

Mathematical Institute
University of Oxford

Evolution PDEs on Graphs and Inverse Problems in Aggregation-Diffusion Equations

László Mikolás
St. Anne's College
University of Oxford

Supervised by
Prof. José A. Carrillo & Prof. Antonio Esposito

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2024

Dedication

First, I want to express my deepest gratitude to José and Antonio for being the most amazing supervisors a student could ask for. You took a chance on me during a very challenging time and I will never forget this. From the bottom of my heart, thank you.

I want to thank my parents for supporting my dream of doing a Ph.D. in mathematics from day one, and never having doubted my ability to achieve it regardless of how far-fetched it seemed when it started. Gracias and köszönöm!

This thesis would not have been possible without the support of the faculty of the Department of Stochastics of the Budapest University of Technology and Economics during my master's degree. In particular, I want to thank Prof. Károly Simon, Prof. Bálint Vető, Prof. Balázs Ráth and Prof. Roland Molontay for their support, guidance and encouragement.

I want to thank La Isla, for preventing me from fully transforming into a robot during my Ph.D. and for giving me some of the most fun and crazy times of my life. I know there are still many to come.

I want to thank Aye, for her support and understanding and for making this last year, as challenging as it has been, one of the happiest and most exciting so far.

Finally, I want to dedicate this thesis to Andrea, the brother I met at Oxford. Grazie broccoli!

Statement of Originality

I declare that the contents of this thesis are, to the best of my knowledge, original and my own work, except where otherwise indicated, cited, or commonly known, nor has any part of this thesis been submitted for a degree at another university.

The work in Chapter 2 of Part I has been published in [EM24]. The works in Chapter 3 and Chapter 4 of Part I are part of ongoing work with Prof. José A. Carrillo and Prof. Antonio Esposito [CEM24].

The work in Chapter 2 of Part II is a joint work with Prof. José A. Carrillo, Prof. Sui Tang and Prof. Gissell Estrada-Rodríguez. It is a preprint submitted for publication [CERMT24].

Abstract

This thesis contains contributions in two areas of research and is therefore divided into two parts. On the one hand we study evolution partial differential equations (PDEs) on graphs, and on the other, we address the inverse problem of learning interaction potentials in aggregation-diffusion equations.

In Part One, we investigate evolution PDEs on co-evolving graphs, i.e. a setting in which the topology of the graph evolves in time with its dynamics coupled to the dynamics on the graph. We establish the well-posedness of the system and investigate scenarios where the graph evolves at different rates—slower or faster—than the dynamics on the graph. Additionally, we derive a discrete-to-continuum limit as the number of vertices tends to infinity. We lift the equation on the co-evolving graph to a type of nonlinear continuity equation on Euclidean space whose velocity field depends on the graph structure. We obtain its well-posedness and comment on discrete-to-continuum limits as well as on the contractivity of solutions of this equation. Furthermore, we describe the long-time behaviour of equations on co-evolving graphs with a particular type of velocity leading to diffusive dynamics.

Part Two addresses the inverse problem of learning the interaction potential of an aggregation-diffusion equation. We use the PDE's numerical solutions as data and propose a quadratic error functional enhanced with sparsity promoting regularization to solve the learning problem. We obtain stability estimates showing that this error functional controls the 2-Wasserstein distance between solutions generated by the true and learned interaction potential and show the effectiveness of our approach with numerical examples in one and two dimensions.

Contents

Introduction v
Structure of the thesis vi

I Evolution PDEs on Graphs **1**

1 Introduction **4**
1.1 Motivation and general framework 4
1.2 Graphs and particle systems with heterogeneous interactions 9
1.3 The contributions of this thesis 12
1.3.1 Chapter 2: evolution PDEs on co-evolving graphs 12
1.3.2 Chapter 3: the graph continuity equation 14
1.3.3 Chapter 4: long-time behaviour 16
1.4 Perspectives and future work 19

2 On evolution PDEs on co-evolving graphs **20**
2.1 Introduction 20
2.2 Preliminaries 22
2.2.1 A priori properties 26
2.3 Well-posedness of the co-evolving non-local continuity equation 30
2.4 Fast-Slow versions of the problem 36
2.5 Discrete-to-continuum limit 41

3 The graph continuity equation **46**
3.1 Introduction 46
3.1.1 Euler Equation 46
3.2 The graph continuity equation 54
3.2.1 Comments on the asymptotic behaviour and discrete-to-continuum limit of (Euler Co-NCL) 66

4 Long-time behaviour **67**
4.1 Introduction . . . 67
4.2 L^∞ theory for (Euler Co-NCL) . . . 67
4.2.1 Well-posedness . . . 71
4.3 Long-time behaviour . . . 74

Appendices **83**

Appendix A Auxiliary results **83**

II Inverse Problems for Aggregation-Diffusion Equations **88**

1 Introduction **90**
1.1 The aggregation-diffusion equation . . . 90
1.2 The role of the interaction potential . . . 94
1.3 The inverse problem for learning interaction potentials . . . 95
1.4 The contributions of this thesis . . . 97
1.5 Perspectives and future work . . . 99

2 Sparse identification of nonlocal interaction kernels in nonlinear gradient flow equations via partial inversion **101**
2.1 Introduction . . . 101
2.2 A regularised variational approach via basis pursuit . . . 104
2.2.1 The error functional . . . 104
2.2.2 The vanilla least squares solutions . . . 106
2.2.3 ℓ^1 regularization via Basis Pursuit . . . 107
2.3 Stability estimates and Γ-convergence . . . 108
2.3.1 No diffusion, the aggregation equation case . . . 109
2.3.2 McKean-Vlasov dynamics with solution dependent diffusion . . . 113
2.3.3 Nonlinear diffusion . . . 114
2.4 Proposed Numerical Schemes . . . 115
2.4.1 Discrete error functional . . . 116
2.4.2 Error bounds . . . 118
2.4.3 Noisy data . . . 120
2.4.4 Support pruning algorithm . . . 121
2.5 Numerical examples . . . 123
2.5.1 Data generation . . . 123

2.5.2 One dimensional examples 125
2.5.3 Two dimensional examples 132
2.6 Conclusion and future work 136

Appendices **137**

Appendix A Proofs of Dobrushin-type stability estimates **137**
A.0.1 Proof of Proposition 2.3.1 137
A.0.2 Assumptions and proof of Proposition 2.3.3 140

Appendix B Error estimate with noisy data **146**

Introduction

Partial Differential Equations (PDEs) are a powerful tool to model the evolution of complex physical systems across multiple scales and domains. For instance, PDEs have been used to study cell populations or to approximate the behaviour of large flocks of birds, among many other examples, see e.g. [BDT17], and the references therein. Although at different scales, in both cases there is a vast number of agents or particles that interact in ways that can give rise to the formation of complex patterns. PDEs can be useful to approximate the average behaviour of such agents as their number becomes very large, making an agent-based description rather complicated.

Recently, evolution equations have also been used to study the behaviour of algorithms in data science and machine learning. Examples include the use of PDEs to capture the training process of large neural networks and certain clustering algorithms [RVE22, CGTS21]. In this context, datasets can be naturally represented by graphs. Indeed, a dataset can be modelled with a graph so that each data point corresponds to a vertex and the edges are given by some measure of similarity between vertices, such as the distance between them. Graphs also naturally arise in other contexts such as transportation networks, opinion dynamics in social networks, or modelling of epidemics. Thus, a question of interest is to understand the behaviour of solutions to PDEs describing the evolution of a certain quantity that is stored on the vertices of a graph and transmitted along the edges. In our previous examples, this quantity could represent the pressure at a given location in a biological transportation network [ABH$^+$17], the amount of support for a given opinion in an opinion formation model [SKCS23, BLSSS20], or the estimated value of a function at a given location in machine learning algorithms [CGTS21].

In this thesis, we study problems related to evolutions on graphs and PDEs motivated by interacting particle systems. On the one hand, we study a class of equations on *co-evolving* graphs, i.e., equations describing the evolution of some mass on the vertices of a graph whose connectivity evolves in time depending on the configuration of the mass on the vertices. Namely, we study the following system

$$\begin{aligned} \partial_t \rho_t &= -\overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]], \\ \partial_t \eta_t &= \omega[\rho_t] - \eta_t, \end{aligned} \tag{Co-NCL}$$

where ρ is a measure representing the mass on the vertices of a graph, $\overline{\nabla}$ and $\overline{\nabla}\cdot$ are the graph gradient and divergence, respectively, F^{Φ} is the flux, which depends on μ, a measure whose support represents the vertices of a graph, η, a function accounting for the weighted edges of a graph, and V_t, a potentially solution-dependent velocity field. The evolution of the edge-weight

function η depends on a function ω which can depend non-locally on the configuration of the mass ρ. The coupling of both equations is what captures the co-evolving nature of the problem. Our contributions on this topic can be found in Chapter 2, Chapter 3 and in Chapter 4.

On the other hand, we study the inverse problem of learning the interaction rules in a large system of interacting particles from solutions to the PDEs approximating the particles' average behaviour. Inverse problems arise in situations in which we can only observe the effect or output of a certain experiment and we want to infer the inputs. In our case the output will be the solution of an aggregation-diffusion equation

$$\partial_t \rho = \nabla \cdot [\rho \nabla (H'(\rho) + V + W * \rho)] \ , \tag{Agg-diff}$$

for $t > 0$, where ρ is a probability density giving the average description of the particle system, H' is the first variation of an internal energy, V is a confinement potential and W is an interaction potential. The interaction potential is the input we try to infer. Our work on this topic can be found in Chapter 2.

Structure of the thesis

This thesis is divided into two parts, each devoted to the study of one of the previously described problems. The first part focuses on the study of evolution PDEs on graphs. In Chapter 1, we present the problem and review the main results in this field. In Chapter 2, we present our work [EM24] about the well-posedness of the equation (Co-NCL), its behaviour on different time-scales, and a discrete-to-continuum limit. Chapter 3 and Chapter 4 contain ongoing work at the time of submission of this thesis [CEM24]. In Chapter 3 we study the link between (Co-NCL) and a particular continuity equation on Euclidean space. In Chapter 4 we study the long time behaviour of (Co-NCL) for a particular class of velocities.

In the second part, we study the inverse problem of learning the interaction potential of (Agg-diff). In Chapter 1 we introduce equation (Agg-diff), review some of its main properties, and present the related literature on the inference of the interaction potential from aggregation-diffusion equations. In Chapter 2, we present our work [CERMT24] in which we propose a method for solving the inverse problem based on a quadratic error functional with a sparsity promoting regularization.

At the conclusion of each part, we shall present open questions on both topics that will be object of future research.

Part I

Evolution PDEs on Graphs

Notation

We list notation we shall use throughout Part I.

- A denotes a generic subset of $\mathbb{R}^d$.
- $\mathcal{B}(A)$: Borel subsets of A.
- $\mathscr{L}^1$ denotes the one-dimensional Lebesgue measure.
- $\mathcal{M}(A)$: Radon measures on A.
- $\mathcal{M}^+(A)$: nonnegative Radon measures on A.
- $\mathcal{P}_p(X)$ is the space of probability measures with finite p-th moments.
- Given $\nu \in \mathcal{M}(\mathbb{R}^d)$ and letting $A \in \mathcal{B}(\mathbb{R}^d)$, we denote by $\nu^+(A) := \sup_{B\in\mathcal{B}(A)} \nu(B)$ and $\nu^-(A) := -\inf_{B\in\mathcal{B}(A)} \nu(B)$ the upper and lower variation measures of ν; the total variation measure of ν is $|\nu|(A) := \nu^+(A) + \nu^-(A)$ and its total variation norm is $\|\nu\|_{\mathrm{TV}} := |\nu|(\mathbb{R}^d)$.
- $\mathcal{M}_{\mathrm{TV}}(A)$: Radon measures on A with finite total variation.
- $\mathcal{M}^+_{\mathrm{TV}}(A) := \mathcal{M}^+(A) \cap \mathcal{M}_{\mathrm{TV}}(A)$.
- $\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d) := \big\{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) : |\rho|(\mathbb{R}^d) \leqslant M\big\}$.
- For any metric space X, $\mathcal{P}(X)$ denotes the space of probability measures over X and $\mathcal{P}_2(X)$ is the space of probability measures with finite second moments.
- For a measurable function f and a measure ν, we denote by $f_{\#}\nu$ the push-forward of ν through f, i.e. for $A \in \mathcal{B}(X)$, $f_{\#}\nu(A) = \nu(f^{-1}(A))$.
- For $i = 1, 2$, $\pi^i : X \times X \to X$ denotes the canonical projeciton map to the i-th coordinate, i.e. $\pi^1(x, y) = x$ and $\pi^2(x, y) = y$.
- For two measures $\nu, \xi \in \mathcal{P}_2(X)$, we denote by $\Pi := \{\gamma \in \mathcal{P}_2(X \times X) \mid {\pi^1}_{\#}\gamma = \nu,\ {\pi^2}_{\#}\gamma = \xi\}$, the space of couplings between ν and ξ.
- For $\nu, \xi \in \mathcal{P}_p(X)$ we denote by $d_p(\nu, \xi) := \inf\left\{\int_{X\times X} |x - y|^p \mathrm{d}\gamma(x, y) \mid \gamma \in \Pi(\nu, \xi)\right\}^{1/p}$, the p-Wasserstein distance between ν and ξ for $p \in [1, \infty)$.
- For a measure $\mu \in \mathcal{P}_2(X)$, $\mathcal{P}_2^\mu(X \times X) := \{\xi \in \mathcal{P}_2(X \times X) \mid {\pi^2}_{\#}\xi = \mu\}$, the space of probability measures with finite second moments over $X \times X$, with second marginal $\mu \in \mathcal{P}_2(X)$.

