \eta(t)|} \right) \mathrm{d}t \right) & \text{if } \eta \in AC_w([0,T];\mathbb{R}^{\mathbb{N}}), \\ +\infty & \text{otherwise.} \end{cases}$$

It is clear that $\mathcal{A}$ is coercive in $\boldsymbol{C}([0,T];\mathbb{R}^{\mathbb{N}})$, and its sublevel sets contain curves that are absolutely continuous with respect to $|\cdot|_\infty$. This follows from the fact that $\sup_{t\in[0,T]} |p_1 \circ \eta|$ is bounded on the sublevel sets of the functional

$$\left(\Phi_1(p_1 \circ \eta(0)) + \int_0^T \bar{\Phi}\left(\frac{|(p_1 \circ \eta)'(t)|}{M(1+|p_1 \circ \eta(t)|)} \right) \mathrm{d}t \right).$$

Now, for every $n \in \mathbb{N}$, we define the marginals $\mathcal{P}(\mathbb{R}^n) \ni \nu_t^n := (\pi_n)_\# \nu_t$ and corresponding vector fields $\boldsymbol{c}_t^n : \mathbb{R}^n \to \mathbb{R}^n$ by

$$p_i \circ \boldsymbol{c}_t^n := \frac{\mathrm{d}\,(\pi_n)_\# \left((p_i \circ \boldsymbol{c}_t) \, \nu_t \right)}{\mathrm{d}\nu_t^n}.$$

Note that $(\nu^n, \boldsymbol{c}_t^n)$ satisfies the continuity equation in $\mathbb{R}^n$. By the fact that $|z_1| \le |z| \le n|z|_\infty$, for $z = (x_1, \dots, x_n) \in \mathbb{R}^n$, we have that

$$\int_0^T \int \frac{|\boldsymbol{c}_t^n|}{1+|x|} \, \mathrm{d}\nu_t \, \mathrm{d}t \le n \int_0^T \int \frac{|\boldsymbol{c}_t^n|_\infty}{1+|x_1|} \, \mathrm{d}\nu_t \, \mathrm{d}t < \infty.$$

Hence, we can apply the finite-dimensional version of [AC08, Theorem 4.4]. Embedding this into $\mathbb{R}^{\mathbb{N}}$, we obtain the probability measure λ_n over $C([0,T], \mathbb{R}^{\mathbb{N}})$, concentrated on absolutely continuous curves satisfying $\dot\gamma = \boldsymbol{c}_t^n(\gamma)$, and such that $(e_t)_\# \lambda = \nu_t^n$. We immediately see that

$$\sup_{n \in \mathbb{N}} \int \mathcal{A}(\gamma) \, \mathrm{d}\lambda_n(\gamma) < \infty,$$

which yields the tightness of λ^n.

Consider any converging sequence λ^n (up to renumbering) and its limit $\lambda \in \mathcal{P}(C([0,T]; \mathbb{R}^{\mathbb{N}}))$. Since the sequence $(\nu_t^n)_{n\in\mathbb{N}}$ clearly converges to $\nu_t := (e_t)_\# \lambda$ in $\mathcal{P}(\mathbb{R}^{\mathbb{N}})$ for every $t \in [0,T]$, it remains to show that λ is concentrated on solutions of $\dot\gamma = \boldsymbol{c}_t(\gamma)$. In fact we will show for each $i \in \mathbb{N}$ and any $0 \le s \le t \le T$ that

$$\int \frac{\left| p_i \circ \gamma(t) - p_i \circ \gamma(s) - \int_s^t p_i \circ \boldsymbol{c}_r(\gamma(r)) \, \mathrm{d}r \right|}{1 + \|p_1 \circ \gamma\|_\infty} \lambda(\mathrm{d}\gamma) = 0$$

Note that it suffices to show that for any vector field $d : [0,T] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ with d_t being k-cylindrical for every $t \in [0,T]$, we have that

$$\int \frac{\left| p_i \circ \gamma(t) - p_i \circ \gamma(s) - \int_s^t d_r(\gamma(r)) \, \mathrm{d}r \right|}{1 + \|p_1 \circ \gamma\|_\infty} \lambda(\mathrm{d}\gamma) \le \int_s^t \int_{\mathbb{R}^{\mathbb{N}}} \frac{|p_i \circ \boldsymbol{c}_r - d_r|}{1+|x_1|} \, \mathrm{d}\nu_r \, \mathrm{d}r, \tag{E.1}$$

since then we can use density of time-dependent cylindrical functions in $L^1((1+|x_1|)^{-1} \nu_s \, \mathrm{d}s)$ and the fact that for all s it holds that $|p_1 \circ \gamma(s)| \le \|p_1 \circ \gamma\|_\infty$.

To prove (E.1), recall that λ^n is concentrated on absolutely continuous solutions of $\dot\gamma_s = \boldsymbol{c}_s^n(\gamma_s)$. Hence,

$$\begin{aligned}
\int \frac{\left| p_i \circ \gamma(t) - p_i \circ \gamma(0) - \int_0^t d_s(\gamma(s)) \, \mathrm{d}s \right|}{1 + \|p_1 \circ \gamma\|_\infty} \lambda^n(\mathrm{d}\gamma) &= \int \frac{\left| \int_0^t (p_i \circ \boldsymbol{c}_s^n(\gamma(s)) - d_s(\gamma(s))) \, \mathrm{d}s \right|}{1 + \|p_1 \circ \gamma\|_\infty} \lambda^n(\mathrm{d}\gamma) \\
&\le \int \frac{\int_0^t |p_i \circ \boldsymbol{c}_s^n - d_s|(\gamma(s)) \, \mathrm{d}s}{1 + |p_1(\gamma(s))|} \lambda^n(\mathrm{d}\gamma) \\
&\le \int_0^t \int_{\mathbb{R}^{\mathbb{N}}} \frac{|p_i \circ \boldsymbol{c}_s^n - d_s|}{1+|x_1|} \, \mathrm{d}\nu_s^n \, \mathrm{d}s
\end{aligned}$$

Note that the integrand on the left-hand side is continuous in γ. Therefore, since for $n \geq k$

$$(p_i \circ \boldsymbol{c}_s^n - d_s)\nu_s^n = (\pi_n)_\#((p_i \circ \boldsymbol{c}_s - d_s)\nu_s),$$

the result then follows after taking the limit $n \to \infty$. □

Remark E.2. If one is only interested in curves in $AC_w([0,T];\mathbb{R}^{\mathbb{N}})$, the theorem also holds whenever

$$\int_0^T \int \frac{|p_i(\boldsymbol{c}_t)|}{1+|x_1|} \, \mathrm{d}\nu_t \, \mathrm{d}t < \infty, \qquad \text{for all } i \in \mathbb{N}.$$

The finite dimensional analog of this statement, set in $\mathbb{R}^n$ with the prefactor $(1 + |x|)^{-1}$, is presented in [AC08, Theorem 4.4]. Moreover, for $\mathbb{R}^{\mathbb{N}}$, in [AT14, Theorem 7.1] the condition reads as

$$\int_0^T \int |p_i(\boldsymbol{c}_t)| \, \mathrm{d}\nu_t \, \mathrm{d}t < \infty, \qquad \text{for all } i \in \mathbb{N}.$$

□

Appendix F

Non-continuous rates

In the proof of Theorem 19.2 we require the vague convergence of $\vartheta^{\pm}_{\mathsf{P}^n}$ and $\mathsf{T}^{n,\mp}_{\#}\bar{\vartheta}^{\pm}_{\mathsf{P}^n}$ under the assumption of narrow convergence of P^n and equiboundedness of the free energy functionals. If the mutation/competition kernels $m^{\pm}$ are continuous, the desired statement would follow directly from the narrow convergence of P^n.

The case of merely bounded measurable $m^{\pm}$ is however less trivial. Note that we do not have setwise convergence of P^n or $\vartheta^{\pm}_{\mathsf{P}^n}$, since although for every fixed n the sublevels of $\mathcal{F}_n$ are sequentially compact with respect to setwise convergence, this is not the case for equibounded sets of $\{\mathcal{F}_n\}_{n\geq 1}$.

Fortunately, due to the connection between Γ-convergence of $\mathcal{F}_n$ and large deviations as discussed in Appendix G, we can modify results from Part I.A on large deviations for interacting systems induced by singular or irregular functionals, specifically Chapter 4. In particular, we obtain the following convergence statement.