- For two measures $\nu,\ \xi \in \mathcal{P}_2^\mu(X \times X)$, we denote by $L^2_\mu d_2(\nu,\xi) := \left(\int_X d_2^2(\nu_x,\xi_x)\mathrm{d}\mu(x)\right)^{1/2}$, where $\{\xi_x\}_{x\in X},\ \{\nu_x\}_{x\in X} \subset \mathcal{P}_2(X)$ are the disintegrations of ξ and ν with respect to μ, respectively, i.e. $\nu = \int_X \nu_x \mathrm{d}\mu(x)$ and $\xi = \int_X \xi_x \mathrm{d}\mu(x)$.
- $C_0(\mathbb{R}^d)$ is the space of continuous functions on $\mathbb{R}^d$ vanishing at infinity.
- $C_b(\mathbb{R}^d)$ is the space of continuous and bounded functions on $\mathbb{R}^d$.
- For a sequence $\{\nu^n\}_{n\in\mathbb{N}} \subset \mathcal{M}^+_{\mathrm{TV}}(X)$ and $\nu \in \mathcal{M}_{\mathrm{TV}}(X)$, we write $\nu^n \overset{*}{\rightharpoonup} \nu$ if $\int \varphi \mathrm{d}\nu^n \to \int \varphi \mathrm{d}\nu$ for any $\varphi \in C_0(X)$.
- For any $x \in X$, δ_x denotes a dirac-mass at x and for any set A, $\mathbf{1}_A(x) = 1$ if $x \in A$ and 0 otherwise.
- $L^p_\mu(X)$ is the Lebesgue space of p-integrable functions over X with respect to a measure μ.
- $\mathbb{R}^{2d}_{\not\diagup} := \left\{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d : x \neq y\right\}$ is the off-diagonal of $\mathbb{R}^d \times \mathbb{R}^d$. We will use an analogous notation for subset $K \subset \mathbb{R}^d$.
- $\mu \in \mathcal{M}^+(\mathbb{R}^d)$ is referred to as the base measure and acts as an abstract notion of vertices.
- $\eta : \mathbb{R}^{2d}_{\not\diagup} \to \mathbb{R}$ is the edge weight function.
- G is the set of edges; i.e., $G = \left\{(x,y) \in \mathbb{R}^{2d}_{\not\diagup} : \eta(x,y) \neq 0\right\}$.
- $\mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\not\diagup})$ is the set of antisymmetric vector fields on $\mathbb{R}^{2d}_{\not\diagup}$; that is, $\mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\not\diagup}) = \{v : \mathbb{R}^{2d}_{\not\diagup} \to \mathbb{R} : v(x,y) = -v(y,x)\}$.
- $\overline{\nabla}\phi(x,y) = \phi(y) - \phi(x)$ denotes the nonlocal gradient for $\phi : \mathbb{R}^d \to \mathbb{R}$.
- $\overline{\nabla} \cdot j$ is a nonlocal divergence for a flux $j \in \mathcal{M}(\mathbb{R}^{2d}_{\not\diagup})$, cf. Definition 2.2.5.
- T is a positive and finite final time.
- $AC([0,T];\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$ is the space of absolutely continuous curves with respect to $\|\cdot\|_{TV}$ from $[0,T]$ to $\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$.
- Given $a \in \mathbb{R}, a_+ := \max\{0,a\}$ and $a_- := (-a)_+$ are its positive and negative parts, respectively.
- We shall denote the time dependence by using subscripts. For instance, for a curve $\rho \in AC([0,T];\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$ we use ρ_t to denote an element in $\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$, for any $t \in [0,T]$.

Chapter 1

Introduction

1.1 Motivation and general framework

There has been a substantial increase in the availability of large and complex datasets in recent years. As vast amounts of data became available, increasing research efforts were directed towards designing efficient methods to extract meaningful information. In particular, it has motivated many developments within unsupervised learning, a field within machine learning mainly used to perform exploratory data analysis tasks. As opposed to supervised learning, where the goal is to predict a given target variable Y using data from some relevant features $X_1, \dots, X_p$, unsupervised learning is concerned with learning certain properties of the features $X_1, \dots, X_p$ themselves.

One of the most popular types of tasks in unsupervised learning is *clustering*. For a given dataset with n observations of p features, $\{x_{i,1}, x_{i,2}, \dots, x_{i,p}\}_{i=1}^n$, the goal of this task is to find groups, or clusters, of observations such that the intra-group similarity is as large as possible whereas the inter-group similarity is as low as possible, for some suitable similarity measure. The output of a clustering procedure is a collection of sets of indices $C_1, C_2, \dots, C_K$, such that, roughly speaking, the observations $\{x_{i,1}, \dots, x_{i,p}\}_{i \in C_j}$, for $j = 1, \dots, K$, are more "similar" among themselves than with observations in some other cluster C_ℓ, with $\ell \neq j$, c.f. [JWH+23, Chapter 12.4] for an introduction to clustering. This technique is widely used in applications such as customer segmentation, where companies try to categorize their customers based on their preferences and purchasing behaviour, or in fraud detection, where clustering can be used to identify potentially fraudulent transactions. We mention that many clustering algorithms can be cast in a dynamical way where the dynamics seek to minimize an energy. For example, the mean-shift algorithm [CPn16, CGTS21] can be written as a gradient descent algorithm that attempts to find local modes of some density.

A dataset with observations $\{x_1, \dots, x_n\}$ can be naturally interpreted as a weighted graph. Indeed, from such a dataset one can build a graph $G = (V, W)$, where the set of vertices $V = \{x_1, \dots, x_n\}$ is given by the data and the weights $W = \{w_{i,j}\}_{i,j=1}^n$ are given by some measure of pairwise similarity between vertices such as the distance $w_{i,j} = d(x_i, x_j)$ for $i, j = 1, \dots, n$ for some distance d, see [NJW01, GTS16, GTS18, Tok18] and the references therein.

However, as aforementioned, the increasing size of datasets highlights the need to have a mathematical framework that can handle not only finite, but also uncountably infinite graphs. This led to the following definition of a graph in [EPSS21] and later contributions, such as, [EM24] (see also [GTS16, GTS18]).

Definition 1.1.1. A graph is given by a pair $\mathcal{G} := (\mu, \eta)$, where $\mu \in \mathcal{M}^+(\mathbb{R}^d)$ is a positive Radon measure whose support describes the set of vertices and $\eta \in C_b(\mathbb{R}^{2d}_{\not\sim})$ is the edge-weight function on the possible set of edges $\mathbb{R}^{2d}_{\not\sim} := \{(x, y) \in \mathbb{R}^{2d} \mid x \neq y\}$.

In view of the previous definition, self-loops are not allowed and, for example, any finite set of observations $\{x_1, \dots, x_n\} \subset \mathbb{R}^{nd}$ corresponds to empirical measures such as $\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$, where δ_{x_i} denotes a dirac-mass at each $x_i \in \mathbb{R}^d$ for $i = 1, \dots, n$. The same framework allows to easily represent vast amounts of data by taking for example $\mu^\infty = \lim_{n\to\infty} \mu^n$, interpreting the limit in a suitable weak-sense. Natural examples of the edge-weight function η are functions of the distance $\eta = f(|x - y|)$. Thus, Definition 1.1.1 includes both finite and infinite weighted graphs.

To the best of the author's knowledge, the literature on evolution PDEs on graphs defined in this framework can be traced back to [EPSS21], although similar frameworks appear in [Maa11, CHLZ12] to study problems related to the gradient-flow interpretation of diffusion processes on graphs and in [Mie11], where the author studies the gradient structure of reaction diffusion systems.

In [EPSS21], the authors develop the graph analogue of the nonlocal interaction equation on $\mathbb{R}^d$ with respect to a corresponding graph version of the Wasserstein distance. We mention that there are several works on dynamical systems on graphs related to optimal transport. In [Maa11, CHLZ12] the authors study suitable metrics for which Markov chains on graphs define a gradient flow with respect to the entropy functional and in [Mie13] the author shows the λ-convexity of the entropy in this setting. The work [CL06] uses diffusion maps, a framework based on diffusion processes to understand geometric properties of datasets. Further works investigating the gradient flow structure of diffusive dynamics on discrete spaces include [DL15, HT23, HST23]. The problem of characterising the discrete-to-continuum limit of graph gradient flows as the number of vertices goes to infinity is considered in [FMP22, GvGO22].

In [EFMM22] the authors define a suitable analog version of the Wasserstein distance on

metric graphs and study the McKean-Vlasov PDE as a gradient flow in the space of probability measures over metric graphs with respect to this distance. In [BHP23] the authors define an analog of the dynamical version of the Wasserstein distance for networks, while [MRT15] studies a similar problem, but define a static analog of the Wasserstein distance on metric graphs. Metric and asymptotic properties of nonlocal Wasserstein distances are considered in [SW23]. We conclude the literature review by mentioning that structures resembling graphs have been also noticed in collision dynamics in kinetic theory, see [EGSS24], where the authors introduce a nonlocal collision metric in order to propose a kinetic interpretation of the nonlocal aggregation equation. The underlying state space resembles a graph and the PDE under study takes the form of a nonlocal continuity equation.

Following [EPSS21], let us turn our attention to dynamics driven by interaction energies, which we begin presenting in the Euclidean setting. On $\mathbb{R}^d$, the nonlocal interaction equation reads

$$\partial_t \rho = \nabla \cdot (\rho \nabla K * \rho) \ , \tag{1.1.1}$$

where $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ is a probability measure with finite second moments and $K \in C(\mathbb{R}^d)$ is a suitably regular interaction kernel. It can be shown that (1.1.1) can be derived as a gradient flow of the interaction energy $\mathcal{E}(\rho) = \frac{1}{2}\iint_{\mathbb{R}^d \times \mathbb{R}^d} K(x-y)\mathrm{d}\rho(x)\mathrm{d}\rho(y)$ with respect to the 2-Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$, see [AFP00] and [CDF+11]. More precisely, we can write

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{E}(\rho)}{\delta \rho} \right) \ ,$$

where $\frac{\delta \mathcal{E}(\rho)}{\delta \rho}$ is the *first variation* of $\mathcal{E}$ at ρ, defined as follows

Definition 1.1.2 ([San15, Definition 7.12])**.** Given a functional $\mathcal{E} : \mathcal{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, we say that $\rho \in \mathcal{P}(\Omega)$ is *regular for* $\mathcal{E}$ if

$$\mathcal{E}((1-\varepsilon)\rho + \varepsilon\tilde{\rho}) < +\infty$$

for every $\varepsilon \in [0,1]$ and every $\tilde{\rho} \in \mathcal{P}(\Omega) \cap L^\infty_c(\Omega)$ (absolutely continuous probabilities with L^∞ density and compact support). If ρ is regular for $\mathcal{E}$, we call $\frac{\delta \mathcal{E}}{\delta \rho}(\rho)$, if it exists, any measurable function such that

$$\frac{d}{d\varepsilon}\mathcal{E}(\rho + \varepsilon\chi)\Big|_{\varepsilon=0} = \int \frac{\delta \mathcal{E}}{\delta \rho}(\rho)\, d\chi$$

for every perturbation $\chi = \tilde{\rho} - \rho$ with $\tilde{\rho} \in L^\infty_c(\Omega) \cap \mathcal{P}(\Omega)$.

We closely follow [PS23, EHS23, Hei24] in our presentation of the continuity equation on graphs and some of its properties. To see the translation of (1.1.1) to the framework of Definition

1.1.1, we break the equation up into the following three components

$$\begin{aligned}\partial_t \rho(x) &= -\nabla \cdot \boldsymbol{j}(x) \ , \\ \boldsymbol{j}(x) &= \rho(x) v[\rho](x) \ , \\ v[\rho](x) &= -\nabla \frac{\delta \mathcal{E}(\rho)}{\delta \rho}(x) \ .\end{aligned}$$

The first equation is the traditional continuity equation for $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ with flux $\boldsymbol{j}$. The second equation is a *kinetic relation*, which states that the flux is the mass multiplied by the velocity. Finally, the last equation is a *constitutive law* describing the velocity in terms of the derivative of the energy. When translated to the graph framework, the previous three equations read

$$\begin{aligned}\partial_t \rho(x) &= -(\overline{\nabla} \cdot \boldsymbol{j})(x) \ , \\ \boldsymbol{j}(x,y) &= \Phi(\rho(x), \rho(y); v[\rho](x,y)) \eta(x,y) \ , \\ v[\rho](x,y) &= -\overline{\nabla} \frac{\delta \mathcal{E}(\rho)}{\delta \rho}(x,y) \ .\end{aligned} \tag{1.1.2}$$

We begin by noting that, in the graph formulation, the gradient and the divergence are replaced by their natural discrete analogues $\overline{\nabla}$ and $\overline{\nabla}\cdot$; the discrete gradient of a function is given by its difference at two points and the discrete divergence is defined by duality, see Definition 2.2.1 for the precise definitions of these quantities. Furthermore, the kinetic relation now depends on an interpolation function Φ relating the mass at any two distinct vertices $x, y \in \mathbb{R}^d$ connected by an edge $(x,y) \in \mathbb{R}^{2d}_{\diagup}$ with the corresponding velocity on (x,y). We note that the kinetic relation also depends on the connectivity given by the edge-weight function η.

The interpolation becomes necessary in the graph setting because the mass is defined on the vertices, whereas the velocity and the flux are defined on the edges. Interpolating the mass at the vertices connected by a given edge allows to compare edge-based quantities. We note that the interpolation can depend on the velocity or not, as we see in relevant cases used in the literature. We begin by mentioning the *upwind interpolation* $\Phi_{upwind}(\rho(x), \rho(y); v(x,y)) := \rho(x) v_+(x,y) - \rho(y) v_-(x,y)$, where $v_+(x,y) = \max\{0, v(x,y)\}$ and $v_-(x,y) = \max\{0, -v(x,y)\}$ is velocity-dependent. Interpolations such as $\Phi(\rho(x), \rho(y); v(x,y)) = \phi(\rho(x), \rho(y)) v(x,y)$, for functions ϕ like the arithmetic mean $\phi(\rho(x), \rho(y)) = \frac{\rho(x)+\rho(y)}{2}$, yield a multiplicative kinetic law as in the Euclidean case and do not depend on the velocity.

The choice of the interpolation has a major impact on the resulting dynamics, hence, it is not unique. Common choices in the literature like the logarithmic interpolation $\Phi_{log}(a,b;w) = \frac{a-b}{\log a - \log b} \cdot w$ (where $\Phi_{log}(0,b;w) = \Phi_{log}(a,0;w) = 0$) as in [Maa11, EM12, CHLZ12], or the geometric mean $\Phi_{geo}(a,b;w) = \sqrt{ab} \cdot w$ do not allow the corresponding dynamics to expand

the support of the solutions for velocities obtained from repulsive interaction energies as in (1.1.2). To see this, consider two vertices $x_{out},\ x_{in} \in \mathbb{R}^d$ such that mass flows from x_{out} to x_{in}, that is, $v(x_{out}, x_{in}) > 0$. Then, we see from the definition of Φ_{log} and Φ_{geo} that $\Phi_{log}(\rho(x_{out}), 0;\ v(x_{out}, x_{in})) = \Phi_{geo}(\rho(x_{out}), 0;\ v(x_{out}, x_{in})) = 0$ meaning that vertices with no mass cannot gain mass from connected vertices that do have mass, and thus, the support will not be expanded. Works that use interpolations such as Φ_{log} typically consider dynamics with a diffusive component by choosing the energy to be the Bolzmann entropy [Maa11], which allows the support to expand.

Other choices such as the arithmetic mean described above can lead to solutions becoming negative. Indeed, it is easy to check that this interpolation does not vanish when $\rho(x_{out})$ vanishes. These considerations motivated the choice $\Phi = \Phi_{upwind}$ in [EPSS21] as it preserves non-negativity of ρ and allows the support of solutions to expand without the presence of diffusion in the dynamics. We also note that this interpolation was used first for graph Wasserstein distances in [CGT18] and that it connects the study of evolution PDEs on graphs with numerical methods for PDEs [EGH00], see e.g., [BCH20, CGT20, BCMS20, HST23] for particular cases of numerical schemes used for gradient flows. We mention that the PDE obtained in [EPSS21] is called the nonlocal nonlocal interaction equation, where the first "nonlocal" stems from the nonlocality of graphs due to their discrete nature.

After [EPSS21], many extensions have been considered. In [EPS23], the authors studied a generalization by considering a class of interpolation functions Φ satisfying certain Lipschitz continuity and linear growth properties, (see Definition 2.2.3 for precise details). We remark that this class encompasses Φ_{upwind}, the interpolation considered in [EPSS21]. By our previous discussion on the effects of the choice of interpolation function on solutions to (1.1.2), in this work, the mass is allowed to become negative. The authors proved the well-posedness of the equation via a Banach-fixed point argument and, when $\rho \ll \mu$ and $\Phi = \Phi_{upwind}$ the authors show non-negativity preservation and L^p regularity of solutions. In [HPS23b], the authors extend the framework of [EPSS21] to the case of a two-species system interacting on graphs. They allow for non-linear mobilities and consider a suitable p-Wasserstein quasi-metric for $p \in (1, \infty)$, generalizing the metric defined in [EPSS21]. The authors show that the multi-species dynamics also define a gradient flow of the corresponding interaction energy with respect to the generalized version of p-Wasserstein metric on graphs. The study of nonlocal cross interaction equations on graphs is continued by the same authors in [HPS23a]. There, the authors investigate certain qualitative properties of solutions for finite graphs. Namely, they give conditions on the interaction and cross-interaction kernels leading to the energy optimality of aggregated or segregated solutions and give conditions for obtaining stationary states. Finally, they study the

stability of stationary states, the formation of patterns and the effect of the connectivity η on the resulting dynamics. They carry out these efforts analytically for small graphs of up to four vertices and present numerical results for graphs of up to one hundred vertices.