Theorem F.1. *Let* $\{\mathsf{P}^n\}_{n\geq 1} \subset \mathcal{P}(\Gamma)$ *be a sequence narrowly converging to* $\mathsf{P} \in \mathcal{P}(\Gamma)$ *with*

$$\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n) < \infty,$$

and let $b^{\pm}$ *satisfy Assumption 19.1. Then we have the vague convergence*

$$\lim_{n\to\infty} \vartheta^{\pm}_{\mathsf{P}^n_t} = \vartheta^{\pm}_{\mathsf{P}_t}, \qquad \lim_{n\to\infty} \mathsf{T}^{n,\mp}_{\#}\bar{\vartheta}^{n,\pm}_{\mathsf{P}^n_t} = \bar{\vartheta}^{\pm}_{\mathsf{P}_t}, \qquad \lim_{n\to\infty} \bar{\vartheta}^{n,\pm}_{\mathsf{P}^n_t} = \bar{\vartheta}^{\pm}_{\mathsf{P}_t}.$$

For convenience, recall that Assumption 19.1 states

$$b^{\pm}[\nu](x) = \psi^{\pm}\left(x, \nu, \int_{\mathcal{T}} m^{\pm}(x,y)\,\mathrm{d}\nu(y)\right).$$

where $\psi^{\pm} \in C(\mathcal{T}\times\Gamma\times\mathbb{R}_{\geq 0})$ are nonnegative continuous functions with the property that: for every $K > 0$ there exists a constant L'_K such that

$$|\psi^{\pm}(x,\mu,z) - \psi^{\pm}(x,\nu,z')| \leq L'_K\left(\|\mu-\nu\|_{TV} + |z-z'|\right), \tag{F.1}$$

for all $x \in \mathcal{T}$, $\mu, \nu \in \Gamma$ and $z \in \mathbb{R}_{\geq 0}$ such that $\max(\mu(\mathcal{T}), \nu(\mathcal{T}), z) \leq K$; there exists a constant $\tilde{M}$ such that

$$\psi^{\pm}(x, \nu, z) \leq \tilde{M}(\nu(\mathcal{T}) + |z|); \tag{F.2}$$

and $m^{\pm} \in B_b^+(\mathcal{T}^2)$.

Remark F.2. In particular, the following holds

$$\lim_{n\to\infty} \int_{\Gamma\times\mathcal{T}} g(x,y)\omega(\nu,x)\nu(\mathrm{d}x)\nu(\mathrm{d}y)\mathsf{P}^n(\mathrm{d}\nu) = \int_{\Gamma\times\mathcal{T}} g(x,y)\omega(\nu,x)\nu(\mathrm{d}x)\nu(\mathrm{d}y)\mathsf{P}(\mathrm{d}\nu),$$
$$\lim_{n\to\infty} \int_{\Gamma\times\mathcal{T}} g(x,y)\omega(\nu,x)\gamma(\mathrm{d}x)\nu(\mathrm{d}y)\mathsf{P}^n(\mathrm{d}\nu) = \int_{\Gamma\times\mathcal{T}} g(x,y)\omega(\nu,x)\gamma(\mathrm{d}x)\nu(\mathrm{d}y)\mathsf{P}(\mathrm{d}\nu),$$

for any $\omega \in C_c(\Gamma \times \mathcal{T})$ and $g \in \mathcal{B}_b(\mathcal{T}^2)$. The result can be easily generalized to bounded measurable functions $g \in \mathcal{B}_b(\mathcal{T}^k)$ for finite $k \in \mathbb{N}$ or n-dependence, but we restrict ourselves to this simple case. □

For the proof of Theorem (F.1) we will need some a priori bounds. Namely, recall from Section 19.1 the generating functionals and their limit

$$G_n(f) := \frac{1}{n}\log\int_\Gamma e^{n\langle f,\nu\rangle}\,\Pi_n(\mathrm{d}\nu), \quad G(f) := \int_{\mathcal{T}} (e^f - 1)\,\mathrm{d}\gamma.$$

For the "interacting" case, namely functionals of the form

$$\frac{1}{n}\log\int_\Gamma e^{n\langle g,\nu^{\otimes 2}\rangle}\,\Pi_n(\mathrm{d}\nu),$$

there is however a problem with the unboundedness of the mass of ν. Nevertheless, upon controlling the mass we can provide the following technical estimate.

Lemma F.3. *Let* $F(\nu) := h(\nu(\mathcal{T}))\langle g, \nu^{\otimes 2}\rangle$ *with* $\operatorname{supp}(h) \in [0, K]$ *and* $g \in \mathcal{B}_b(\mathcal{T}^2)$. *Then*

$$\tfrac{1}{n}\log\int_\Gamma e^{n|F|}\,\mathrm{d}\Pi_n \leq \left(\int_{\mathcal{T}^2} e^{4K\|h\|_\infty |g|(x,y)}\mathrm{d}\gamma^{\otimes 2}\right)^{1/2} + \frac{1}{n}\left(K\|g\|_\infty\|h\|_\infty\right) - \gamma(\mathcal{T}) \tag{F.3}$$

and, in particular,

$$\limsup_{n\to\infty}\frac{1}{n}\log\int_\Gamma e^{n|F|}\,\mathrm{d}\Pi_n \leq \left(\int_{\mathcal{T}^2} e^{4K\|h\|_\infty |g|(x,y)}\mathrm{d}\gamma^{\otimes 2}\right)^{1/2}.$$

Proof. Suppose that

$$\limsup_{n\to\infty} \mathcal{F}_n(\mathsf{P}^n) =: C < \infty,$$

and let us consider the following interaction energy functional:

$$E_g^N(x_1, \dots, x_N) := \frac{1}{N^2} \sum_{i,j\neq i} |g|(x_i, x_j).$$

From a Hoeffding's decomposition argument, see Lemma 4.10 of Part I.A, we have for every $N \geq 2$, $\alpha \geq 0$ the estimate

$$\frac{1}{N} \log \frac{1}{\gamma(\mathcal{T})^N} \int_{\mathcal{T}^N} e^{\alpha N E_g^N(x_1,\dots,x_N)}\, \mathrm{d}\gamma^{\otimes N} \leq \tfrac{1}{2} \log \left(\frac{1}{\gamma(\mathcal{T})^2} \int_{\mathcal{T}\times\mathcal{T}} e^{\frac{2\alpha N}{N-1}|g|(x,y)}\, \mathrm{d}\gamma^{\otimes 2} \right).$$

Moreover, since $N/(N-1) \leq 2$ for $N \geq 2$, and

$$\sum_{i,j} |g|(x_i, x_j) = \sum_{i,j\neq i} |g|(x_i, x_j) + \sum_{i} |g|(x_i, x_i) \leq \sum_{i,j\neq i} |g|(x_i, x_j) + N\|g\|_\infty,$$

we find that

$$\frac{1}{N} \log \left(\frac{1}{\gamma(\mathcal{T})^N} \int_{\mathcal{T}^N} e^{\frac{\alpha}{N} \sum_{i,j} |g|(x_i,x_j)}\, \mathrm{d}\gamma^{\otimes N} \right) \leq \frac{1}{2} \log \left(\frac{1}{\gamma(\mathcal{T})^2} \int_{\mathcal{T}\times\mathcal{T}} e^{4\alpha|g|(x,y)}\, \mathrm{d}\gamma^{\otimes 2} \right) + \frac{\alpha\|g\|_\infty}{N}.$$

Recall that $L_n(x_1, \dots, x_N) := \frac{1}{n} \sum \delta_{x_i}$. Since the mass $L_n(x_1, \dots, x_N)(\mathcal{T}) = N/n$ is bounded by K on the support of F we have for $N \geq 2$:

$$|F|(L_n) \leq |h|(L_n(\mathcal{T})) \tfrac{1}{n^2} \sum_{i,j} |g|(x_i, x_j) \leq \frac{K\|h\|_\infty}{nN} \sum_{i,j} |g|(x_i, x_j),$$

while for $N = 1$ we have the trivial estimate $|F|(L_n) \leq \frac{K}{n}\|h\|_\infty \|g\|_\infty$, and hence for all $N \geq 0$,

$$\frac{1}{\gamma(\mathcal{T})^N} \int_{\mathcal{T}^N} e^{n|F|(L_N)}\, \mathrm{d}\gamma^{\otimes N} \leq e^{K\|g\|_\infty \|h\|_\infty} \left(\frac{1}{\gamma(\mathcal{T})^2} \int_{\mathcal{T}^2} e^{4K\|h\|_\infty |g|(x,y)} \mathrm{d}\gamma^{\otimes 2} \right)^{N/2}.$$

Using the representation for Π_n we can therefore estimate

$$\begin{aligned}
\int_\Gamma e^{n|F|}\, \mathrm{d}\Pi_n &= \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^\infty \frac{(n\gamma(\mathcal{T}))^N}{N!} \int_{\mathcal{T}^N} e^{n|F|} \mathrm{d}\gamma^{\otimes N} / \gamma(\mathcal{T})^N \\
&\leq \frac{1}{e^{n\gamma(\mathcal{T})}} \sum_{N=0}^\infty \frac{(n\gamma(\mathcal{T}))^N}{N!} e^{K\|g\|_\infty \|h\|_\infty} \left(\frac{1}{\gamma(\mathcal{T})^2} \int_{\mathcal{T}^2} e^{4K\|h\|_\infty |g|(x,y)} \mathrm{d}\gamma^{\otimes 2} \right)^{N/2} \\
&= \frac{e^{K\|g\|_\infty \|h\|_\infty}}{e^{n\gamma(\mathcal{T})}} \exp\left\{ n\gamma(\mathcal{T}) \left(\frac{1}{\gamma(\mathcal{T})^2} \int_{\mathcal{T}^2} e^{4K\|h\|_\infty |g|(x,y)} \mathrm{d}\gamma^{\otimes 2} \right)^{1/2} \right\},
\end{aligned}$$

which proves (F.3). The final desired statement follows directly after taking limits. □

With the above estimate in hand, we can now prove our convergence statement by approximating any $m^{\pm} \in \mathcal{B}_b(\mathcal{T}^2)$ with a sequence of continuous g_ε such that

$$\lim_{\varepsilon\to 0} \int_{\mathcal{T}^2} e^{\beta|m^{\pm}-g_\varepsilon|(x,y)} \mathrm{d}\gamma^{\otimes 2} = 0, \qquad \text{for all } \beta > 0. \tag{F.4}$$

The existence of such a sequence follows similarly as for density statements in $L^p(\gamma)$, see for example Theorem C.5 of Part I.

Proof of Theorem F.1. (F.1) (F.2)

Fix a $m^{\pm} \in \mathcal{B}_b^+(\mathcal{T}^2)$. We will establish the vague convergence of $\vartheta^{\pm}_{\mathsf{P}^n}$ to $\vartheta^{\pm}_{\mathsf{P}^n}$, namely that for any $\omega \in C_c(\Gamma \times \mathcal{T})$

$$\begin{aligned}
&\lim_{n\to\infty} \int_{\Gamma\times\mathcal{T}} \omega(\nu, x)\psi^+\left(x, \nu, \int_{\mathcal{T}} m^+(x, y)\,\mathrm{d}\nu(y)\right) \gamma(\mathrm{d}x)\mathsf{P}^n(\mathrm{d}\nu) \\
&\quad = \int_{\Gamma\times\mathcal{T}} \omega(\nu, x)\psi^+\left(x, \nu, \int_{\mathcal{T}} m^+(x, y)\,\mathrm{d}\nu(y)\right) \gamma(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu), \\
&\lim_{n\to\infty} \int_{\Gamma\times\mathcal{T}} \omega(\nu, x)\psi^-\left(x, \nu, \int_{\mathcal{T}} m^-(x, y)\,\mathrm{d}\nu(y)\right) \nu(\mathrm{d}x)\mathsf{P}^n(\mathrm{d}\nu) \\
&\quad = \int_{\Gamma\times\mathcal{T}} \omega(\nu, x)\psi^-\left(x, \nu, \int_{\mathcal{T}} m^-(x, y)\,\mathrm{d}\nu(y)\right) \nu(\mathrm{d}x)\mathsf{P}(\mathrm{d}\nu).
\end{aligned} \tag{F.5}$$

By the uniform continuity and compact support of ω and the Lipschitz property of $\psi^{\pm}$ on bounded sets it is straightforward to show by an approximation argument that $\mathsf{T}^{n,\mp}_{\#}\bar{\vartheta}^{n,\pm}_{\mathsf{P}^n_t}$ and $\bar{\vartheta}^{n,\pm}_{\mathsf{P}^n_t}$ converge vaguely as well.

Now, let $\{m^{\pm}_\varepsilon\}_{\varepsilon>0} \subset C_b(\mathcal{T}^2)$ be a sequence approximating $m^{\pm}$ in the sense of (F.4). The narrow convergence of P^n implies that for any $\omega \in C_c(\Gamma \times \mathcal{T})$ and any $\varepsilon > 0$ we have the convergence of (F.5) with $m^{\pm}$ substituted by $m^{\pm}_\varepsilon$.

Again by the compact support of ω and the Lipschitz property of $\psi^{\pm}$ on bounded sets, it is clear that to show (F.5) we need to verify for every $K > 0$:

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} \int_{\nu(\mathcal{T})\le K} \left(\int_{\mathcal{T}^2} |m^- - m^-_\varepsilon|(x, y)\nu(\mathrm{d}x)\nu(\mathrm{d}y)\right) \mathsf{P}^n(\mathrm{d}\nu) = 0, \tag{F.6a}$$

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} \int_{\nu(\mathcal{T})\le K} \left(\int_{\mathcal{T}^2} |m^+ - m^+_\varepsilon|(x, y)\gamma(\mathrm{d}x)\nu(\mathrm{d}y)\right) \mathsf{P}^n(\mathrm{d}\nu) = 0. \tag{F.6b}$$

Let us consider (F.6a), and set

$$F_{\varepsilon,K}(\nu) := 1_{\nu(\mathcal{T})\le K} \int_{\mathcal{T}^2} |g - g_\varepsilon|(x, y)\nu(\mathrm{d}x)\nu(\mathrm{d}y).$$

By duality of the entropy and Lemma F.3, we have for every $n \geq 1, \varepsilon > 0, K > 0$, and $\beta > 0$,

$$\begin{aligned}\int_\Gamma F_{\varepsilon,K}(\nu)\mathsf{P}^n(\mathrm{d}\nu) &\leq \frac{1}{\beta n}\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) + \frac{1}{\beta n}\log\int_\Gamma e^{n\beta F_{\varepsilon,K}}\,\mathrm{d}\Pi_n \\ &\leq \frac{1}{\beta n}\mathcal{E}\mathrm{nt}(\mathsf{P}^n|\Pi_n) + \frac{1}{\beta}\left(\int_{\mathcal{T}^2} e^{4\beta K|m^- - m_\varepsilon^-|(x,y)}\mathrm{d}\gamma^{\otimes 2}\right)^{1/2} \\ &\qquad + \frac{1}{\beta n}\left(K\beta\|m - m_\varepsilon^-\|_\infty - \log(e^{n\gamma(\mathcal{T})} - 1)\right).\end{aligned}$$

Taking subsequently the limits $n \to \infty$ and $\varepsilon \to 0$, we deduce

$$\limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\int_\Gamma F_{\varepsilon,K}(\nu)\,\mathsf{P}^n(\mathrm{d}\nu) \leq \frac{C}{\beta}.$$

But, since $\beta > 0$ was arbitrary, we conclude that the right-hand side reduces to zero. Similarly, for (F.6b), let

$$f_\varepsilon(x) := \int_{\mathcal{T}} |m^+ - m_\varepsilon^+|(y,x)\,\gamma(\mathrm{d}y), \qquad F_\varepsilon(\nu) := \int_{\mathcal{T}} f_\varepsilon(x)\,\nu(\mathrm{d}x).$$

Then by duality and Lemma 19.10 we obtain

$$\begin{aligned}\limsup_{n\to\infty}\int_\Gamma F_{\varepsilon,K}(\nu)\mathsf{P}^n(\mathrm{d}\nu) &\leq \frac{C}{\beta} + \int_{\mathcal{T}}(e^{\beta f_\varepsilon(x)} - 1)\mathrm{d}\gamma \\ &\leq \frac{C}{\beta} + \frac{1}{\gamma(\mathcal{T})}\int_{\mathcal{T}^2} e^{\beta\gamma(\mathcal{T})|m^+ - m_\varepsilon^+|(x,y)}\mathrm{d}\gamma^{\otimes 2},\end{aligned}$$

where the last inequality follows by applying Jensen's inequality inside the exponential. Again taking the limit $\varepsilon \to 0$ and thereafter $\beta \to \infty$ we conclude the proof. □

Appendix G

Motivation from large deviations

We introduced a new variational and dissipation structure for the forward Kolmogorov equation and later showed convergence in the large-population limit to a structure that was lifted from the mean-field dynamics. Moreover, already in Part II.A the relation between existing variational structures for the FKE was discussed, and their connection to asymptotic probabilities non-interacting particle system as treated in large deviation theory. We now discuss the relation between the EDP-convergence of variational structures and their connection to large deviations. All calculations are purely formal and are meant for illustratory purposes.