In [EHS23], the authors studied a natural question posed in [EPSS21]: does one recover the nonlocal interaction equation on $\mathbb{R}^d$ as the graph localizes? In other words, as the range of connections between different vertices decreases, but the weight of each connecting edge increases, do the dynamics in (1.1.2) converge to the dynamics in (1.1.1)? In [EHS23], the authors show that, as the graph localizes, the obtained dynamic is the nonlocal interaction equation on Euclidean space, but with a tensor mobility $\mathbb{T}$. This tensor depends on the geometry of the graph through the connectivity η and the base measure μ. Furthermore, the authors show that the resulting localized PDE is a Riemannian gradient flow on $(\mathcal{P}_2(\mathbb{R}^d_{\mathbb{T}}), d_2)$, where $\mathbb{R}^d_{\mathbb{T}}$ is the d-dimensional Euclidean space endowed with a metric induced by $\mathbb{T}^{-1}$. This result establishes a link between the Finslerian gradient flow structure obtained in [EPSS21] and the Riemannian gradient flow obtained in the limiting Euclidean case. The graph-to-local limit result was extended by the same authors for a system of N nonlocal cross interaction equations on graphs in [EHPS23].

1.2 Graphs and particle systems with heterogeneous interactions

A parallel stream of the literature has considered dynamical systems on graphs, albeit with a different framework than the one presented in the previous section and in this thesis. This approach considers interacting particle systems (IPSs) with heterogeneous interactions, depending on the particles' identities. Heterogeneous interactions arise naturally in contexts such as social dynamics where the interaction between two agents can differ strongly depending on whether they are acquainted or not, for example. Thus, unlike traditional *exchangeable* IPSs where the system is completely invariant under any permutation of the variables' labels, particles are no longer exchangeable.

A natural way to keep track of the identities of the particles and their connections is to use a weighted graph. Namely, the labels can be encoded into a vertex set $V = \{1, \dots, N\}$ and the heterogeneous interactions can be modelled by a family of interaction potentials $(W^{i,j})_{i,j=1}^N$ depending on which pair of particles $(i, j) \in E$ is being considered, where $E = V \times V$ is the set of edges. Consider for example, the following IPS composed of particles $X^i_t \in \mathbb{R}^d$ with labels $i = 1, \dots, N$

$$\begin{cases} \frac{d}{dt} X^i_t = -\frac{1}{N} \sum_{i \neq j}^N \kappa(i,j) \nabla f(X^i_t - X^j_t) \ , \\ X^i_0 \in \mathbb{R}^d, \end{cases} \tag{1.2.1}$$

where the family of potentials $(W^{i,j})_{i,j=1}^N$ is given by $W^{i,j} := \kappa(i,j)f(X_t^i - X_t^j)$, and $\kappa : E \to \mathbb{R}$ encodes the differences in the interactions between the particles depending on their labels. Thus, the vertices of this graph can be viewed as a collection of point masses that move in space, with the trajectory of each point mass being influenced by the rest. We refer to this approach as the *fixed-mass* approach, in contrast with the framework described in the previous section where the positions of the masses are fixed, but the mass is allowed to evolve. We refer to the latter point of view as the *evolving-mass* approach.

The distinction between these approaches has relevant implications in fields such as data science. As we mentioned earlier, when modelling datasets, the vertices can be interpreted as datapoints which are fixed and the weighted edges can encode some measure of similarity such as the distance between these points. Then, for example, the dynamics on this graph can model the behaviour of some algorithm, such as the mean-shift algorithm for clustering [CGTS21]. The mean-shift algorithm is an iterative procedure that attempts to find the modes of a density from a sample by iteratively replacing the "position" of a particle by its mean in some neighbourhood. Note that this mean might not be one of the datapoints in the dataset in the fixed-mass approach framework. As highlighted in [EPSS21], considering this type of dynamics in the graph setting (evolving-mass approach) ensures that the discovered node will be a datapoint.

The study of IPSs on heterogeneous graphs, i.e. not all-to-all, can be traced back to the series of seminal papers by Medvedev [Med14a, Med14b] studying a nonlinear heat-equation on a graph and its continuum limit as the number of vertices (particles) goes to infinity. To the best of the author's knowledge, these contributions were the first to rigorously justify the use of continuum evolution equations, called the *graph limit*, to approximate dynamics on large dense graphs by using *graphons*, a notion of limiting graph for sequences of dense graphs [Lov12]. The graph limit describes the trajectory of an agent with a given label when the number of agents goes to infinity. More precisely, under suitable conditions on the regularity of the interaction potentials, one obtains, following for example, [APD21], that, as $N \to \infty$, the particle system (1.2.1) lifted to a function space, converges in suitable norm to a function $X : [0,T] \times \mathbb{V} \to \mathbb{R}^d$ describing the position of an agent $v \in \mathbb{V}$, where $\mathbb{V}$ is a suitably normalized vertex set (typically $\mathbb{V} = [0,1]$, the continuum generalization of the particles' labels) and solving an integro-differential equation of the following form

$$\partial_t X(t,v) = \int_{\mathbb{V}} \kappa(v,u)\nabla f(X(t,v) - X(t,u))\mathrm{d}u \ . \tag{1.2.2}$$

The notion of graph limit in the context of IPSs has been studied extensively, see for example [PCG23, AD23, GKX23, FJMAT24, Has24, EBFE23, BPCJ24]. We mention that many works study a suitable way of defining the mean-field limit in the case of non-exchangeable particles,

see [JZ23, ADP24, APD21, PT24, AD24, GJKM22, IY23, GKX22] and the references therein. We will revisit this topic in greater detail when we discuss the contents of Chapter 3.

Recently, the study of heterogeneous IPSs on *evolving* as well as *co-evolving* or *adaptive* graphs has started to attract an increasing amount of attention. As its name suggests, in the evolving graph setting, the graph is no longer static, but has time dependent dynamics. Evolving graphs have been used to study technological, social and biological networks. For example, in [IRB+11] the vertices are the patients in a pediatrical ward and the edge-weights depend on the proximity of the patients. This gives rise to a network that evolves in time since the patients can be moved to a different ward. We refer the reader to review papers [HS12, AS14, Hol15] on temporal networks for multiple different models and applications of dynamical networks.

Although considering evolving networks already increases the modelling capabilities, co-evolving or adaptive networks cover even more realistic scenarios. In this setting, the graph's time-evolution is allowed to depend on the dynamics of the system defined on it. Indeed, many problems can be thought of as a dynamical system on a network such that the topology of the network affects the dynamics and the dynamics of the system affect the network's topology. For example, in [GKX23], the authors study a Kuramoto model where the weights determining the coupling of the oscillators is allowed to evolve in time and to depend on the phase of the oscillators themselves. In [APD21], the authors study a model of opinion dynamics on a fully connected graph where the influence of each agent, i.e. the weights of the edges also evolve in time with their dynamics coupled to the opinions' evolution. A non exhaustive list of further examples includes [IMP+24, LPS+24, AJM+24, GKX23, HKT22]. We also refer the reader to the survey article [BGK+23] and the manuscript [Say14] for an introduction to the field of adaptive networks.

To the best of our knowledge, the study of co-evolving graphs in the evolving-mass case has received less attention. As an example we mention [MK19], where the authors study a transport network in which each vertex is assigned a pressure, and the flow between vertices depends on the pressure differential as well as the conductance (weight) of the edge connecting them. The conductances are allowed to evolve in time with their evolution depending positively on pressure differentials hence defining an adaptive network evolution. Part of the contributions of this thesis is to extend the framework of [EPSS21] and [EPS23] to the co-evolving case which, as we shall see, allows for a broader class of edge-weight functions η.

1.3 The contributions of this thesis

1.3.1 Chapter 2: evolution PDEs on co-evolving graphs

Our first contribution on co-evolving graphs, presented in Chapter 2, is the work [EM24], where we extend [EPS23] to the co-evolving graph case. More precisely, we study the well-posedness and certain stability properties of the system

$$\begin{aligned} \partial_t \rho_t &= -\overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]], \\ \partial_t \eta_t &= \omega_t[\rho_t] - \eta_t, \end{aligned} \tag{Co-NCL}$$

where, for $t \in [0,T]$, $\rho_t \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$ describes the (possibly negative) mass on the vertices, $\eta_t \in C_b(\mathbb{R}^{2d}_{\diagup})$ is the, now co-evolving, edge-weight function, V_t is a solution dependent velocity field and F^{Φ} is the flux, which depends on an admissible interpolation function Φ satisfying certain Lipschitz continuity properties. Unlike in previous contributions, the edge-weight function η also evolves in time and its dynamics depend on the dynamics of the mass defined on the network through the (potentially non-local) function $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$. We obtain well-posedness of weak solutions (in duality with $C_0(\mathbb{R}^d)$ functions) to (Co-NCL) using the Banach fixed-point theorem in the metric space $(C([0,T], \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)) \times C([0,T], C_b(\mathbb{R}^{2d}_{\diagup})),\ d_\infty)$, where $d_\infty = \|\cdot\|_{\infty, \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} + \|\cdot\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})}$. We summarise the result in the following theorem.

Theorem 1.3.1 (Lemma 2.3.1, Theorem 2.3.1)**.** Let V and ω satisfy suitable continuity and boundedness assumptions. Then, there exists a unique solution (ρ, η) to (Co-NCL) with initial datum (ρ^0, η^0) on $[0,T]$.

Next, we consider the dynamics of (Co-NCL) when the graph η evolves at a slower timescale than the mass ρ, that is, we consider the problem

$$\begin{cases} \partial_t \rho_t = -\overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]] \\ \partial_t \eta_t = \varepsilon(\omega_t[\rho_t] - \eta_t) \\ \rho_0 \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d),\ \eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})\ , \end{cases} \tag{Co$-$NCL$_{\mathrm{S}}$}$$

where $\varepsilon > 0$ is a small parameter. Formally, one can see that (Co$-$NCL$_{\mathrm{S}}$) can be obtained from (Co-NCL) by considering η to evolve at a time scale $\tau = \varepsilon t$ and then re-expressing the evolution of η for time scale t. Then, we obtain the following statement to characterize the limit of (Co$-$NCL$_{\mathrm{S}}$) as $\varepsilon \to 0^+$.

Theorem 1.3.2 (Theorem 2.4.1)**.** Let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ a vanishing sequence, that is $\varepsilon_n \to 0^+$. Consider the sequence of solutions $\{(\rho^n, \eta^n)\}_{n\in\mathbb{N}}$ to (Co$-$NCL$_{\mathrm{S}}$) with $\rho^n_0 = \rho^0 \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$, and

$\eta_0^n = \eta^0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. Let V and ω satisfy suitable assumptions. Then it holds that

$$d_\infty((\rho^n, \eta^n), (\overline{\rho}, \eta^0)) \to 0, \text{ as } n \to \infty,$$

where

$$\overline{\rho}_t + \int_0^t \overline{\nabla} \cdot F^\Phi[\mu, \eta^0; \overline{\rho}_s, V_s(\overline{\rho}_s)]\mathrm{d}s = \rho^0 \ ,$$

to be understood in duality with $C_0(\mathbb{R}^d)$.

A further natural problem is to study the case in which η evolves at a faster time scale than ρ, i.e.

$$\begin{cases} \partial_t \rho_t = -\overline{\nabla} \cdot F^\Phi[\mu, \eta_t; \rho_t, V_t[\rho_t]] \\ \varepsilon \partial_t \eta_t(x,y) = \omega_t[\rho](x,y) - \eta_t(x,y), \\ \rho_0 \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d), \ \eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup}) \ . \end{cases} \tag{Co$-$NCL$_{\mathrm{F}}$}$$

This problem can be obtained, formally, from (Co-NCL) by considering η evolving at a time scale $\tau = \frac{t}{\varepsilon}$ and then re-writing it at time scale t. The statement characterising the limit of (Co$-$NCL$_{\mathrm{F}}$) as $\varepsilon \to 0^+$ is as follows.

Theorem 1.3.3 (Theorem 2.4.2)**.** Let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ such that $\varepsilon_n \to 0^+$ as $n \to \infty$, and consider a sequence of solutions $\{(\rho^n, \eta^n)\}_{n\in\mathbb{N}}$ to (Co$-$NCL$_{\mathrm{F}}$) with $\eta_0^n \in C_b(\mathbb{R}^{2d}_{\diagup})$ and $\rho_0^n \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$ satisfying $\|\eta_0^n - \omega_0[\rho_0^n]\|_\infty \to 0$, as $n \to \infty$. If V and ω satisfy suitable continuity and boundedness assumptions, then

$$d_\infty((\rho^n, \eta^n), (\rho, \omega)) \to 0 \text{ as } n \to \infty \ ,$$

where, for a.e. $t \in [0,T]$, ρ solves

$$\rho_t + \int_0^t \overline{\nabla} \cdot F^\Phi[\mu, \omega_s[\rho_s]; \rho_s, V_s[\rho_s]]\mathrm{d}s = \rho_0 \ , \tag{1.3.1}$$

to be understood in duality with $C_0(\mathbb{R}^d)$.

Finally, we obtain a discrete-to-continuum limit result in the case $\Phi \equiv \Phi_{upwind}$, for simplicity, although it holds for a wider class of interpolations which we specify later in Chapter 2. More precisely, we show that for a sequence of graphs with an increasing number of vertices, given by a sequence of base measures μ^n converging in a suitable sense to a base measure μ^*, the associated sequence of problems (Co-NCL) converge to the problem on the limiting graph given by the limit base measure μ^*. The statement is as follows

Theorem 1.3.4 (Theorem 2.5.1)**.** Fix $\Phi \equiv \Phi_{\text{upwind}}$ and consider a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^+_{\text{TV}}(\mathbb{R}^d)$ such that $\mu^n \xrightarrow{*} \mu \in \mathcal{M}^+_{\text{TV}}(\mathbb{R}^d)$. Let V and ω satisfy suitable continuity and boundedness assumptions. Let us consider a sequence of solutions $\{(\rho^n, \eta^n)\}_{n\in\mathbb{N}}$ to (Co-NCL) associated to $\{\mu_n\}$, and let (ρ, η) be the solution to (Co-NCL) depending on μ. If $\|\rho_0^n - \rho^0\|_{\text{TV}} \to 0$ and $\|\eta_0^n - \eta_0\|_\infty \to 0$ as $n \to \infty$, then

$$\lim_{n\to\infty} d_\infty((\rho^n, \eta^n), (\rho, \eta)) = 0 \ .$$

1.3.2 Chapter 3: the graph continuity equation

In Chapter 3, we link (Co-NCL) to a particular nonlinear continuity equation on $\mathbb{R}^{d+1}$ which we shall call the *graph continuity equation* and is relevant as a stepping stone to understanding discrete-to-continuum limits. Indeed, similar continuity equations have been obtained in the study of the mean-field limit of IPSs with heterogeneous interactions, see for example [KVM18, CM19a, CM19b, Kue20, GK22, KX22] for static graphs, although this list is not exhaustive. In the case of co-evolving graphs, we highlight the recent work [GKX22] on the Kuramoto model on a co-evolving graph, where the authors prove the mean-field limit and provide several examples both for dense as well as sparse graphs. These works rely on the recent theory of graph limits [Lov12, BS22]. We also mention [ADP24, PT24] for recent reviews on understanding the different approaches to defining the mean-field limit of an IPS on graphs. In this sense, the work in this chapter can also be viewed as an application of the ideas in [PT24, Section 5.2.3] on how to obtain a continuity equation starting from the solution of a nonlinear Euler equation, although our setting is not included in [PT24].

A common point in many of the aforementioned works and the one presented in this chapter is the approach to obtain the well-posedness of the nonlinear continuity equation. Namely, the well-posedness is obtained by studying a family of characteristic equations parametrized by the vertices of the underlying graph extending the classical work of Neunzert [Neu84]. To the best of the author's knowledge, the first application of this approach in the context of all-to-all graphs can be found in [Lan05] and, in the case of heterogeneous graphs can be found in [KVM18].

Throughout this chapter we will restrict μ to be a probability measure with finite second moment. To establish the link between (Co-NCL) and the graph continuity equation, we first study the following version of (Co-NCL) in the case $\rho \ll \mu$ with $r = \frac{\mathrm{d}\rho}{\mathrm{d}\mu}$, under suitable assumptions on Φ and η

$$\begin{aligned} \partial_t r_t(x) &= -\int_{\mathbb{R}^d\setminus\{x\}} \Phi(r_t(x), r_t(y); V_t[r_t](x,y))\eta_t(x,y)\mathrm{d}\mu(y) \ , \\ \partial_t \eta_t &= \omega_t[r_t] - \eta_t, \end{aligned} \tag{Euler Co-NCL}$$

for given initial data $r^0 \in L^2_\mu(\mathbb{R}^d)$ and $\eta^0 \in C_b(\mathbb{R}^{2d}_{\nearrow})$. This is a nonlinear Euler equation motivating the name (Euler Co-NCL). We prove well-posedness of this equation in Theorem 3.1.1 using the Banach fixed-point theorem.

To obtain the graph continuity equation, we lift (Euler Co-NCL) to a space of measures by considering $\sigma := \delta_r \otimes \mu$ given by

$$(\delta_r \otimes \mu)(A \times B) = \int_B \delta_{r(x)}(A) \mathrm{d}\mu(x), \tag{1.3.2}$$

on any Borel set $A \times B \in \mathcal{B}(\mathbb{R} \times \mathbb{R}^d)$, where (r, η) is the solution of (Euler Co-NCL) with base measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then, we can show that σ solves the following continuity equation

$$\begin{cases} \partial_t \sigma + \partial_\xi(\sigma \mathcal{X}[\sigma, r, \eta]) = 0, \\ \sigma_0 = \bar{\sigma} \in \mathcal{P}(\mathbb{R} \times \mathbb{R}^d), \end{cases} \tag{graph-CE}$$

where, for any $\sigma \in \mathcal{P}(\mathbb{R} \times \mathbb{R}^d)$, the velocity field is given by

$$\mathcal{X}[\sigma, r, \eta](t, \xi, x) := -\int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(\xi, \xi'; V_t[r_t](x, x'))\eta_t(x, x') \mathrm{d}\sigma(\xi', x') \ , \tag{1.3.3}$$

for a solution (r, η) to (Euler Co-NCL). We refer to (graph-CE) as the *graph-continuity equation* since the velocity field depends on the solution of the co-evolving graph problem. As the base measure μ does not depend on time, we obtain the well-posedness of (graph-CE) by first considering its *disintegrated version*

$$\begin{cases} \partial_t \sigma_{t,x} + \partial_\xi(\sigma_{t,x} \mathcal{X}[\sigma, r, \eta]) = 0, \\ \sigma_{0,x} = \bar{\sigma}_x, \end{cases} \tag{1.3.4}$$

for μ-a.e. $x \in \mathbb{R}^d$, where $\sigma_t = \int_{\mathbb{R}^d} \sigma_{t,x} \mathrm{d}\mu(x)$ and $\bar{\sigma} = \int_{\mathbb{R}^d} \bar{\sigma}_x \mathrm{d}\mu(x)$. We solve (graph-CE) in the space $\mathcal{P}^\mu_2(\mathbb{R} \times \mathbb{R}^d)$ of probability measures with finite second moments whose second marginal is μ, equipped with the distance

$$L^2_\mu d_2(\nu^1, \nu^2) := \left(\int_{\mathbb{R}^d} d_2^2(\nu^1_x, \nu^2_x) \mathrm{d}\mu(x) \right)^{\frac{1}{2}} \ , \quad \nu^1, \nu^2 \in \mathcal{P}^\mu_2(\mathbb{R} \times \mathbb{R}^d) \ ,$$

which is a complete metric space, see Proposition A.0.1. After a thorough study of some properties of the characteristics associated to (1.3.4), we consider the flow-maps given by the

solution to

$$\begin{aligned} &\frac{d}{dt} f_x[\sigma, r, \eta](t, u) = \mathcal{X}[\sigma, r, \eta](t, f_x[\sigma, r, \eta](t, u), x) \ , \\ &f_x[\sigma, r, \eta](0, u) = u \in \mathbb{R} \ , \end{aligned} \tag{1.3.5}$$

for μ-a.e. $x \in \mathbb{R}^d$. We summarise the statement on the well-posedness of (graph-CE) in the following theorem.

Theorem 1.3.5 (Theorem 3.2.1)**.** Let $V : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ be continuous in time and satisfy suitable boundedness and integrability assumptions. Let (r, η) be the solution of (Euler Co-NCL). Then, there exists a unique solution $\sigma \in C([0,T], \mathcal{P}^\mu_2(\mathbb{R} \times \mathbb{R}^d))$ to (graph-CE) with initial condition $\nu \in \mathcal{P}^\mu_2(\mathbb{R} \times \mathbb{R}^d)$ given by $\sigma = \int_{\mathbb{R}^d} \sigma_{t,x} \mathrm{d}\mu(x)$ where $\sigma_{t,x} = f_x[\sigma, r, \eta](t, \cdot)_\# \nu_x$ for μ-a.e. $x \in \mathbb{R}^d$.

Furthermore, we obtain the following stability result with respect to the initial data.

Theorem 1.3.6 (Theorem 3.2.2)**.** Assume $V : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ to be continuous in time and to satisfy (3.1.9a) and (3.1.9b). Let (r, η) be the solution to (Euler Co-NCL) with initial data $(r_0, \eta_0) \in L^2_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ and let $\sigma^i \in C([0,T], \mathcal{P}^\mu_2(\mathbb{R} \times \mathbb{R}^d))$ be the solutions to the corresponding (graph-CE) with initial conditions $\tilde{\sigma}^i_0 \in \mathcal{P}^\mu_2(\mathbb{R} \times \mathbb{R}^d)$ for $i = 1, 2$. Then we have the following stability estimate

$$L^2_\mu d^2_2(\sigma^1_t, \sigma^2_t) \leqslant 2 e^{\tilde{C}_T + 2C(T)T^2} L^2_\mu d^2_2(\tilde{\sigma}^1_0, \tilde{\sigma}^2_0) \ ,$$

where $\tilde{C}_T$ and $C(T) > 0$ are constants depending on η_0.

We conclude the chapter by commenting on the difficulties faced to obtain a discrete-to-continuum limit and the long-time behaviour of (Co-NCL).

1.3.3 Chapter 4: long-time behaviour

In Chapter 4, overcoming the difficulties faced with the previous approach, we study the long-time behaviour of (Euler Co-NCL) for a particular class of velocities. More precisely, we consider velocity fields that are pointwise, i.e., that can be written as $V_t[r](x, x') = V_t(r_t(x), r_t(x'))$, for any vertices $x, x' \in \mathbb{R}^d$ and *monotonic*, that is, velocities that satisfy the following definition.

Definition 1.3.1. We say that a pointwise velocity $V : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ is *monotonic* if, for $r \in L^\infty_\mu(\mathbb{R}^d)$, any $t \geqslant 0$, and μ-a.e. $(x, x') \in \mathbb{R}^{2d}_{\diagup}$, we have that:

- $r(x) > r(y)$, then $V^+_t(r(x), r(x')) > V^+_t(r(y), r(x'))$,
- $r(x) < r(y)$, then $V^-_t(r(x), r(x')) > V^-_t(r(y), r(x'))$,

- $r(x) = r(y)$, then $V_t(r(x), r(x')) = V_t(r(y), r(x'))$,
- $r(x) = r(x')$, then $V_t(r(x), r(x')) = 0$.

This definition captures the intuitive idea that the mass flows in the direction of the negative discrete gradient of the mass between vertices. Furthermore, with upwind dynamics, the first condition ensures that a vertex with more mass will transmit more mass to its neighbours than vertices with less mass. In view of the pointwise nature of the velocity fields, we obtain the well posedness of (Euler Co-NCL) for $r \in L^\infty_\mu(\mathbb{R}^d)$ by the Banach-fixed point theorem in Theorem 4.2.1. We then move to the study of the long-time behaviour of (Euler Co-NCL) in this framework. To this end, we restrict the co-evolving graph to be defined on a compact set $K \subset \mathbb{R}^d$ and fix the upwind interpolation Φ_{upwind}. We carry over the notation used on $\mathbb{R}^d$ to this setting, and thus, $K^2_{\not=} := \{x, x' \in K \mid x \neq x'\}$. We obtain the long-time behaviour by showing that, as one would expect by the monotonicity of the velocities, the supremum norm of r decreases monotonically in time, while the infimum increases montonically in time, resulting in a stabilization of the dynamics at $\frac{M}{\mu(K)}$, the uniform mass distribution, where $\int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) = M$.

The approach to solve this problem follows similar ideas to [BDS22] for obtaining the long-time behaviour of large systems of interacting agents on time-evolving graphs. Indeed, the long-time behaviour of the mass dynamics on a co-evolving graph is linked to the problem of *consensus formation* for systems of *non-exchangeable* interacting particles. More precisely, the consensus formation problem is to determine whether there exists an element x^∞ such that

$$\lim_{t\to\infty} |x_i(t) - x^\infty| = 0$$

for any agent $i \in \{1, \dots, N\}$ in an IPS. In our case, we obtain that $x^\infty = \frac{M}{\mu(K)}$, the uniform mass distribution. Consensus formation has received a vast amount of attention, see for example [BHOT05, CFPT15, CFPT13, DM11, Mor05, OSM04, WT21] and the references therein. The literature on this topic is very extensive as techniques have to be adapted to the particular type of interaction and connectivity in each case. Our approach is close to the works that tackle the problem by obtaining *diameter* decay estimates. We obtain that the rate at which the "vertex with maximum mass" ("vertex with minimum mass") loses (gains) mass depends on the connectivity of the system through the quantities η_*, the smallest connectivity for any vertex in the history of the co-evolving graph, and α'_*, the smallest rate of outflow for the velocities. This is closely related to the *scrambling-coefficient* [Sen79] obtained for example in [BDS22], a quantity which is positive if either a pair of agents is interacting or they follow a common third agent. Our condition is stronger, since we impose that the graph is strongly connected, i.e. $\eta_t > 0$ for $t \geqslant 0$. This is due to the co-evolving nature of η which is not present in most of

the previously mentioned works, where the connectivity of the topology is, at most, allowed to evolve in time, but not to depend on the dynamics of the agents as in our case. Finding weaker conditions for η yielding convergence to consensus could be achieved upon imposing greater structure on the function ω, but this is left for future investigation.

To obtain the result, we further restrict to a class of bounded montonic pointwise velocities of the form

$$V(r_t(x), r_t(x')) = \alpha(r_t(x)) - \alpha(r_t(x'))$$

where $\alpha \in C_b^1(\mathbb{R}^d)$ and $\alpha'_* := \inf_{x\in K} \alpha'(x)$, where $\alpha'(y) = \frac{d}{dy}\alpha(y)$. We also use the notation

$$\eta_* := \inf_{t\in[0,\infty)} \inf_{(x,x')\in K^2_{\diagup}} \eta_t(x, x') \ ,$$

to denote the smallest connectivity in the history of the co-evolving graph and $r_{t,*} := \inf_{x\in K} r_t(x)$. The decay/gain results are obtained in Theorem 4.3.3 and 4.3.4, respectively. We summarise these findings in the following theorem

Theorem 1.3.7 (Theorem 4.3.3 and Theorem 4.3.4)**.** Fix the interpolation $\Phi \equiv \Phi_{upwind}$. Let $(r, \eta) \in AC(\mathbb{R}^+, L^\infty_\mu(K)) \times AC(\mathbb{R}^+, C_b(K^2_{\diagup}))$ be the solution of (Euler Co-NCL) with a pointwise monotonic velocity field $V : \mathbb{R}^+ \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathcal{V}^{as}(K^2_{\diagup})$ and $\omega : [0,\infty) \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathbb{R}$, satisfying suitable assumptions. Assume $r_0 \geqslant 0$ with $\int_K r_0(x)\mathrm{d}\mu(x) = M$ and that $\eta^0 \in C_b(K^2_{\diagup})$ is symmetric and positive. Then, we have

$$\|r_t\|_{L^\infty_\mu(K)} \leqslant \frac{\|r_0\|_{L^\infty_\mu(K)}\alpha'_*\eta_* M e^{\alpha'_*\eta_* Mt}}{\alpha'_*\eta_*\left[\|r_0\|_{L^\infty_\mu(K)}\mu(K)(e^{\alpha'_*\eta_* Mt} - 1) + M\right]} \ , \tag{1.3.6}$$

and

$$r_{t,*} \geqslant \frac{r_{0,*}\eta_*\alpha'_* M e^{\eta_*\alpha'_* Mt}}{\eta_*\alpha'_*(M + (e^{\eta_*\alpha'_* Mt} - 1)\mu(K)r_{0,*})} \ . \tag{1.3.7}$$

Using the previous result, the long-time behaviour of (Euler Co-NCL) is an easy consequence obtained in Theorem 4.3.5 which we present below.

Theorem 1.3.8 (Theorem 4.3.5)**.** Under the assumption of Theorem 1.3.7, it holds that for μ-a.e. $x \in \mathbb{R}^d$

$$\lim_{t\to\infty} r_t(x) = \frac{M}{\mu(K)} \ ,$$

the uniform distribution over K.

1.4 Perspectives and future work

There are multiple open questions left for future work and investigation. First, it would be interesting to analytically study the long-time behaviour and steady states of (Co-NCL) under broader conditions for V and η in the infinite graph case beyond the diffusive case considered in this work. More precisely, for an interaction potential K, it would be interesting to give a characterisation of conditions on K and co-evolving η leading to aggregation. Furthermore, obtaining conditions that guarantee the contractivity of solutions is also an open direction of enquiry. It would be also interesting to try to obtain a gradient flow interpretation for (Co-NCL) for suitable classes of energies. This is not a trivial question as this coupled equation has a different gradient structure than the ones considered in previous work.