Throughout, for simplicity, let $\mathcal{T}$ be a finite set. Recall the reacting particle system formulation described by (15.3), i.e. as particles with labels $A_t^1, \dots, A_t^{N_t}$ and traits $X_t^1, \dots, X_t^{N_t} \in \mathcal{T}$, and with

$$
\begin{aligned}
A_t^i &\to A_t^i + A_t^{N_t+1}, &\quad \text{with rate} \quad & m\left(X_t^i, X_t^{N_1+1}\right)\gamma\left(X_t^{N_1+1}\right), \\
A_t^i + A_t^j &\to A_t^j, &\quad \text{with rate} \quad & n^{-1}c\left(X_t^i, X_t^j\right).
\end{aligned}
$$

Let L_t^n be the rescaled empirical measure

$$
L_t^n(x) := \sum_{i=1}^{N_t} \delta_{X_t^i}(x),
$$

and $W_t^{n,\pm}$ the *integrated birth/death fluxes*:

$$
W_t^{n,\pm}(x) := \frac{1}{n}\#\Big\{\text{Number of births}(+)/\text{deaths}(-) \text{ with trait } x \text{ in the time-window } [0,t)\Big\}.
$$

Moreover, assume that the particles are initially distributed at time $t = 0$ as π_n. Then by the work of [PR19], one can derive under suitable assumptions that the triple $(L_t^n, W_t^{n,\pm})$ is a well-defined Markov process and satisfies a *large-deviation principle* as $n \to \infty$ with rate function $\mathcal{I}(\nu, \lambda^+, \lambda^+)$ in the sense that asymptotically (as $n \to \infty$)

$$\mathrm{Prob}\left(L_t^n \approx \nu_t,\ W_t^{n,\pm} \approx \int_0^t \lambda_s^\pm \,\mathrm{d}s,\ \forall t \in [0,T] \right) \asymp e^{-n\left(\mathcal{I}^0(\nu_0)+\mathcal{I}_{MF}(\nu,\lambda^+,\lambda^-)\right)}$$

where $\mathcal{I}^0(\nu) := \mathcal{E}\mathrm{nt}(\nu|\gamma)$ and $\mathcal{I}_{MF}$ the mean-field rate function from Chapter 16:

$$\mathcal{I}_{MF}(\nu,\lambda^+,\lambda^-) := \int_0^T \mathcal{L}_{MF}(\nu_t,\lambda_t^+,\lambda_t^-)\,\mathrm{d}t,\ \mathcal{L}_{MF}(\nu,\lambda^+,\lambda^-) := \mathcal{E}\mathrm{nt}(\lambda^+|\chi_\nu^+)+\mathcal{E}\mathrm{nt}(\lambda^-|\chi_\nu^-).$$

Moreover, recall Section 17.2 the relation between mass fluxes $\mathsf{J}^\pm$ and edge fluxes $j^\pm$, and the FKE-rate functional $\mathcal{I}(\mathsf{P}, j; \kappa_n)$ from Remark 17.18 given as

$$\mathcal{I}(\rho, j; \kappa_n) = n\mathcal{I}_n(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-).$$

Due to the origin of $\bar{\mathcal{I}}(\rho, j; \kappa_n)$ in large deviations for independent particles (or via variational representations as found in [DE97]), one would expect that if $F_t \in C_b(\Gamma)$ for all $t \in [0,T]$, we would have for all $n > 0$ the following representation formula for the expectation:

$$\frac{1}{n}\log\mathbb{E}\left[e^{-n\int_0^T F_t(L_t^n)\,\mathrm{d}t}\right] = \inf_{(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)}\left\{\int_0^T\int_\Gamma F_t(\nu_t)\mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}t + \frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}_0|\Pi_n) + \mathcal{I}_n(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\right\}.$$

On the other hand, by the large deviation principle of $(L_t^n, W_t^{n,\pm})$ as $n \to \infty$, and Varadhan's Lemma (see Theorem 2.1, Part I.A), it holds that

$$-\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[e^{-n\int_0^T F_t(L_t^n)\,\mathrm{d}t}\right] = \inf_{(\nu,\lambda^+,\lambda^-)}\left\{\int_0^T F_t(\nu_t)\,\mathrm{d}t + \mathcal{E}\mathrm{nt}(\nu_0|\gamma) + \mathcal{I}_{MF}(\nu,\lambda^+,\lambda^-)\right\}.$$

Consequently,

$$\begin{aligned}\lim_{n\to\infty}\inf_{(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)}&\left\{\int_0^T\int_\Gamma F_t(\nu_t)\mathsf{P}_t(\mathrm{d}\nu)\,\mathrm{d}t + \frac{1}{n}\mathcal{E}\mathrm{nt}(\mathsf{P}_0|\Pi_n) + \mathcal{I}_n(\mathsf{P},\mathsf{J}^+,\mathsf{J}^-)\right\}\\ &= \inf_{(\nu,\lambda^+,\lambda^-)}\left\{\int_0^T F_t(\nu_t)\,\mathrm{d}t + \mathcal{E}\mathrm{nt}(\nu_0|\gamma) + \mathcal{I}_{MF}(\nu,\lambda^+,\lambda^-)\right\}.\end{aligned}$$

Note that the lower bound of this equality follows from Theorem 19.2 and the superposition principle in Theorem 18.9. Moreover, we expect that the large-deviation principle implies evolutionary Γ-convergence of $\mathcal{I}_n$ in a suitable topology—an implication studied in [Kra19] in a general setting.

It then begs the question if one can reverse this procedure, namely using evolutionary Γ-convergence to establish large-deviation principles similar to the non-evolutionary setting of [Mar12]. This approach was successfully applied in the case of certain diffusion processes [Fat16] and discussed for more general processes in [KJZ19].

Part III

Summary and discussion

Discussion and open questions

We have established mean-field limits for interacting diffusions with singular interactions, in the form of almost sure convergence of the empirical measures over path space to the McKean–Vlasov equation, and for measure-dependent birth/death processes, in the form of concentration of the solution to the forward Kolmogorov equation (FKE) around the mean-field limit. In the case of the former, we even established a large deviation principle (LDP), and for the latter, we showed convergence of variational structures in terms of Energy-Dissipation Principles (EDPs).

Now, let us review our results and revisit some of the underlying open questions.

I.A Large deviations for singular functionals and Gibbs measures

Here we presented an extension of Varadhan's Integral Lemma, which allowed us to prove LDPs for systems that are tilted by a (possibly) singular energy functional, where we in essence used an approximation and convexity argument. This was applied to interacting particle systems that are characterized by an interaction potential V and an underlying non-interacting particle system, with the laws of the interacting particle system forming Gibbs measures. Moreover, we generalized these arguments to Gibbs-like systems, to apply this to interacting diffusions of Part I.B.

The main restriction was our assumption of *strong exponential integrability* of the interaction potential V. As we discussed, one can go beyond this, as done in the works of [BO19, LW20], in effect also using convexity but merely to establish the large deviation upper bound, and using a stability argument for the large deviation lower bound, usually requiring that V is integrable. Without the integrability of V, a mean-field large deviation principle might not exist, as can be seen in [HLSS18] for hypersingular Riesz gases. The question then arises if we can use similar arguments for our Gibbs-like systems, opening the way for large deviations for interacting diffusions with strong repulsive interactions.

Moreover, we briefly discussed how for the interacting particle systems the large deviation principle in combination with a unique minimizer of the rate function not just implies almost sure convergence, but can also imply entropic chaoticity under additional assumptions. With applications to entropic propagation of chaos in mind, this avenue might be worth exploring as well.

I.B Large deviations for singularly interacting diffusions

In this part, we investigated systems of interacting diffusions with measure-dependent and possibly singular drift. We proved large deviation principles for the empirical measures over path space by relating the interacting system to a non-interacting one by Girsanov's theorem, using the necessary stochastic estimates and the tools of Part I.A. Moreover, we established the uniqueness of the minimizer of the rate function, which is the solution to the associated McKean–Vlasov equation, resulting in the convergence of the empirical measures.