Following from the discussion in Section 3.2.1 of Chapter 3, obtaining a characterisation of (graph-CE) as the discrete-to-continuum limit of sequences of (Euler Co-NCL) depending on increasing base measures is still a very interesting open problem. Indeed, studying (graph-CE) in $(\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d), L^2_{\mu}d_2)$ is advantageous as it allows us to have less restrictive assumptions on the regularity of the maps $(x,x')\mapsto V_\cdot[\cdot](x,x')$ and $(x,x')\mapsto \omega_\cdot[\cdot](x,x')$. However, by definition of $\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d)$, we can not consider a sequence of problems (graph-CE) associated to a sequence of base measures $\{\mu^n\}$ on the same space. A potential way to circumvent this issue would be to consider the space $\mathcal{P}_2(\mathbb{R}\times\mathbb{R}^d)$ and equip it with a different distance that would not require the additional regularity of V and η, but this is left for future work.

Investigating the case of multiple species on a co-evolving graph, where the cross-interactions not only affect the mass dynamics, but also the dynamics of the graph is also an appealing area of study.

As discussed in Section 1.1, the nonlocal interaction equation on graphs gives rise to non-trivial interaction equation in the local limit. While this problem has been successfully tackled in [EHS23], obtaining quantitative convergence estimates is still an open question.

Other interesting directions include the study of an optimal control problems by including a control in the evolution of the edge-weight function η, as well as considering a time dependent base measure μ, for different models capturing the possible growth of a network.

Chapter 2

On evolution PDEs on co-evolving graphs

2.1 Introduction

We consider a class of nonlocal continuity equations on co-evolving graphs extending the results in [EPSS21, EPS23], where the underlying graph is, instead, static. More precisely, in [EPSS21, EPS23] the authors consider equations on the time interval $[0, T]$ of the form

$$\partial_t \rho_t + \overline{\nabla} \cdot \boldsymbol{j}_t = 0, \tag{2.1.1a}$$

with nonlocal divergence (cf. Defintion 2.2.1)

$$\overline{\nabla} \cdot \boldsymbol{j}_t(\mathrm{d}x) = \int_{\mathbb{R}^d \setminus \{x\}} \eta(x, y) \mathrm{d}\boldsymbol{j}_t(x, y),$$

for a time-dependent measure flux, $\boldsymbol{j}_t \in \mathcal{M}(G)$, on the set of edges defined by an edge weight function $\eta : \mathbb{R}^d \times \mathbb{R}^d \backslash \{x = y\} \to [0, \infty)$, i.e. $G = \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : x \neq y \text{ and } \eta(x, y) > 0\}$; throughout the chapter we set $\mathbb{R}^{2d}_{\not\Delta} := \mathbb{R}^d \times \mathbb{R}^d \backslash \{x = y\}$. In particular, points in $\mathbb{R}^d$ are possible vertices and the edges, G, are defined through the function η. Equation (2.1.1a) describes the time-evolution of a probability measure, ρ_t, representing the mass at a vertex $x \in \mathbb{R}^d$, according to a nonlocal continuity equation on the graph. The variation of the mass on each vertex is given as an average of all possible outgoing and ingoing fluxes. We note this is a substantial difference with equations on $\mathbb{R}^d$, where one usually describes the flow of moving particles with a given mass. On graphs, particles are vertices which are fixed, in contrast to the mass which is transported along the edges. In the graph setting, fluxes and velocities, v, are defined on the edges, whereas the mass is a vertex-based quantity. Therefore, one needs to consider a suitable

interpolation of the mass at the vertices in order to have an edge-based quantity for the mass as well. Hence, on graphs, the relation between flux and velocity can be nonlinear as it depends on the interpolation function chosen, denoted by Φ, as well as on the vertices, which are represented by a measure $\mu \in \mathcal{M}^+(\mathbb{R}^d)$. In view of these considerations, (2.1.1a) is complemented with the constitutive equation

$$\boldsymbol{j} = F^{\Phi}[\mu; \rho, v], \tag{2.1.1b}$$

for admissible interpolation functions specified in Definition 2.2.3 and Example 2.2.4. In particular, admissible fluxes F^{Φ} (see Definition 2.2.5) depend on the interpolation function chosen as in (2.1.1b).

We consider a scenario where the connection of the graph evolves in time, possibly depending on the dynamics of the mass on the edges determined by the interpolation function. Specifically, for $\eta : [0,T] \times (\mathbb{R}^d \times \mathbb{R}^d \backslash \{x = y\}) \to \mathbb{R}$, we shall study the following class of systems

$$\begin{aligned} &\partial_t \rho_t + \overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]] = 0, \\ &\partial_t \eta_t(x,y) = H(t,x,y,\eta_t,\rho_t), \end{aligned} \tag{2.1.2}$$

for an admissible flux F^{Φ} satisfying Definition 2.2.5. The velocity field is given as $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$ (set of antisymmetric velocity fields), including possible dependence on the mass ρ. Among others, a driving example is given by

$$V_t[\rho_t](x,y) = -\overline{\nabla}(K * \rho_t)(x,y) = \int_{\mathbb{R}^d} K(x,z) - K(y,z) \mathrm{d}\rho_t(z) \ ,$$

for $K : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ being an interaction potential, studied in [EPSS21, EHS23, EHPS23]. We shall focus on

$$H(t,x,y,\eta_t,\rho_t) := \omega_t[\rho_t](x,y) - \eta_t(x,y),$$

where $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ is a function satisfying a set of assumptions specified in Section 2.3. With this choice for the function H the evolution of the graph's weights is influenced by two processes. First, the connection between two vertices can depend (in a nonlocal way) on the mass configuration on the graph, as well as on edges $(x,y) \in G$ and time. A possible example for the function ω could be a nonlocal functional on $\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$, depending on the edge considered as well as on time, such as

$$\omega_t[\sigma](x,y) = \int_{\mathbb{R}^d} K(t,x,y,z) \mathrm{d}\sigma(z), \tag{2.1.3}$$

where $K \in C_b([0,T] \times \mathbb{R}^{2d}_{\diagup} \times \mathbb{R}^d)$ is a convolution kernel. The second term in the definition H

can be thought of as a relaxation effect in time, since, in case the first term is constant, the weight function would converge to a steady state. A similar choice is also considered, e.g., in the recent work [Bur22] in a different framework of co-evolving networks. Under these assumptions, system (2.1.2) reads

$$\begin{aligned} \partial_t \rho_t &= -\overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]], \\ \partial_t \eta_t &= \omega_t[\rho_t] - \eta_t, \end{aligned} \tag{Co-NCL}$$

for given initial data $\rho^0 \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$ and $\eta^0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. The acronym NCL stands for nonlocal conservation law as we allow the velocity field to depend on the solution itself. We add the prefix "Co", standing for "co-evolving", to make clear we consider coupled dynamics for the weight function. Differently from previous contributions in the literature, we allow the weight function to become negative and not be symmetric. We analyse different time scales and note that these can lead to weight functions depending on the mass configuration at the vertices, which have not been explored in depth so far, to the best of our knowledge. More precisely, when the graph evolves faster than the mass on the vertices, one obtains $\eta_t = \omega_t[\rho]$, leading to a possible further nonlocality in the flux provided by η, in case of ω as in (2.1.3). In particular, the PDE takes the form

$$\partial_t \rho_t = -\overline{\nabla} \cdot F^{\Phi}[\mu_t, \omega_t[\rho_t]; \rho_t, V_t[\rho_t]].$$

In this scenario we still talk about *co-evolving* graphs, though edge weights only change according to the mass configuration and not a coupled dynamics. Furthermore, we also provide a discrete-to-continuum limit for solutions of (Co-NCL) when the number of vertices tends to infinity. In particular, we rigorously justify the continuum equation on the graph as the many vertices limit of the discrete, finite vertices, model.

Structure of the chapter

We introduce the set up of the problem in Section 2.2, including preliminary results. In Section 2.3, we prove well-posedness for (Co-NCL) by means of a fixed-point argument. Section 2.4 is devoted to the analysis of different time scales for (Co-NCL) distinguishing between the graph evolving at a faster or slower rate than the dynamics of the weight function. We conclude the chapter with a discrete-to-continuum limit for (Co-NCL) in Section 2.5.

2.2 Preliminaries

A graph is identified by a pair (μ, η): $\mu \in \mathcal{M}^+(\mathbb{R}^d)$ is referred to as *base measure* and $\eta : \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ is the *weight function*, the analogue of the weighted edges for a discrete finite graph. We define

the set of edges as $G := \{(x,y) \in \mathbb{R}^{2d}_{\not\sim} : \eta(x,y) \neq 0\}$. More precisely, the pair (μ, η) defines a weighted graph. In previous works, the weight function η is sometimes non-negative, whilst in this chapter we allow for possible negative values. Furthermore, we do not restrict to symmetric weights.

The mass configuration at the vertices is described by a measure with finite total variation, i.e. $\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$, as defined in the notation list. We equip the set $\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$ with the total variation norm:

$$\|\sigma\|_{\mathrm{TV}} = |\sigma|[\mathbb{R}^d] = \sup\left\{\langle \varphi, \sigma\rangle \ : \ \varphi \in C_0(\mathbb{R}^d),\ \|\varphi\|_\infty \leqslant 1\right\},$$

for any $\sigma \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$, where $\langle \varphi, \sigma\rangle := \int_{\mathbb{R}^d} \varphi \,\mathrm{d}\sigma$ the dual product between $C_0(\mathbb{R}^d)$ and $\mathcal{M}(\mathbb{R}^d)$. The dynamics we consider preserve the mass, as well as the bound on the total variation, therefore, we shall consider the set

$$\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d) := \left\{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) : |\rho|(\mathbb{R}^d) \leqslant M\right\}.$$

As we are not in the traditional Euclidean setting, we recall the notion of gradient and divergence for a function defined on graphs, similar to [EPS23]; the only difference is that we do not restrict fluxes on the set G, as this is changing in time, but rather consider their (possible) extension to $\mathbb{R}^{2d}_{\not\sim}$, cf. Remark 2.2.1.

Definition 2.2.1 (Nonlocal gradient and divergence)**.** For any $\phi : \mathbb{R}^d \to \mathbb{R}$, we define its nonlocal gradient $\overline{\nabla}\phi : \mathbb{R}^{2d}_{\not\sim} \to \mathbb{R}$ by

$$\overline{\nabla}\phi(x,y) = \phi(y) - \phi(x), \qquad \text{for all } (x,y) \in \mathbb{R}^{2d}_{\not\sim}.$$

For any Radon measure $\boldsymbol{j} \in \mathcal{M}(\mathbb{R}^{2d}_{\not\sim})$, its nonlocal divergence $\overline{\nabla} \cdot \boldsymbol{j} \in \mathcal{M}(\mathbb{R}^d)$ is defined as the adjoint of $\overline{\nabla}$, i.e., for any $\phi : \mathbb{R}^d \to \mathbb{R}$ in $C_0(\mathbb{R}^d)$, there holds

$$\begin{aligned}\int_{\mathbb{R}^d} \phi \mathrm{d}\overline{\nabla} \cdot \boldsymbol{j} &= -\frac{1}{2}\iint_{\mathbb{R}^{2d}_{\not\sim}} \overline{\nabla}\phi(x,y)\mathrm{d}\boldsymbol{j}(x,y) \\ &= \frac{1}{2}\int_{\mathbb{R}^d} \phi(x) \int_{\mathbb{R}^d\setminus\{x\}} (\mathrm{d}\boldsymbol{j}(x,y) - \mathrm{d}\boldsymbol{j}(y,x)).\end{aligned}$$

In particular, for $\boldsymbol{j}$ antisymmetric, that is, $\boldsymbol{j} \in \mathcal{M}(\mathbb{R}^{2d}_{\not\sim})$ and $\mathrm{d}\boldsymbol{j}(x,y) = -\mathrm{d}\boldsymbol{j}(y,x)$, denoted $\boldsymbol{j} \in \mathcal{M}^{\mathrm{as}}(\mathbb{R}^{2d}_{\not\sim})$, we have

$$\int_{\mathbb{R}^d} \phi \mathrm{d}\overline{\nabla} \cdot \boldsymbol{j} = \iint_{\mathbb{R}^{2d}_{\not\sim}} \phi(x)\mathrm{d}\boldsymbol{j}(x,y)\ .$$

Solutions of the nonlocal continuity equations are intended as curves defined on a time interval $[0,T]$, for $T>0$. Therefore, we shall denote by $\mathrm{AC}([0,T];\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$ the set of curves from $[0,T]$ to $\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$ absolutely continuous with respect to the TV norm, that is, the set of curves $\rho\colon [0,T]\to\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$ such that there exists $m\in L^1([0,T])$ with

$$\|\rho_t-\rho_s\|_{\mathrm{TV}}\leqslant\int_s^t m(r)\mathrm{d}r,\qquad\text{for all } 0\leqslant s<t\leqslant T.$$

With these preliminary notions we specify the definition of weak solution for the nonlocal continuity equation, referred to as (NCE).

Definition 2.2.2 (Weak solution for the NCE)**.** A measurable pair of curves $(\rho,\boldsymbol{j})\colon[0,T]\to\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)\times\mathcal{M}(\mathbb{R}^{2d}_{\diagup})$ is a *weak* solution to the nonlocal continuity equation, denoted as

$$\partial_t\rho+\overline{\nabla}\cdot\boldsymbol{j}=0,\tag{NCE}$$

provided that, for any $\varphi\in C_0(\mathbb{R}^d)$, it holds:

(i) $(t\mapsto\rho_t)\in\mathrm{AC}([0,T];\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$;

(ii) $(\boldsymbol{j}_t)_{t\in[0,T]}$ is Borel measurable and the map $(t\mapsto\langle\varphi,\overline{\nabla}\cdot\boldsymbol{j}_t\rangle)\in L^1([0,T])$;

(iii) $(\rho,\boldsymbol{j})$ satisfies

$$\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_t(x)-\frac{1}{2}\int_0^t\iint_{\mathbb{R}^{2d}_{\diagup}}\overline{\nabla}\varphi(x,y)\mathrm{d}\boldsymbol{j}_s(x,y)\mathrm{d}s=\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_0(x)\qquad\text{for a.e. } t\in[0,T];$$

in this case, we write $(\rho,\boldsymbol{j})\in\mathrm{NCE}_{\mathrm{T}}$.

For the sake of completeness, we specify that absolute continuity of ρ is guaranteed by the integrability of the flux divergence. The nonlocal continuity equation above is the driving dynamics we consider in this work, already considered, e.g., in [EPSS21, EPS23], though the concept of solution is slightly different, as it is given using duality between C_0 and signed Radon measures. The latter is more suitable when dealing with the discrete-to-continuum limit (Section 2.5).

As discussed in the introduction, in order to specify the evolution on graphs we need to suitably define the flux, F^Φ, which is an edge-based quantity, as well as the velocities. As the mass is solely vertex-based, one has to interpolate the quantities located at the vertices to obtain an edge-based quantity. As a byproduct, the relation between the flux and the velocity can vary, depending on the application. We specify below interpolations and fluxes used in this chapter, following [EPS23], allowing for a linear dependence on the weight function, η, in the flux.