In the case where the drift is characterized by a function $\varphi(x, y)$, representing the force between two particles, we established large deviations and convergence if $\int_0^T |\varphi|^2(X_t, Y_t)\,\mathrm{d}t$ is strongly exponentially integrable with respect to the Wiener measure. We gave sufficient conditions under the assumption of suitable L^p estimates of φ, but these are not likely to be tight, and various questions still remain on which singularities precisely fall into the desired class.

Meanwhile, various results exist on quantitative propagation of chaos, as done for example in [JW18, BJW19], both for strongly repulsive and attractive interaction potentials. One then wonders if we can develop asymmetric generalizations of the tools in I.A and I.B to deal with these strongly repulsive interactions as well. Moreover, it seems to be that it is still not known whether there are interaction potentials for which propagation of chaos holds but a large deviation principle does not.

II.A Variational and dissipation structures for jump processes

Here we introduced variational structures for the FKE corresponding to a jump kernel κ. This energy-dissipation structure consisted of three parts: the relative entropy with respect to a given reference measure π as a free energy functional; a dissipation potential $\mathcal{R}$ acting on unidirectional edge fluxes, relating the free energy functional and the evolution, and a Fisher information functional $\mathcal{D}$.

It was a generalization of existing variational structures, such as those found in [PRST22, KJZ18, PRS21], that incorporated net fluxes and assumed either the detailed balance condition or that π is an invariant measure. In our setting π is merely a fixed reference measure, which has as a consequence that the free energy does not actually dissipate along the solution. Even though, we showed that many of the same concepts of standard EDPs still apply. Moreover, we showed the equivalence of our variational structure to the rate function arising from large deviations of independent particles.

In order to apply the tools to Part II.B we considered three assumptions on the kernel κ: the kernel was bounded; the kernel was unbounded but the corresponding fluxes were bounded; the kernel and the fluxes were (possibly) unbounded but the kernel satisfied the detailed balance condition.

Despite our results, the unboundedness of the kernel still has left various gaps in our understanding. Namely, we do not know yet the precise assumptions under which the different notions of solutions coincide and are unique, and whether under those conditions a corresponding Markov process exists, the chain rule holds, or the rate functional and the EDP-functional are equivalent. For example, it would be interesting to see if one could even construct rapidly increasing kernels on a countable space such that all the desired equivalences do not hold.

Moreover, we considered a generalization of corresponding minimizing movement schemes for the solution to the FKE, which does not satisfy the assumptions of [PRST22], but the question of its convergence still remains open.

II.B Variational convergence for population dynamics

Finally, we considered a generalization of the Bolker–Pacala–Dieckmann–Law model, namely one where the trait space $\mathcal{T}$ is a compact Polish space. We equipped the forward Kolmogorov equation with the corresponding variational structure of Part II.A, trading in edge fluxes for mass fluxes that track the amount of birth and death. Moreover, we showed convergence of these structures in the EDP-sense to one corresponding to the Liouville equation, which itself was lifted from the expected mean-field limit and, in particular, we obtained a variational formulation for the mean-field limit itself. Finally, under suitable conditions on the initial data, we established entropic propagation of chaos as well.

The free energy was defined as the relative entropy with respect to the law of a Poisson point process determined by a fixed finite measure γ over $\mathcal{T}$, simplifying the corresponding calculations and proof of its Γ-convergence but necessitating the structure of Part II.A due to it not being the invariant measure of the system.

We considered two cases: one where we assumed finite entropy and super-quadratic moments, in terms of the mass, on the initial data (which corresponded to the case of bounded fluxes of Part II.A); and one where we merely assumed finite entropy on the initial data, but restricted the kernel to satisfy the detailed balance condition. Aside from $\mathcal{T}$ being compact Polish, the mutation and competition kernels were assumed to be bounded and measurable.

The restriction of compactness might seem strict, but in fact can be removed after various technical modifications, which all rely on the measure γ being *finite*. For locally finite γ over a locally compact Polish space, we do not know whether all results carry over, placing classical results such as [FM04] or the more recent [FKKK15] out of reach of our assumptions. It is possible that some of the full equivalences for all possible admissible curves listed in this thesis might not always hold, but do hold for the solution to the FKE itself.

Next, although we proved entropic propagation of chaos the question remains whether our proposed structure can be used to obtain *quantitative* results, an area that for example is well established for the systems of Part I.B.

Finally, another observation we made was that formally an enhanced version of our variational convergence seems equivalent to a large deviation principle. Large deviations for the BPDL model over a finite space can be derived from known results on large deviations for reacting particle systems, but the question remains whether improvements on our results can be used to prove LDPs for arbitrary trait spaces.

Summary

Mean-field limits and beyond

Large deviations for singular interacting diffusions and variational convergence for population dynamics

In this thesis, we study interacting particle systems, where particles are moving, created, or annihilated. We focus on weak interactions, also called *mean-field* interactions, where the interaction between a particle and all the other particles only depends in a suitable way on the corresponding empirical measure, a rescaled discrete measure that described the positions of all the particles.

For the systems we consider, we prove that as the number of particles goes to infinity the empirical measure converges to the so-called *mean-field limit*. Moreover, we establish either *large deviation principles* (LDPs), which are asymptotic probabilities of deviating from these limits, or show convergence of corresponding variational structures in terms of *Energy-Dissipation Principles* (EDPs).

Aside from the discussion in Part III, the thesis consists of two parts: Part I, with the goal of establishing large deviations for singularly interacting diffusions in Part I.B; and Part II, leading up to Part II.B, where we describe the EDP-convergence for variational structures for the forward Kolmogorov equation (FKE) corresponding to the Bolker–Pacala–Dieckmann–Law (BPDL) model, which is used in ecology and population dynamics. We develop various necessary techniques to establish our results, some of which are delegated to Parts I.A and II.A. We will now briefly summarize these parts.

In Part I.A we establish an extension of Varadhan's Integral Lemma, in order to show LDPs for systems that are tilted by a (possibly) singular energy functional. We employ this to study interacting particle systems of *Gibbs*-type, which are characterized by some interaction potential V, and modify this argument to handle the systems of Part I.B. Moreover, as a consequence, we obtain almost sure convergence and additional estimates if the minimizer of the rate function is unique.

Next, we consider a system of interacting diffusions in Part I.B with measure-dependent and possibly singular drift. We first prove the large deviation principle,

where the minimizer of the rate function corresponds to a solution of the McKean–Vlasov equation. By establishing the uniqueness of these solutions we obtain convergence and propagation of chaos.

In Part II.A we turn to jump processes instead of diffusions. We introduce a variational structure for the FKE under suitable assumptions on the jump kernel, which is not necessarily bounded. The structure is determined by a fixed reference measure, while for existing EDPs the reference measure is also an invariant measure of the system. This allows us to treat irreversible processes without knowing the corresponding invariant measures.

This technique is then put to use in Part II.B, where we consider variational structures for the FKE associated to the BPDL model, with the reference measure now the law of Poisson point process. We show convergence of these structures to one corresponding to the expected mean-field limit, establishing convergence and entropic propagation of chaos along the way.

Acknowledgements

It might be a bit cliché to say that the success of my PhD and the completion of this thesis would not have been possible without certain people, but I'm still going to say it: all of this would not have been possible without all of you, and all of you have my deepest thanks. You know who you are.

Due to the sickness and eventual death of my dear father during my PhD, and the fact that some of the attributes that make me good at mathematics make me horrible at certain other things, it was not always guaranteed I would finish the road I started on. But I did, and that certainly is due to the support of close friends, family, and colleagues. While I cannot name everyone who helped me over the years or simply cheered me up by being there, at least let me name a few.

Oliver, first I would like to thank you, for your guidance, kindness, and support. Looking back, I realize that in the beginning I was much a more stubborn student than I thought I was, but you have always been nothing but patient. If I was amid the umpteenth obsession, question, or very important dilemma (like which notation do you use when you are running out of letters), you would indulge me and jump on in right with me, until you would eventually subtly indicate we would need to take a step back and look at the bigger picture.

You moved me from the joint PhD/TA program by buying me out with your grant, you and Mario took over our article for the time I had to be in Groningen a lot, and you organized a conference just so that I and Anastasiia could have another one. Aside from that, I have learned in Schiermonnikoog that in a game of billiards it is infinitely better to be in your team than to have to play against you.

Next, I thank you, Mark, both for your continued support during these years, and for introducing me to the wonderful world of large deviations and gradient flows, all the way back during the elective with Francesca when I just had finished my bachelor's. You were always ready to answer any one of the thousand questions I would ask, followed up by your own, and I still remember fondly my earliest attempts at writing, which would be a mass of red comments after you were done with it.

Mario, I am grateful for our collaboration and your patience, and it was a pleasure visiting you in York (aside from the thing they dare to call breakfast). Although I still call myself an “analyst who sometimes dabbles in probability”, you are responsible for quite some of that dabbling.

I also want to thank all the committee members for their valuable comments and insight, and their help in my neverending hunt for typos.

At CASA I was surrounded by wonderful colleagues, and the last five years there have brought me great joy. I especially want to thank my past and present office mates: Saeed, Koondi, Mikola, Anastasiia, Alberto, Oxana, Giacomo, and Evi (and our next-door neighbours Jelle and Mahefa). Anastasiia, we have now spent almost our entire PhDs together and it was a blast goofing off, gigglingly brainstorming presents for other people, all amidst very seriously trying to uncover the underlying secrets of gradient flows. Jens, I remember fondly your one-sided attempts at hide-and-seek and thank you for reintroducing me to improv theatre, albeit just before the pandemic hit. Lotte, your enthusiasm at the SSC got me back into working out properly (at least for a time), even though my hardest workout seemed to be your warm-up. Barry, of all the courses I tutored in, I enjoyed Introduction to Numerical Analysis the most, and to this day no one can top the audacity of the student that once claimed on an exam that the necessary steps for a certain algorithm to end was -7.

Throughout it all, I was surrounded by my incredible sisters, Suzanne and Nienke, and by a collection of verifiably crazy but amazing friends. Florian, you have been an absolute rock, and I look forward to continue beating you at FIFA. Marjolein, you are forever allowed to crush me during your hugs. Kris, Jeannette, and Maaike, thank you for all your support these years, and all the delicious dinners at the Stadsbrouwerij. And to the last three and the rest of the Leuk-group, I still find it hilarious that one horrible play brought part of the group together, and hope we will all remain friends for life.

Finally, I want to thank my father. I know you saw a lot of yourself in me, both the good and the bad, but you have always been supportive of my choices and were very happy when I finally switched to math during my studies. Funnily enough, the first time I tutored math was because of you, although merely because you were not able to help my sister with her homework and I had to step in. You are no longer around to witness the end of my PhD, or to always remind me that if I finished it I would have done something you did not have, but this thesis is still for you.

Curriculum Vitae

Jasper Hoeksema was born on April 1st, 1987, in Groningen, the Netherlands. After finishing high school in 2004 at the Zernike College in Haren and enjoying a gap year filled with theatre, he moved to Eindhoven. After a brief stint studying Architectural Engineering at the Eindhoven University of Technology, his love for math won him over, and he started studying Industrial and Applied Mathematics in 2006, completing his bachelor's studies in 2015 and his master's studies in 2017 (cum laude). He graduated with a thesis on large deviations for singular functionals, under the supervision of dr. O.T.C. Tse and prof. dr. M.A. Peletier, which laid the groundwork for the investigations described in Part I of this dissertation.

He continued his research in this direction as a PhD student/teaching assistant at the Centre for Analysis, Scientific computing and Applications (CASA), starting in October 2017. From October 2018 onward, he was a regular PhD student as part of dr. O.T.C. Tse's VIDI-project *Dynamical-Variational Transport Costs and Application to Variational Evolutions*, under the NWO Vidi grant 016.Vidi.189.102, and eventually his work shifted to applying variational convergence techniques to population dynamics, the subject of Part II of this dissertation. The results obtained during both these periods are described in this thesis.

Bibliography

[AB06] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis*. Springer Berlin, Heidelberg, 2006.

[AC08] Luigi Ambrosio and Gianluca Crippa. Existence, Uniqueness, Stability and Differentiability Properties of the Flow Associated to Weakly Differentiable Vector Fields. In *Transport equations and multi-D hyperbolic conservation laws*, volume 5 of *Lect. Notes Unione Mat. Ital.*, pages 3–57. Springer, Berlin, 2008.

[ADdR22] Daniel Adams, Manh H. Duong, and Gonçalo dos Reis. Entropic regularization of nongradient systems. *SIAM J. Math. Anal.*, 54(4):4495–4535, 2022.

[AT14] Luigi Ambrosio and Dario Trevisan. Well-posedness of Lagrangian flows and continuity equations in metric measure spaces. *Anal. PDE*, 7(5):1179–1234, 2014.

[BBBO21] Giada Basile, Dario Benedetto, Lorenzo Bertini, and Carlo Orrieri. Large deviations for Kac-like walks. *J. Stat. Phys.*, 184(1):1–27, 2021.

[BBC17] Martina Baar, Anton Bovier, and Nicolas Champagnat. From stochastic, individual-based models to the canonical equation of adaptive dynamics in one step. *The Annals of Applied Probability*, 27(2):1093–1170, 2017.

[BDF12] Amarjit Budhiraja, Paul Dupuis, and Markus Fischer. Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.*, 40(1):74–102, 2012.

[BDSG+15] Lorenzo Bertini, Alberto De Sole, Davide Gabrielli, Giovanni Jona-Lasinio, and Claudio Landim. Macroscopic fluctuation theory. *Rev. Modern Phys.*, 87(2):593–636, 2015.

[Ber18] Robert J. Berman. On large deviations for Gibbs measures, mean energy and gamma-convergence. *Constr. Approx.*, 48(1):3–30, 2018.

[BFGM19] Lisa Beck, Franco Flandoli, Massimiliano Gubinelli, and Mario Maurelli. Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electron. J. Probab.*, 24:1–72, 2019.

[BG99] Thierry Bodineau and Allice Guionnet. About the stationary states of vortex systems. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 35(2):205–237, 1999.

[BGSRS20] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Fluctuation theory in the Boltzmann-Grad limit. *J. Stat. Phys.*, 180(1-6):873–895, 2020.

[BJW19] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. On mean-field limits and quantitative estimates with a large class of singular kernels: Application to the Patlak–Keller–Segel model. *Comptes Rendus Mathematique*, 357(9):708–720, 2019.

[BO19] Robert J. Berman and Magnus Önnheim. Propagation of chaos for a class of first order models with singular mean field interactions. *SIAM J. Math. Anal.*, 51(1):159–196, 2019.

[Bog07] Vladimir I. Bogachev. *Measure Theory*, volume II. Springer Berlin Heidelberg, 2007.

[BP97] Benjamin Bolker and Stephen W Pacala. Using Moment Equations to Understand Stochastically Driven Spatial Pattern Formation in Ecological Systems. *Theoret. Population Biol.*, 52(3):179–197, 1997.

[But98] Giuseppe Buttazzo. *Semicontinuity, relaxation and integral representation in the calculus of variations*. Pitman Research Notes in Mathematics Series. Longman Scientific and Technical, Harlow, 1998.

[CDFM20] Michele Coghi, Jean-Dominique Deuschel, Peter K. Friz, and Mario Maurelli. Pathwise McKean-Vlasov theory with additive noise. *Ann. Appl. Probab.*, 30(5):2355–2392, 2020.

[CdR20] Paul E. Chaudru de Raynal. Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.*, 130(1):79–107, 2020.

[CFM06] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theoret. Population Biol.*, 69(3):297–321, 2006.

[CFM08] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. From Individual Stochastic Processes to Macroscopic Models in Adaptive Evolution. *Stochastic Models*, 24:2–44, 2008.

[CGZ14] Djalil Chafaï, Nathael Gozlan, and Pierre André Zitt. First-order global asymptotics for confined particles with singular pair repulsion. *Annals of Applied Probability*, 24(6):2371–2413, 2014.

[CHLZ12] Shui-Nee Chow, Wen Huang, Yao Li, and Haomin Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal.*, 203(3):969–1008, 2012.

[CL22] Giovanni Conforti and Christian Léonard. Time reversal of Markov processes with jumps under a finite entropy condition. *Stochastic Process. Appl.*, 144:85–124, 2022.

[Coh13] Donald L. Cohn. *Measure theory*. Birkhäuser Advanced Texts: Basler Lehrbücher. Birkhäuser/Springer, New York, second edition, 2013.

[DE97] Paul Dupuis and Richard S. Ellis. *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, 1997.