Definition 2.2.3 (Admissible flux interpolation)**.** A measurable function $\Phi : \mathbb{R}^3 \to \mathbb{R}$ is called an admissible flux interpolation provided that the following conditions hold:

(i) Φ satisfies

$$\Phi(0,0;v) = \Phi(a,b;0) = 0, \qquad \text{for all } a,b,v \in \mathbb{R};$$

(ii) Φ is argument-wise Lipschitz in the sense that, for some $L_\Phi > 0$, any a, $b, c, d, v, w \in \mathbb{R}$, it holds

$$|\Phi(a,b;w) - \Phi(a,b;v)| \leqslant L_\Phi(|a| + |b|)|w - v|;$$
$$|\Phi(a,b;v) - \Phi(c,d;v)| \leqslant L_\Phi(|a-c| + |b-d|)|v|;$$

(iii) Φ is positively one-homogeneous in its first and second arguments, that is, for all $\alpha > 0$ and $(a,b,w) \in \mathbb{R}^3$, it holds

$$\Phi(\alpha a, \alpha b; w) = \alpha \Phi(a,b;w).$$

Example 2.2.4. In [EPSS21], the authors consider the upwind interpolation given by

$$\Phi_{\text{upwind}}(a,b;w) = aw_+ - bw_-, \qquad \text{for } (a,b,w) \in \mathbb{R}^3.$$

Another example is provided by the mean multipliers

$$\Phi_{\text{prod}}(a,b;w) = \phi(a,b)w, \qquad \text{for } (a,b,w) \in \mathbb{R}^3,$$

where common choices for ϕ include: $\phi(a,b) = \frac{a+b}{2}$, or $\phi(a,b) = \max\{a,b\}$.

The fluxes considered are defined as follows.

Definition 2.2.5 (Admissible flux)**.** Let Φ be an admissible flux interpolation, and let $\rho \in \mathcal{M}_{\text{TV}}(\mathbb{R}^d)$, $w \in \mathcal{V}^{\text{as}}(\mathbb{R}^{2d}_{\not\sim}) := \{v : \mathbb{R}^{2d}_{\not\sim} \to \mathbb{R} | v(x,y) = -v(y,x)\}$, and $\eta : \mathbb{R}^{2d}_{\not\sim} \to \mathbb{R}$ measurable. Furthermore, take a reference measure $\lambda \in \mathcal{M}^+(\mathbb{R}^{2d})$ such that $\rho \otimes \mu, \mu \otimes \rho \ll \lambda$. Then, the admissible flux $F^\Phi[\mu,\eta;\rho,w] \in \mathcal{M}(\mathbb{R}^{2d}_{\not\sim})$ at (ρ,w) is defined by

$$\mathrm{d}F^\Phi[\mu,\eta;\rho,w] = \Phi\left(\frac{\mathrm{d}(\rho\otimes\mu)}{\mathrm{d}\lambda}, \frac{\mathrm{d}(\mu\otimes\rho)}{\mathrm{d}\lambda}; w\right)\eta\,\mathrm{d}\lambda.$$

Remark 2.2.1. Note that, in view of the one-homogeneity of Φ, the definition of admissible flux is independent of the choice of λ as long as the absolute continuity assumption is satisfied. For instance, one could choose $\lambda = |\rho| \otimes \mu + \mu \otimes |\rho|$. Differently from [EPSS21, EPS23], in this chapter η evolves in time and, consequently, so does the set of edges G. For this reason, we include the weight function in the flux, which is now a measure on $\mathbb{R}^{2d}_{\not\sim}$ and not on G. We leave to a future investigation possible nonlinear relations between the flux and the weight function η.

As mentioned in the introduction, the evolution of η is determined by a function $H : [0,T] \times \mathbb{R}^{2d}_{\diagup} \times C_b(\mathbb{R}^{2d}_{\diagup}) \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \to \mathbb{R}$ given by

$$H(t,x,y,\eta_t,\rho_t) := \omega_t[\rho_t](x,y) - \eta_t(x,y),$$

where $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ satisfies properties we postpone to Section 2.3. We focus on the system

$$\begin{aligned} \partial_t \rho_t &= -\overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]], \\ \partial_t \eta_t &= \omega_t[\rho_t] - \eta_t, \end{aligned} \tag{Co-NCL}$$

for given initial data $\rho^0 \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$ and $\eta^0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. A solution to (Co-NCL) is intended as follows, complementing Definition 2.2.2 with the coupled equation for η.

Definition 2.2.6 (Weak solution to the initial value problem (Co-NCL))**.** Given an admissible flux interpolation Φ, a velocity field $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$, and function $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$, a pair $(\rho, \eta) : [0,T] \to \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ is a solution to the initial value problem (Co-NCL) if, for any $\varphi \in C_0(\mathbb{R}^d)$,

(i) $\rho \in AC([0,T], \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$, $\eta \in AC([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$;

(ii) the maps $t \mapsto \langle \varphi, \overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]] \rangle$ and $t \mapsto \omega_t[\rho_t] - \eta_t$ belong to $L^1([0,T])$;

(iii) for a.e. $t \in [0,T]$, every $(x,y) \in \mathbb{R}^{2d}_{\diagup}$, for any $\varphi \in C_0(\mathbb{R}^d)$, the following conditions hold

$$\int_{\mathbb{R}^d} \varphi \mathrm{d}\rho_t = \int_{\mathbb{R}^d} \varphi \mathrm{d}\rho_0 + \frac{1}{2} \int_0^t \iint_{\mathbb{R}^{2d}_{\diagup}} \overline{\nabla} \varphi \mathrm{d}F^{\Phi}[\mu, \eta_s, \rho_s; V_s[\rho_s]] \mathrm{d}s \tag{2.2.1}$$

$$\eta_t(x,y) = \eta_0(x,y) + \int_0^t (\omega_s[\rho_s](x,y) - \eta_s(x,y)) \, \mathrm{d}s. \tag{2.2.2}$$

In the following, for any curve $\gamma \in C([0,T], S)$, for some normed space $(S, \|\cdot\|_S)$, we will write

$$\|\gamma\|_{\infty,S} := \sup_{t \in [0,T]} \|\gamma_t\|_S \ .$$

We denote by $\|\cdot\|_\infty$ the usual sup-norm when this does not create confusion.

2.2.1 A priori properties

Our strategy to prove well-posedness for (Co-NCL) relies on the application of the Banach fixed-point theorem. In the next proposition we collect properties of solutions important for the definition of the solution map, under assumptions milder than those needed later on the velocity fields.

Proposition 2.2.1. Let Φ be an admissible flux interpolation, $\rho_0 \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)$, $\eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. Assume $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ and $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$ satisfy, for some $C_V > 0$ and $C_\omega > 0$,

$$\int_0^T \sup_{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d \setminus \{x\}} |V_s[\rho_s](x,y)| \, \mathrm{d}\mu(y)\mathrm{d}s \leqslant C_V, \tag{2.2.3a}$$

$$\int_0^T \sup_{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{(x,y) \in \mathbb{R}^{2d}_{\diagup}} |\omega_s[\rho_s](x,y)| \mathrm{d}s \leqslant C_\omega. \tag{2.2.3b}$$

Suppose the map $(x,y) \in \mathbb{R}^{2d}_{\diagup} \mapsto \omega_t[\sigma](x,y)$ is continuous, for any $t \in [0,T]$ and $\sigma \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$. For a pair $(\rho, \eta) : [0,T] \to \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ satisfying (iii) in Definition 2.2.6, the following properties hold:

1. for any $\varphi \in C_0(\mathbb{R}^d)$ the maps

$$t \mapsto \iint_{\mathbb{R}^{2d}_{\diagup}} \overline{\nabla}\varphi \mathrm{d}F^\Phi[\mu, \eta_t, \rho_t; V_t[\rho_t]] \in L^1([0,T]),$$

$$t \mapsto \omega_t[\rho_t](x,y) - \eta_t(x,y) \in L^1([0,T]),$$

   and $\rho \in L^\infty([0,T]; \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$, $\eta \in L^\infty([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$ (flux integrability and time boundedness);

2. $\rho_t[\mathbb{R}^d] = \rho_0[\mathbb{R}^d]$ for all $t \in [0,T]$ (mass preservation);

3. $\rho \in AC([0,T], \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d))$ and $\eta \in AC([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$ (absolute continuity);

4. if $\operatorname{supp}\rho_0 \subseteq \operatorname{supp}\mu$, then $\operatorname{supp}\rho_t \subseteq \operatorname{supp}\mu$ for a.e. $t \in [0,T]$ (support inclusion).

Proof. Time boundedness and flux integrability - We start proving these properties hold for η. From (2.2.2) we have

$$|\eta_t(x,y)| \leqslant |\eta_0(x,y)| + \int_0^t (|\omega_s[\rho_s](x,y)| + |\eta_s(x,y)|)\mathrm{d}s,$$

whence

$$\|\eta_t\|_\infty \leqslant \|\eta_0\|_\infty + \int_0^t \sup_{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{(x,y) \in \mathbb{R}^{2d}_{\diagup}} |\omega_s[\rho_s](x,y)| \mathrm{d}s + \int_0^t \|\eta_s\|_\infty \mathrm{d}s.$$

Then, an application of Grönwall's inequality yields

$$\|\eta_t\|_\infty \leqslant \left(\|\eta_0\|_\infty + \int_0^T \sup_{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{(x,y) \in \mathbb{R}^{2d}_{\diagup}} |\omega_s[\rho_s](x,y)| \mathrm{d}s \right) e^T. \tag{2.2.4}$$

By taking the supremum for $t \in [0,T]$, using assumption (2.2.3b) on ω, yields the result for η, meaning $\eta \in L^\infty([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$. Next, we prove an analogous result for ρ. By definition of admissible flux and nonlocal divergence, for any $\varphi \in C_0(\mathbb{R}^d)$, we have

$$\begin{aligned}\langle \varphi, \overline{\nabla} \cdot F^\Phi\left[\mu, \eta_t; \rho_t, V_t[\rho_t]\right]\rangle &= -\frac{1}{2}\iint\limits_{\mathbb{R}^{2d}_{\diagup}} \overline{\nabla}\varphi(x,y) \mathrm{d}F^\Phi\left[\mu, \eta_t; \rho_t, V_t[\rho_t]\right](x,y) \\ &= -\frac{1}{2}\iint\limits_{\mathbb{R}^{2d}_{\diagup}} \overline{\nabla}\varphi(x,y)\Phi\left(\frac{\mathrm{d}\left(\rho_t \otimes \mu\right)}{\mathrm{d}\lambda}, \frac{\mathrm{d}\left(\mu \otimes \rho_t\right)}{\mathrm{d}\lambda}; V_t[\rho_t]\right)\eta_t(x,y)\mathrm{d}\lambda(x,y).\end{aligned}$$

Using (i), (ii) in Definition 2.2.3, [Rud87, Theorem 6.13], the boundedness of η and the antisymmetry of the velocity field, we have, for any $\varphi \in C_0(\mathbb{R}^d)$ and $t \in [0,T]$:

$$\begin{aligned}&\int_0^t \left|\langle \varphi, \overline{\nabla} \cdot F^\Phi\left[\mu, \eta_s; \rho_s, V_s[\rho_s]\right]\rangle\right| \mathrm{d}s \\ &\leqslant \frac{L_\Phi \|\eta\|_\infty}{2}\int_0^t \iint\limits_{\mathbb{R}^{2d}_{\diagup}} |\overline{\nabla}\varphi|\left|V_s[\rho_s]\right|\left(\mathrm{d}\left|\rho_s\right| \otimes \mu + \mathrm{d}\mu \otimes \left|\rho_s\right|\right)\mathrm{d}s \\ &\leqslant L_\Phi \|\eta\|_\infty \|\varphi\|_\infty \int_0^t \iint\limits_{\mathbb{R}^{2d}_{\diagup}} \left|V_s[\rho_s](x,y)\right| \mathrm{d}\mu(y)\mathrm{d}\left|\rho_s\right|(x)\mathrm{d}s \\ &\leqslant L_\Phi \|\eta\|_\infty \|\varphi\|_\infty \int_0^t \overline{v}_s \left|\rho_s\right|[\mathbb{R}^d]\mathrm{d}s, \end{aligned} \tag{2.2.5}$$

where we set $\overline{v}_s := \sup_{\rho \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d \setminus \{x\}} \left|V_s\left[\rho_s\right](x,y)\right| \mathrm{d}\mu(y)$. From (2.2.1) in Definition 2.2.6, (2.2.5), and the definition of the total variation norm, we infer

$$|\rho_t|[\mathbb{R}^d] \leqslant |\rho_0|[\mathbb{R}^d] + L_\Phi \|\eta\|_\infty \int_0^t \overline{v}_s |\rho_s|[\mathbb{R}^d]\mathrm{d}s.$$

An application of Grönwall's inequality and assumption (2.2.3a) provide

$$|\rho_t|[\mathbb{R}^d] \leqslant |\rho_0|[\mathbb{R}^d] e^{L_\Phi \|\eta\|_\infty C_V} < \infty. \tag{2.2.6}$$

Hence, we can conclude the time integrability of the flux and that $\rho \in L^\infty([0,T], \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d))$. *Mass preservation* - Let $\varphi^R(x) := e^{-|x|^2/R^2}$. Note that $\varphi^R \in C_0^\infty(\mathbb{R}^d)$ and $\varphi^R(x) \to 1$ pointwise as $R \to \infty$. Using Definition 2.2.6 (iii), we infer

$$\begin{aligned}&\left|\int_{\mathbb{R}^d} \varphi^R(x)\mathrm{d}\rho_t(x) - \int_{\mathbb{R}^d} \varphi^R(x)\mathrm{d}\rho_0(x)\right| \\ &\quad\leqslant \frac{1}{2}\int_0^t \left|\iint\limits_{\mathbb{R}^{2d}_{\diagup}} (\varphi^R(y) - \varphi^R(x))\eta_s \Phi\left(\frac{\mathrm{d}\left(\rho_t \otimes \mu\right)}{\mathrm{d}\lambda}, \frac{\mathrm{d}\left(\mu \otimes \rho_t\right)}{\mathrm{d}\lambda}; V_t[\rho_t]\right)\mathrm{d}\lambda(x,y)\right|\mathrm{d}s\end{aligned}$$

$$\leqslant L_\Phi \|\eta\|_\infty \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} |\varphi^R(y) - \varphi^R(x)| \, |V_s[\rho_s]| \mathrm{d}|\rho_s|(x) \mathrm{d}\mu(y) \mathrm{d}s \tag{2.2.7}$$

$$\leqslant L_\Phi 2\|\eta\|_\infty \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} |V_s[\rho_s]| \mathrm{d}|\rho_s|(x) \mathrm{d}\mu(y) \mathrm{d}s \ . \tag{2.2.8}$$

Note that the map $(t,x,y) \mapsto V_t[\rho_t](x,y)$ is integrable since

$$\int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} |V_s[\rho_s]| \mathrm{d}|\rho_s|(x) \mathrm{d}\mu(y) \mathrm{d}s \leqslant \int_0^t \overline{v}_s |\rho_s|[\mathbb{R}^d] \mathrm{d}s$$

$$\leqslant C_V |\rho_0|[\mathbb{R}^d] e^{L_\Phi \|\eta\|_\infty C_V} \ .$$

where we used the bounds (2.2.3a) and (2.2.6). Hence, by means of the Dominated Convergence Theorem, as we let $R \to \infty$ in (2.2.7), we obtain

$$\left| \int_{\mathbb{R}^d} \varphi^R(x) \mathrm{d}\rho_t(x) - \int_{\mathbb{R}^d} \varphi^R(x) \mathrm{d}\rho_0(x) \right| \longrightarrow 0 \text{ as } R \to \infty$$

Thus, noting that $\rho_t[\mathbb{R}^d] = \lim_{R\to\infty} \int_{\mathbb{R}^d} \varphi^R(x) \mathrm{d}\rho_t(x)$, which follows again from the Dominated Convergence theorem, as well as $\rho_0[\mathbb{R}^d] = \lim_{R\to\infty} \int_{\mathbb{R}^d} \varphi^R(x) \mathrm{d}\rho_0(x)$, we conclude the proof.