[Del19] François Delarue. Restoring uniqueness to mean-field games by randomizing the equilibria. *Stoch. Partial Differ. Equ. Anal. Comput.*, 7(4):598–678, 2019.

[DFMS18] Jean-Dominique Deuschel, Peter K. Friz, Mario Maurelli, and Martin Slowik. The enhanced Sanov theorem and propagation of chaos. *Stochastic Processes and their Applications*, 128(7):2228–2269, 2018.

[DG87] Donald A. Dawson and Jürgen Gärtner. Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[dH08] Frank den Hollander. *Large Deviations*. Fields Institute monographs. American Mathematical Society, 2008.

[DLR20] Paul Dupuis, Vaios Laschos, and Kavita Ramanan. Large deviations for configurations generated by Gibbs distributions with energy functionals consisting of singular interaction and weakly confining potentials. *Electron. J. Probab.*, 25:1–41, 2020.

[DMZ03] Pierre Del Moral and Tim Zajic. A note on the Laplace–Varadhan integral lemma. *Bernoulli*, 9(1):49–65, 2003.

[DO21] Manh H. Duong and Michela Ottobre. Non-reversible processes: GENERIC, Hypocoercivity and fluctuations. *arXiv preprint arXiv:2111.00286*, 2021.

[dRST19] Gonçalo dos Reis, William Salkeld, and Julian Tugaut. Freidlin–Wentzell LDP in path space for McKean–Vlasov equations and the functional iterated logarithm law. *Ann. Appl. Probab.*, 29(3):1487–1540, 2019.

[DZ10] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer Berlin Heidelberg, 2010.

[EFLS16] Matthias Erbar, Max Fathi, Vaios Laschos, and André Schlichting. Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Discrete Contin. Dyn. Syst.*, 36(12):6799–6833, 2016.

[EG15] Lawrence C. Evans and Ronald F. Gariepy. *Measure Theory and Fine Properties, revised edition*. Chapman and Hall/CRC, 2015.

[Eic04] Peter Eichelsbacher. Refined large deviations for von Mises statistics. *Teor. Veroyatn. Primen.*, 49(1):197–204, 2004.

[EK05] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., 2005.

[Erb14] Matthias Erbar. Gradient flows of the entropy for jump processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 50(3):920–945, 2014.

[Erb16] Matthias Erbar. A gradient flow approach to the Boltzmann equation. *arXiv preprint arXiv:1603.00540*, 2016.

[ES98] Peter Eichelsbacher and Uwe Schmock. Exponential approximations in completely regular topological spaces and extensions of Sanov's theorem. *Stochastic Processes and their Applications*, 77(2):233–251, 1998.

[ES02] Peter Eichelsbacher and Uwe Schmock. Large deviations of U-empirical measures in strong topologies and applications. *Annales de l'institut Henri Poincare Probability and Statistics*, 38(5):779–797, 2002.

[EZ03] Peter Eichelsbacher and Tim Zajic. Moderate deviations for mean-field Gibbs measures. *Bernoulli*, 9(1):67–95, 2003.

[Fat16] Max Fathi. A gradient flow approach to large deviations for diffusion processes. *Journal des Mathematiques Pures et Appliquees*, 106(5):957–993, 2016.

[Fed09] Ennio Fedrizzi. Uniqueness and Flow Theorems for solutions of SDEs with low regularity of the drift. Master's thesis, University of Pisa, 2009.

[FF11] Ennio Fedrizzi and Franco Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.

[FHM14] Nicolas Fournier, Maxime Hauray, and Stéphane Mischler. Propagation of chaos for the 2D viscous vortex model. *Journal of the European Mathematical Society*, 16(7):1423–1466, 2014.

[Fis14] Markus Fischer. On the form of the large deviation rate function for the empirical measures of weakly interacting systems. *Bernoulli*, 20(4):1765–1801, 2014.

[FJ17] Nicolas Fournier and Benjamin Jourdain. Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.*, 27(5):2807–2861, 2017.

[FKK09] Dmitri Finkelshtein, Yuri G. Kondratiev, and Oleksandr Kutoviy. Individual based model with competition in spatial ecology. *SIAM J. Math. Anal.*, 41(1):297–317, 2009.

[FKK10] Dmitri Finkelshtein, Yuri G. Kondratiev, and Oleksandr Kutoviy. Vlasov scaling for stochastic dynamics of continuous systems. *J. Stat. Phys.*, 141(1):158–178, 2010.

[FKK21] Dmitri Finkelshtein, Yuri G. Kondratiev, and Peter Kuchling. Markov dynamics on the cone of discrete Radon measures. *Methods Funct. Anal. Topology*, 27(2):173–191, 2021.

[FKKK15] Dmitri Finkelshtein, Yuri G. Kondratiev, Kozitsky Yuri, and Oleksandr Kutoviy. The statistical dynamics of a spatial logistic model and the related kinetic equation. *Math. Models Methods Appl. Sci.*, 25(2):343–370, 2015.

[FM04] Nicolas Fournier and Sylvie Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.*, 14(4):1880–1919, 2004.

[Fon04] Joaquín Fontbona. Uniqueness for a weak nonlinear evolution equation and large deviations for diffusing particles with electrostatic repulsion. *Stochastic Processes and their Applications*, 112(1):119–144, 2004.

[FPZ19] Franco Flandoli, Enrico Priola, and Giovanni Zanco. A mean-field model with discontinuous coefficients for neurons with spatial interaction. *Discrete Contin. Dyn. Syst.*, 39(6):3037–3067, 2019.

[GHM12] Alexander Grigor'yan, Xueping Huang, and Jun Masamune. On stochastic completeness of jump processes. *Mathematische Zeitschrift*, 271(3-4):1211–1239, 2012.

[GNP21] Nir Gavish, Pierre Nyquist, and Mark A. Peletier. Large deviations and gradient flows for the Brownian one-dimensional hard-rod system. *Potential Anal.*, pages 1–51, 2021.

[GQ15] David Godinho and Cristobal Quiñinao. Propagation of chaos for a subcritical Keller–Segel model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(3):965–992, 2015.

[GZ19] David García-Zelada. A large deviation principle for empirical measures on Polish spaces: Application to singular Gibbs measures on manifolds. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(3):1377–1401, 2019.

[Has59] Rafail Z. Has'minskiĭ. On positive solutions of the equation $\mathcal{U}+Vu=0$. *Theor. Probability Appl.*, 4:309–318, 1959.

[HH16] Yuu Hariya and Kaname Hasegawa. On divergence of expectations of the Feynman–Kac type with singular potentials. *J. Math. Soc. Japan*, 68(3):1271–1296, 2016.

[HHMT20] Jasper Hoeksema, Thomas Holding, Mario Maurelli, and Oliver Tse. Large deviations for singularly interacting diffusions. *arXiv preprint arXiv:2002.01295 (accepted, with minor revisions, by Ann. Inst. Henri Poincaré Probab. Stat.)*, 2020.

[HLSS18] Douglas P. Hardin, Thomas Leblé, Edward B. Saff, and Sylvia Serfaty. Large deviation principles for hypersingular Riesz gases. *Constr. Approx.*, 48(1):61–100, 2018.

[HNVW16] Tuomas Hytönen, Jan Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, volume 63 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer, Cham, 2016.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[Hoe17] Jasper Hoeksema. Quasi-continuity and large-deviation principles for singularly interacting particle systems. Master's thesis, Technical University of Eindhoven, 2017.

[HPST20] Bastian Hilder, Mark A. Peletier, Upanshu Sharma, and Oliver Tse. An inequality connecting entropy distance, Fisher information and large deviations. *Stochastic Process. Appl.*, 130(5):2596–2638, 2020.

[HRZ22] Zimo Hao, Michael Röckner, and Xicheng Zhang. Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions. *arXiv preprint arXiv:2204.07952*, 2022.

[HS22] Bastian Hilder and Upanshu Sharma. Quantitative coarse-graining of markov chains. *arXiv preprint arXiv:2201.10256*, 2022.

[HT22] Jasper Hoeksema and Oliver Tse. Generalized gradient structures for measure-valued population dynamics and their large-population limit. *arXiv preprint arXiv:2207.00853 (accepted, with minor revisions, by Calc. Var. Partial Differential Equations)*, 2022.

[HvS21] William R. P. Hammersley, David Šiška, and Łukasz Szpruch. McKean-Vlasov SDEs under measure dependent Lyapunov conditions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(2):1032–1057, 2021.