Absolute continuity - This is a direct consequence of the flux integrability proven by exploiting (2.2.1) and (2.2.2).

Support inclusion for ρ - Let $A = \mathbb{R}^d \backslash \operatorname{supp} \mu$ and $(\rho_t)_{t\in[0,T]}$ be a solution to (2.2.1). As we want to evaluate $|\rho_t|(A)$, we consider test functions $\varphi \in C_c(A)$, due to [AFP00, Proposition 1.47]. Using (2.2.1), we estimate:

$$\begin{aligned}
\left| \int_A \varphi(x) \mathrm{d}\rho_t(x) \right| &\leqslant \left| \int_A \varphi(x) \mathrm{d}\rho_0(x) \right| \\
&\quad + \frac{L_\Phi}{2} \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta} \cap \operatorname{supp}\mu \times A} |\varphi(y)| \, |\eta_s| \, |V_s[\rho_s](x,y)| \mathrm{d}\mu(x) \mathrm{d}|\rho_s|(y) \\
&\quad + \frac{L_\Phi}{2} \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta} \cap A \times \operatorname{supp}\mu} |\varphi(x)| \, |\eta_s| \, |V_s[\rho_s](x,y)| \mathrm{d}|\rho_s|(x) \mathrm{d}\mu(y) \\
&\leqslant \left| \int_A \varphi(x) \mathrm{d}\rho_0(x) \right| + L_\Phi \|\eta\|_\infty \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta} \cap A \times \operatorname{supp}\mu} |\varphi(x)| \, |V_s[\rho_s](x,y)| \mathrm{d}|\rho_s|(x) \mathrm{d}\mu(y) \\
&\leqslant \left| \int_A \varphi(x) \mathrm{d}\rho_0(x) \right| + L_\Phi \|\eta\|_\infty \int_0^t \int_A |\varphi(x)| \mathrm{d}|\rho_s|(x) \overline{v}_s \mathrm{d}s \ .
\end{aligned}$$

Taking the supremum over all $\varphi \in C_c(A)$ such that $\|\varphi\|_\infty \leqslant 1$ we infer, by [AFP00, Proposition

1.47] and Grönwall's inequality, that

$$|\rho_t|[A] \leqslant e^{C_V L_\Phi \|\eta\|_\infty} |\rho_0|[A] = 0$$

by definition of A. □

2.3 Well-posedness of the co-evolving non-local continuity equation

In this section we show that the system (Co-NCL) is well-posed by means of a Banach fixed-point argument. From now on we fix $\rho_0 \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)$, $\eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})$, and assume Φ is an admissible flux interpolation according to Definition 2.2.3. Let us denote

$$\mathcal{AC}_T := AC([0,T]; \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)) \times AC([0,T]; C_b(\mathbb{R}^{2d}_{\diagup})),$$

and equip this space with the metric defined by

$$d_\infty((\rho^1,\eta^1),(\rho^2,\eta^2)) := \|\rho^1 - \rho^2\|_{\infty,\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} + \|\eta^1 - \eta^2\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})},$$

where

$$\|\rho^1 - \rho^2\|_{\infty,\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} = \sup_{t\in[0,T]} \|\rho_t^1 - \rho_t^2\|_{\mathrm{TV}},$$
$$\|\eta^1 - \eta^2\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} = \sup_{t\in[0,T]} \|\eta_t^1 - \eta_t^2\|_\infty.$$

We assume the velocity field $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$ to fulfil, for a constant $C_V > 0$, the uniform compressibility assumption

$$\sup_{t\in[0,T]} \sup_{\rho\in\mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_t[\rho](x,y)| \,\mathrm{d}\mu(y) \leqslant C_V. \tag{2.3.1}$$

Assumption (2.3.1) is an L^∞ bound for the nonlocal divergence needed later in our argument to obtain a contraction. We note this is a stronger requirement with respect to time, than (2.2.3a), which was used to obtain the time-continuity of solutions in Proposition 2.2.1.

Remark 2.3.1. We refer to (2.3.1) as the *uniform compressibility* assumption in relation to the terminology used for the continuity equation $\partial_t \rho_t + \nabla \cdot (b_t \rho_t) = 0$, for a vector field $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$. In this context an L^∞-bound is required on $\nabla \cdot b$, which is replaced in our framework by (2.3.1). We refer the reader to [EPS23, Remark 3.6 and Section 4] for further details.

We fix $\omega : [0,T] \times \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\not\Delta} \to \mathbb{R}$ satisfying the following conditions:

($\boldsymbol{\omega}1$) the map $(x,y) \in \mathbb{R}^{2d}_{\not\Delta} \mapsto \omega_\cdot[\cdot](x,y)$ is continuous and $(t \mapsto \omega_t[\cdot](\cdot,\cdot)) \in L^1([0,T])$;

($\boldsymbol{\omega}2$) there exists a constant $L_\omega \geqslant 0$ such that for any $\rho, \sigma \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$ it holds

$$\sup_{t\in[0,T]} \sup_{(x,y)\in\mathbb{R}^{2d}_{\not\Delta}} |\omega_t[\sigma](x,y) - \omega_t[\rho](x,y)| \leqslant L_\omega \|\sigma - \rho\|_{\mathrm{TV}};$$

($\boldsymbol{\omega}3$) ω is bounded, that is, there exists a constant $C_\omega > 0$ such that

$$\sup_{t\in[0,T]} \sup_{\rho\in\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}_{\not\Delta}} \big|\omega_t[\rho](x,y)\big| \leqslant C_\omega.$$

We define the solution maps $\mathcal{S} := (\mathcal{S}_V, \mathcal{S}_\omega) : \mathcal{AC}_T \to \mathcal{AC}_T$ by

$$\mathcal{S}_V(\rho,\eta)(t) := \int_{\mathbb{R}^d} \varphi \mathrm{d}\rho_0 + \frac{1}{2}\int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} \overline{\nabla}\varphi \mathrm{d}F^\Phi[\mu, \eta_s, \rho_s; V_s[\rho_s]]\mathrm{d}s, \tag{2.3.2a}$$

$$\mathcal{S}_\omega(\rho,\eta)(t)(x,y) := \eta^0(x,y) + \int_0^t \omega_s[\rho_s](x,y) - \eta_s(x,y)\mathrm{d}s. \tag{2.3.2b}$$

for any $\varphi \in C_0(\mathbb{R}^d)$, $t \in [0,T]$ and $(x,y) \in \mathbb{R}^{2d}_{\not\Delta}$. We observe that the map $\mathcal{S} : \mathcal{AC}_T \to \mathcal{AC}_T$ is well-defined due to our assumptions on the functions V, (2.3.1), and ω, ($\boldsymbol{\omega}1$) and ($\boldsymbol{\omega}3$), taking into account Proposition 2.2.1.

Remark 2.3.2. In Section 5, we shall compare different solutions $(\rho,\eta), (\widetilde{\rho},\widetilde{\eta})$ of (Co-NCL), with potentially different base measures $\mu, \widetilde{\mu}$. Note that we can take $\lambda := |\rho|\otimes\mu + \mu\otimes|\rho| + |\widetilde{\rho}|\otimes\widetilde{\mu} + \widetilde{\mu}\otimes|\widetilde{\rho}|$ in Definition 2.2.5 in order to be able to compare fluxes associated with (ρ,η) and $(\widetilde{\rho},\widetilde{\eta})$.

First, we show the following contraction with respect to d_∞, crucial for the application of the Banach fixed-point theorem.

Lemma 2.3.1. Let $V : [0,T] \times \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\not\Delta} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\not\Delta})$ satisfy assumption (2.3.1) and assume there is a constant $L_V \geqslant 0$ such that, for all $t \in [0,T]$ and all $\rho, \sigma \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$,

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_t[\rho](x,y) - V_t[\sigma](x,y)| \,\mathrm{d}\mu(y) \leqslant L_V \|\rho - \sigma\|_{\mathrm{TV}} \,. \tag{2.3.3}$$

Assume ω satisfies assumptions ($\boldsymbol{\omega}1$)—($\boldsymbol{\omega}3$) and that the solution map, $\mathcal{S}$, is defined as in (2.3.2). Then, for any $(\rho,\eta), (\widetilde{\rho},\widetilde{\eta}) \in \mathcal{AC}_T$, the following contraction estimate holds

$$d_\infty\big(\mathcal{S}(\rho,\eta), \mathcal{S}(\widetilde{\rho},\widetilde{\eta})\big) \leqslant \kappa(T) T d_\infty((\rho,\eta),(\widetilde{\rho},\widetilde{\eta})),$$

where $\kappa(T) := \kappa(M, L_\Phi, L_V, C_\omega, \|\eta_0\|_\infty, T) \geqslant 0$ and is monotonically increasing. In particular, there exists a $T^* > 0$ solving $\kappa(T^*)T^* = 1$ such that we have a unique solution (ρ, η) to (Co-NCL) on $[0, T]$, for $T < T^*$, and $(\rho(0), \eta(0)) = (\rho_0, \eta_0) \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$.

Proof. Let (ρ, η) and $(\widetilde{\rho}, \widetilde{\eta})$ belong to $\mathcal{AC}_T$. Upon using our assumptions on the flux Φ in Definition 2.2.5 we obtain

$$
\begin{aligned}
|S_V(\rho,\eta)(t) - S_V(\widetilde{\rho},\widetilde{\eta})(t)| &\leqslant \frac{1}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} \Big| (\varphi(y)-\varphi(x))\Big(\eta_s \Phi\left(\frac{\mathrm{d}\rho_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\rho_s}{\mathrm{d}\lambda}; V_s[\rho_s]\right) \\
&\qquad - \widetilde{\eta}_s \Phi\left(\frac{\mathrm{d}\widetilde{\rho}_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\widetilde{\rho}_s}{\mathrm{d}\lambda}; V_s[\widetilde{\rho}_s]\right)\Big)\Big| \mathrm{d}\lambda \mathrm{d}s \\
&\leqslant \frac{1}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} \Big| (\varphi(y)-\varphi(x))\eta_s\Big(\Phi\left(\frac{\mathrm{d}\rho_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\rho_s}{\mathrm{d}\lambda}; V_s[\rho_s]\right) \\
&\qquad - \Phi\left(\frac{\mathrm{d}\widetilde{\rho}_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\widetilde{\rho}_s}{\mathrm{d}\lambda}; V_s[\rho_s]\right)\Big)\Big| \mathrm{d}\lambda \mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} \Big| (\varphi(y)-\varphi(x))\,(\eta_s - \widetilde{\eta}_s) \\
&\qquad \times \Phi\left(\frac{\mathrm{d}\widetilde{\rho}_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\widetilde{\rho}_s}{\mathrm{d}\lambda}; V_s[\rho_s]\right)\Big| \mathrm{d}\lambda \mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} \Big| (\varphi(y)-\varphi(x))\widetilde{\eta}_s\Big(\Phi\left(\frac{\mathrm{d}\widetilde{\rho}_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\widetilde{\rho}_s}{\mathrm{d}\lambda}; V_s[\rho_s]\right) \\
&\qquad - \Phi\left(\frac{\mathrm{d}\widetilde{\rho}_s \otimes \mathrm{d}\mu}{\mathrm{d}\lambda}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\widetilde{\rho}_s}{\mathrm{d}\lambda}; V_s[\widetilde{\rho}_s]\right)\Big)\Big| \mathrm{d}\lambda ds \\
&\leqslant \frac{L_\Phi}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} |\varphi(y)-\varphi(x)|\,|V_s[\rho_s]]\,|\eta_s(x,y)|(\mathrm{d}\,(|\rho_s - \widetilde{\rho}_s| \otimes \mu) \\
&\qquad + \mathrm{d}\,(\mu \otimes |\rho_s - \widetilde{\rho}_s|))\mathrm{d}s \\
&\quad + \frac{L_\Phi}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} |\varphi(y)-\varphi(x)||V_s[\rho_s]| \\
&\qquad \times |\eta_s - \widetilde{\eta}_s|(\mathrm{d}\,(\mu \otimes |\widetilde{\rho}_s|) + \mathrm{d}\,(|\widetilde{\rho}_s| \otimes \mu))\mathrm{d}s \\
&\quad + \frac{L_\Phi}{2}\int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} |\varphi(y)-\varphi(x)||\widetilde{\eta}_s|\,|V_s[\rho_s] - V_s[\widetilde{\rho}_s]| \\
&\qquad \times (\mathrm{d}\,(|\widetilde{\rho}_s| \otimes \mu) + \mathrm{d}\,(\mu \otimes |\widetilde{\rho}_s|))\,\mathrm{d}s \\
&=: I + II + III\ .
\end{aligned}
$$

We then have the following estimates for the terms I, II, III. Starting with I, using antisymmetry of V and our assumption (2.3.1) we have

$$
I \leqslant L_\Phi \|\varphi\|_\infty \int_0^t \int_{\mathbb{R}^{2d}_{\diagup}} |V_s[\rho_s]]\,|\eta_s|\,(\mathrm{d}\,(|\rho_s - \widetilde{\rho}_s| \otimes \mu) + \mathrm{d}\,(\mu \otimes |\rho_s - \widetilde{\rho}_s|))\,\mathrm{d}s
$$

$$
\begin{aligned}
&\leqslant 2L_\Phi\|\varphi\|_\infty\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\int_0^t\int_{\mathbb{R}^{2d}_{\not\wedge}}|V_s\left[\rho_s\right]|\left(\mathrm{d}\left(\mu\otimes|\rho_s-\widetilde{\rho}_s|\right)(y,x)\right)\mathrm{d}s\\
&\leqslant 2L_\Phi\|\varphi\|_\infty\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\int_0^t\int_{\mathbb{R}^d}\mathrm{d}|\rho_s-\widetilde{\rho}_s|(x)\\
&\qquad\times\left(\sup_{\rho\in\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d\setminus\{x\}}|V_s[\rho_s](x,y)|\mathrm{d}\mu(y)\right)\mathrm{d}s\\
&\leqslant 2L_\Phi\|\varphi\|_\infty C_V\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)}T
\end{aligned}
$$

For the term II, using antisymmetry of V and (2.3.1) we get

$$
\begin{aligned}
II&\leqslant L_\Phi\|\varphi\|_\infty\int_0^t\int_{\mathbb{R}^{2d}_{\not\wedge}}|V_s[\rho_s]||\eta_s-\widetilde{\eta}_s|(\mathrm{d}\left(\mu\otimes|\widetilde{\rho}_s|\right)+\mathrm{d}\left(|\widetilde{\rho}_s|\otimes\mu\right))\mathrm{d}s\\
&\leqslant 2L_\Phi\|\varphi\|_\infty\|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\int_0^t\int_{\mathbb{R}^{2d}_{\not\wedge}}|V_s[\rho_s](x,y)||\mathrm{d}\mu(y)\mathrm{d}|\widetilde{\rho}_s|(x)\mathrm{d}s\\
&\leqslant 2L_\Phi\|\varphi\|_\infty C_V M\|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}T
\end{aligned}
$$