[Jab19] Jean-Francois Jabir. Rate of propagation of chaos for diffusive stochastic particle systems via Girsanov transformation. *arXiv preprint arXiv:1907.09096*, 2019.

[JST22] Ansgar Jüngel, Ulisse Stefanelli, and Lara Trussardi. A minimizing-movements approach to GENERIC systems. *Math. Eng.*, 4(1):1–18, 2022.

[JTT18] Jean-François Jabir, Denis Talay, and Milica Tomašević. Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic Keller-Segel model without smoothing. *Electron. Commun. Probab.*, 23:1–14, 2018.

[JW18] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Inventiones Mathematicae*, 214(1):523–591, 2018.

[Kal17] Olav Kallenberg. *Random Measures, Theory and Applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, 2017.

[KJZ18] Marcus Kaiser, Robert L. Jack, and Johannes Zimmer. Canonical structure and orthogonality of forces and currents in irreversible Markov chains. *J. Stat. Phys.*, 170(6):1019–1050, 2018.

[KJZ19] Marcus Kaiser, Robert L. Jack, and Johannes Zimmer. A variational structure for interacting particle systems and their hydrodynamic scaling limits. *Commun. Math. Sci.*, 17(3):739–780, 2019.

[KLU99] Yuri G. Kondratiev, Eugene W Lytvynov, and Georgi F Us. Analysis and geometry on marked configuration space. *Methods Funct. Anal. Topology*, 5(1):29–64, 1999.

[KR05] Nicolai V. Krylov and Michael Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[Kra19] Richard C. Kraaij. Gamma convergence on path-spaces via convergence of viscosity solutions of Hamilton-Jacobi equations. *arXiv preprint arXiv:1905.08785*, 2019.

[Kra21] Richard C. Kraaij. Flux large deviations of weakly interacting jump processes via well-posedness of an associated Hamilton-Jacobi equation. *Bernoulli*, 27(3):1496–1528, 2021.

[Lac18] Daniel Lacker. On a strong form of propagation of chaos for McKean–Vlasov equations. *Electron. Commun. Probab.*, 23:1–11, 2018.

[Lac22] Daniel Lacker. Quantitative approximate independence for continuous mean field Gibbs measures. *Electron. J. Probab.*, 27:1–21, 2022.

[LD00] Richard Law and Ulf Dieckmann. Moment Approximations of Individual-based Models. In *The Geometry of Ecological Interactions*, pages 252–270. Cambridge Univ. Press., 2000.

[Léo87] Christian Léonard. Large deviations and law of large numbers for a mean field type interacting particle system. *Stochastic processes and their applications*, 25:215–235, 1987.

[LMPR17] Matthias Liero, Alexander Mielke, Mark A. Peletier, and D. R. Michiel Renger. On microscopic origins of generalized gradient structures. *Discrete Contin. Dyn. Syst. Ser. S*, 10(1):1–35, 2017.

[LS01] Robert S. Liptser and Albert N. Shiryaev. *Statistics of random processes: I. General theory*, volume 5 of *Applications of Mathematics*. Springer, 2001.

[LW20] Wei Liu and Liming Wu. Large deviations for empirical measures of mean-field Gibbs measures. *Stochastic Process. Appl.*, 130(2):503–520, 2020.

[Maa11] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

[Mar12] Mauro Mariani. A Gamma-convergence approach to large deviations. *arXiv preprint arXiv:1204.0640*, 2012.

[Mie11] Alexander Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.

[Mie13] Alexander Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations*, 48(1-2):1–31, 2013.

[Mie16] Alexander Mielke. On evolutionary Gamma-convergence for gradient systems. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, pages 187–249. Lect. Notes Appl. Math. Mech., 3, Springer, Cham, 2016.

[MM20] Jan Maas and Alexander Mielke. Modeling of chemical reaction systems with detailed balance using gradient structures. *J. Stat. Phys.*, 181(6):2257–2303, 2020.

[MPR14] Alexander Mielke, Mark A. Peletier, and D.R. Michiel Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Potential Anal.*, 41(4):1293–1327, 2014.

[MS19] Sima Mehri and Wilhelm Stannat. Weak solutions to Vlasov–McKean equations under Lyapunov-type conditions. *Stoch. Dyn.*, 19(6):1950042, 2019.

[MSW22] Alberto Montefusco, Christof Schütte, and Stefanie Winkelmann. A route to the hydrodynamic limit of a reaction-diffusion master equation using gradient structures. *arXiv preprint arXiv:2201.02613*, 2022.

[MV20] Yuliya Mishura and Alexander Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theor. Probability and Math. Statist.*, 103:59–101, 2020.

[Orr20] Carlo Orrieri. Large deviations for interacting particle systems: joint mean-field and small-noise limit. *Electron. J. Probab.*, 25:1–44, 2020.

[Osh13] Yoichi Oshima. *Semi-Dirichlet forms and Markov processes*, volume 48 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2013.

[PdH96] Paolo D. Pra and Frank den Hollander. McKean–Vlasov limit for interacting random processes in random media. *Journal of Statistical Physics*, 84(3-4):735–772, 1996.

[Pel14] Mark A. Peletier. Variational modelling: Energies, gradient flows, and large deviations. *arXiv:1402.1990*, 2014.

[PR19] Robert I. A. Patterson and D. R. Michiel Renger. Large deviations of jump process fluxes. *Math. Phys. Anal. Geom.*, 22(3):1–32, 2019.

[PRS21] Robert I. A. Patterson, D. R. Michiel Renger, and Upanshu Sharma. Variational structures beyond gradient flows: a macroscopic fluctuation-theory perspective. *arXiv preprint arXiv:2103.14384*, 2021.

[PRST22] Mark A. Peletier, Riccarda Rossi, Giuseppe Savaré, and Oliver Tse. Jump processes as generalized gradient flows. *Calc. Var. Partial Differential Equations*, 61(1):1–85, 2022.

[PS19] Mark A. Peletier and Mikola C. Schlottke. Large-deviation principles of switching Markov processes via Hamilton-Jacobi equations. *arXiv preprint arXiv:1901.08478*, 2019.

[PS22] Mark A. Peletier and André Schlichting. Cosh gradient systems and tilting. *arXiv:2203.05435*, 2022.

[Ren18] D. R. Michiel Renger. Flux large deviations of independent and reacting particle systems, with implications for macroscopic fluctuation theory. *J. Stat. Phys.*, 172(5):1291–1326, 2018.

[Rey18] Julien Reygner. Equilibrium large deviations for mean-field systems with translation invariance. *Ann. Appl. Probab.*, 28(5):2922–2965, 2018.

[RZ21a] D. R. Michiel Renger and Johannes Zimmer. Orthogonality of fluxes in general nonlinear reaction networks. *Discrete Contin. Dyn. Syst. Ser. S*, 14(1):205–217, 2021.

[RZ21b] Michael Röckner and Xicheng Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli*, 27(2):1131–1158, 2021.

[Sch19] André Schlichting. Macroscopic limit of the Becker-Döring equation via gradient flows. *ESAIM Control Optim. Calc. Var.*, 25:1–36, 2019.

[Sch21] Mikola C. Schlottke. *Large Deviations of Irreversible Processes*. PhD thesis, Eindhoven University of Technology, 2021.

[Ser11] Sylvia Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011.

[SS04] Etienne Sandier and Sylvia Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

[SW15] René L. Schilling and Jian Wang. Lower bounded semi-Dirichlet forms associated with Lévy type operators. In *Festschrift Masatoshi Fukushima*, volume 17 of *Interdiscip. Math. Sci.*, pages 507–526. World Sci. Publ., Hackensack, NJ, 2015.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In Paul-Louis Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*, pages 165–251. Springer Berlin Heidelberg, 1991.

[Szn98] Alain-Sol Sznitman. *Brownian motion, obstacles and random media*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[Tan84] Hiroshi Tanaka. Limit theorems for certain diffusion processes with interaction. In *Stochastic Analysis*, volume 32 of *North-Holland Mathematical Library*, pages 469–488. Elsevier, 1984.

[Tom20] Milica Tomasevic. Propagation of chaos for stochastic particle systems with singular mean-field interaction of L^q–L^p type. *hal preprint hal:03086253*, 2020.

[Tou09] Hugo Touchette. The large deviation approach to statistical mechanics. *Phys. Rep.*, 478(1-3):1–69, 2009.

[Var84] S.R. Srinivasa Varadhan. *Large Deviations and Applications*. Society for Industrial and Applied Mathematics, 1984.