Finally, we have the following bound for term III using our assumption (2.3.3) on V:

$$
\begin{aligned}
III&\leqslant L_\Phi\|\varphi\|_\infty\int_0^t\int_{\mathbb{R}^{2d}_{\not\wedge}}|\widetilde{\eta}_s|\left|V_s\left[\rho_s\right]-V_s\left[\widetilde{\rho}_s\right]\right|\left(\mathrm{d}\left(|\widetilde{\rho}_s|\otimes\mu\right)+\mathrm{d}\left(\mu\otimes|\widetilde{\rho}_s|\right)\right)\mathrm{d}s\\
&\leqslant 2L_\Phi\|\varphi\|_\infty\|\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\int_0^t\int_{\mathbb{R}^{2d}_{\not\wedge}}\left|V_s\left[\rho_s\right](x,y)-V_s\left[\widetilde{\rho}_s\right](x,y)\right|\mathrm{d}\mu(y)\mathrm{d}|\widetilde{\rho}_s|(x)\mathrm{d}s\\
&\leqslant 2L_\Phi\|\varphi\|_\infty L_V M\|\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)}T.
\end{aligned}
$$

Thus, taking the supremum over all $\varphi\in C_0(\mathbb{R}^d)$ with $\|\varphi\|_\infty\leqslant 1$ and recalling the definition of the TV-norm we have the following estimate

$$
\begin{aligned}
\|\mathcal{S}_V(\eta,\rho)-\mathcal{S}_V(\widetilde{\eta},\widetilde{\rho})\|_{\infty,\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}&\leqslant 2L_\Phi C_V\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}T\\
&\quad+2L_\Phi C_V M\|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}T\\
&\quad+2L_\Phi L_V M\|\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\wedge})}\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}T.
\end{aligned}
\tag{2.3.4}
$$

On the other hand for $(x,y)\in\mathbb{R}^{2d}_{\not\wedge}$ and $t\in[0,T]$ we have

$$
\begin{aligned}
|\mathcal{S}_\omega(\rho,\eta)(x,y)(t)-\mathcal{S}_\omega(\widetilde{\rho},\widetilde{\eta})(x,y)(t)|&\leqslant\int_0^t|\eta_s(x,y)-\widetilde{\eta}_s(x,y)|\mathrm{d}s\\
&\quad+\int_0^t|\omega_s[\rho_s](x,y)-\omega_s[\widetilde{\rho}_s](x,y)|\mathrm{d}s\ .
\end{aligned}
\tag{2.3.5}
$$

Thus, by assumption ($\boldsymbol{\omega}2$),

$$\begin{aligned}\|\mathcal{S}_\omega(\rho,\eta)-\mathcal{S}_\omega(\widetilde{\rho},\widetilde{\eta})\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} &\leqslant \|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}T \\ &\quad+\int_0^T \sup_{(x,y)\in\mathbb{R}^{2d}_{\diagup}} |\omega_t[\rho_t](x,y)-\omega_t[\widetilde{\rho}_t](x,y)|\,\mathrm{d}t \\ &\leqslant \Big(\|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} \\ &\quad+L_\omega\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}\Big)T.\end{aligned}$$

According to Proposition 2.2.1, Eq. (2.2.4), having assumption ($\boldsymbol{\omega}3$) we know

$$\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} \leqslant (\|\eta_0\|_\infty + C_\omega T)\, e^T. \tag{2.3.6}$$

Combining the previous estimates we infer

$$\begin{aligned}d_\infty(\mathcal{S}(\eta,\rho),\mathcal{S}(\widetilde{\eta},\widetilde{\rho})) &\leqslant (2L_\Phi(\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}C_V \\ &\quad+\|\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}L_V M)+L_\omega)\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}T \\ &\quad+(2L_\Phi C_V M+1)\|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}T \\ &\leqslant \underbrace{(2L_\Phi(C_V+L_V M)(\|\eta_0\|_\infty+C_\omega T)e^T+L_\omega)}_{=:\alpha(T)}\|\rho-\widetilde{\rho}\|_{\infty,\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)}T \\ &\quad+\underbrace{(2L_\Phi C_V M+1)}_{=:\beta}\|\eta-\widetilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}T \\ &\leqslant \max\{\alpha(T),\beta\}Td_\infty((\rho,\eta),(\widetilde{\rho},\widetilde{\eta})).\end{aligned} \tag{2.3.7}$$

When $T<\frac{1}{\max\{\alpha(T),\beta\}}$ we have that the solution map is a contraction. We check this is not an issue in case the maximum is $\alpha(T)=L_\omega+2L_\Phi(C_V+L_V M)\|\eta_0\|_\infty e^T+2L_\Phi(C_V+L_V M)C_\omega T e^T$, which we rewrite as $\alpha(T):=\sigma+\gamma e^T+\chi T e^T$, for some positive constants σ,γ,χ. The inequality $T<\frac{1}{\alpha(T)}$ is true for $T<T^*$, where $T^*=T^*(\sigma,\gamma,\chi)$ is fixed and solves $\alpha(T^*)T^*=1$. Denoting by $\kappa:=\max\{\alpha,\beta\}$ we obtain existence and uniqueness when $T<\frac{1}{\kappa}$ as a direct consequence of the Banach fixed-point theorem. □

Remark 2.3.3. For the sake of completeness we observe that Banach fixed-point theorem is applied to $C_T=C([0,T];\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d))\times C([0,T];C_b(\mathbb{R}^{2d}_{\diagup}))$. Absolute continuity in time follows from the fact that the flux and the map $(t\mapsto\omega_t[\rho_t]-\eta_t)$ belong to $L^1([0,T])$, as proven in Proposition 2.2.1.

Remark 2.3.4. Although this is not required for our result to hold, we note that, if in addition to the current assumptions on ω we impose

$$\inf_{(x,y)\in\mathbb{R}^{2d}_{\not\Delta}} \eta_0(x,y) \geqslant \|\omega_-\|_\infty (e^T - 1) \ ,$$

where $\|\omega_-\|_\infty = \sup_{t\in[0,T]} \sup_{\rho\in\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}_{\not\Delta}} \left|(\omega_t)_-[\rho](x,y)\right|$, we obtain non-negativity preservation of η from the explicit solution

$$\eta_t(x,y) = e^{-t}\left(\eta_0(x,y) + \int_0^t e^s \omega_s[\rho_s](x,y)\mathrm{d}s\right) \ . \tag{2.3.8}$$

We point the reader to Lemma A.0.1 in Appendix A for more details. We note that in [EPS23] only non-negative weight functions were considered, while here we allow for weights to become negative which is used in many applications, see for example [SAB19]. Furthermore, having proved existence and uniqueness of a solution of (Co-NCL) over a suitable time interval, we can use (2.3.8) to improve the bound in (2.3.6) to

$$\|\eta_t\|_\infty \leqslant \|\eta_0\|_\infty + C_\omega \ \ , \tag{2.3.9}$$

for $t \in [0, T^*]$. Thus, we obtain $\alpha(T) = \tilde{\alpha} := L_\omega + 2L_\Phi(C_V + L_V M)\|\eta_0\|_\infty + 2L_\Phi(C_V + L_V M)C_\omega$ in (2.3.7).

Now we are ready to prove existence and uniqueness of the initial value problem (Co-NCL), which is the content of the following theorem.

Theorem 2.3.1 (Existence and uniqueness for (Co-NCL))**.** Let $V : [0,T] \times \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\not\Delta} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\not\Delta})$ and $\omega : [0,T]\times\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)\times\mathbb{R}^{2d}_{\not\Delta} \to \mathbb{R}$ satisfy (2.3.1), (2.3.3) and ($\boldsymbol{\omega}$1)—($\boldsymbol{\omega}$3), respectively. Furthermore, let the assumptions of Lemma 2.3.1 hold. Then there exists a unique solution (ρ, η) to (Co-NCL) such that $(\rho_0, \eta_0) = (\rho^0, \eta^0)$.

Proof. Let κ be as in Lemma 2.3.1. The condition $T < \frac{1}{\kappa}$ can be also interpreted as $T < a := \min\left\{\frac{1}{\beta}, T^*(\sigma,\gamma,\chi)\right\}$. Assume $T \geqslant a$, let $\tau := a/2$. Then by Lemma 2.3.1, we know there exists a unique solution in $\mathcal{AC}_T$ to (Co-NCL) on $[0,\tau]$. Denote this solution by (ρ^1, η^1) and note that $(\rho^1, \eta^1) \in \mathcal{AC}_{0,\tau}$, where $\mathcal{AC}_{0,\tau} = AC([0,\tau], \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)) \times AC([0,\tau], C_b(\mathbb{R}^{2d}_{\not\Delta}))$, where by Remark 2.3.4, the estimate (2.3.9) holds. In particular, writing $\alpha^1(t)$ for the constant α in Lemma 2.3.1 corresponding to (ρ^1, η^1) we obtain $\alpha^1(t) = \tilde{\alpha}$ for $t \in [0,\tau]$. Next, let us consider (Co-NCL) with initial condition $(\rho^1_\tau, \eta^1_\tau)$ on the interval $[\tau, \tilde{\tau}]$, where $\tilde{\tau}$ has to be determined. By Lemma 2.3.1

there is a time $\tilde{\tau}$, possibly different from τ, depending on

$$2L_\Phi(C_V + L_V M)(\|\eta_\tau^1\|_\infty + C_\omega(\tilde{\tau}-\tau)e^{\tilde{\tau}-\tau}) + L_\omega \\ \leqslant 2L_\Phi(C_V + L_V M)(\|\eta_0\|_\infty + C_\omega + C_\omega(\tilde{\tau}-\tau)e^{\tilde{\tau}-\tau}) + L_\omega =: \alpha^2(\tilde{\tau}) \ ,$$

where in the second inequality we have used the estimate (2.3.9) to bound $\|\eta^1\|_\infty$, such that there exists a unique solution $(\rho^2, \eta^2) \in \mathcal{AC}_{\tau,\tilde{\tau}}$. Having established existence and uniqueness of the solution over the interval $[0, \tilde{\tau}]$, we can improve the bound on $\|\eta_t^2\|_\infty$ to

$$\|\eta_t^2\|_\infty \leqslant \|\eta_0\|_\infty + C_\omega \tag{2.3.10}$$

for any $t \in [0, \tilde{\tau}]$. Note that following this same procedure we can obtain the existence and uniqueness of (ρ^3, η^3) over $[\tilde{\tau}, 2\tilde{\tau}]$ by noting that due to (2.3.10) we obtain $\alpha^3(\tilde{\tau}) = \alpha^2(\tilde{\tau})$. Denote by $N := \left[\frac{T-\tau}{\tilde{\tau}}\right]$. Iterating this procedure we can construct a sequence of solutions

$$(\rho^i, \eta^i) \in \mathcal{AC}_{(i-2)\tilde{\tau},(i-1)\tilde{\tau}} \ \text{ for all } i \in \{3, \dots, N+2\}, \qquad (\rho^{N+2}, \eta^{N+2}) \in \mathcal{AC}_{N\tilde{\tau},T} \ .$$

We can now define the curve $(\rho, \eta) \in \mathcal{AC}_{0,T} = \mathcal{AC}_T$ by

$$\begin{cases} (\rho_t, \eta_t) = (\rho_t^1, \eta_t^1) & \text{for } t \in [0, \tau) \\ (\rho_t, \eta_t) = (\rho_t^2, \eta_t^2) & \text{for } t \in [\tau, \tilde{\tau}) \\ (\rho_t, \eta_t) = (\rho_t^i, \eta_t^i) & \text{for all } t \in [(i-2)\tilde{\tau}, (i-1)\tilde{\tau}) \text{ and } i \in \{3, \dots, N+1\} \\ (\rho_t, \eta_t) = (\rho_t^{N+2}, \eta_t^{N+2}) & \text{for all } t \in [N\tilde{\tau}, T] \end{cases}$$

which, by construction, is the unique solution to (Co-NCL). □

Remark 2.3.5. In a similar spirit of [EPS23, Section 5], we could extend the result of this section to the L^1_μ setting, i.e. $\rho \ll \mu$, which can be obtained upon assuming $\rho_0 \ll \mu$ and choosing $\lambda = \mu \otimes \mu$. In this scenario, for $\eta_t \geqslant 0$ and symmetric for any $t \in [0, T]$ — obtained by assuming $\eta_0(x, y)$ and $\omega_\cdot[\cdot](x, y)$ to be symmetric and η to satisfy the condition in Remark 2.3.4 — we can prove non-negativity preservation for ρ_t as in [EPS23, Proposition 5.2], i.e., if the initial condition is non-negative then the solution $\rho_t(x) \geqslant 0$ for a.e. $t \in [0, T]$ and μ-a.e. $x \in \mathbb{R}^d$.

2.4 Fast-Slow versions of the problem

In this section we are concerned with different time scales for the evolution of the weights, motivated by [Bur22, Section 2.2]. We emphasize that the notion of solution outlined in

Definition 2.2.6 is the one used throughout. Let $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}^{M}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ and $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}^{M}(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ satisfy (2.3.1) and ($\boldsymbol{\omega}1$)—($\boldsymbol{\omega}3$), respectively. We start with the following version of (Co-NCL):

$$\begin{cases} \partial_t \rho_t = -\overline{\nabla} \cdot F^{\Phi}[\mu, \eta_t; \rho_t, V_t[\rho_t]] \\ \partial_t \eta_t = \varepsilon(\omega_t[\rho_t] - \eta_t) \\ \rho_0 \in \mathcal{M}_{\mathrm{TV}}^{M}(\mathbb{R}^d),\ \eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})\ , \end{cases} \tag{Co$-$NCL$_{\mathrm{S}}$}$$

where $\varepsilon > 0$ is a small parameter. This problem corresponds to the case in which the weights of the graph evolve at a slower time scale than the mass on the vertices. Indeed, if we formally consider that the graph evolves at a time scale $\tau = \varepsilon t$, then expressing the evolution of η for the time scale t yields (Co$-$NCL$_{\mathrm{S}}$).

Theorem 2.3.1 implies this problem is well-posed for $\varepsilon > 0$. Then, it is natural to study the behaviour of the system (Co$-$NCL$_{\mathrm{S}}$) as ε approaches 0^+. The following result shows that, as $\varepsilon \to 0^+$, the solution to (Co$-$NCL$_{\mathrm{S}}$) converges to the solution of a PDE on a static graph, meaning the weight function is given by the initial condition, η_0, which does not change in time.

Theorem 2.4.1. Let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ a vanishing sequence, that is $\varepsilon_n \to 0^+$. Consider the sequence of solutions $\{(\rho^n, \eta^n)\}_{n\in\mathbb{N}}$ to (Co$-$NCL$_{\mathrm{S}}$) with $\rho_0^n = \rho^0 \in \mathcal{M}_{\mathrm{TV}}^{M}(\mathbb{R}^d)$, and $\eta_0^n = \eta^0 \in C_b(\mathbb{R}^{2d}_{\diagup})$ satisfying (2.3.1), (2.3.3) and ($\boldsymbol{\omega}1$)–($\boldsymbol{\omega}3$). It holds that

$$d_\infty((\rho^n, \eta^n), (\overline{\rho}, \eta^0)) \to 0, \text{ as } n \to \infty,$$

where

$$\overline{\rho}_t + \int_0^t \overline{\nabla} \cdot F^{\Phi}[\mu, \eta^0; \overline{\rho}_s, V_s(\overline{\rho}_s)]\mathrm{d}s = \rho^0\ ,$$

understood in duality with $C_0(\mathbb{R}^d)$.

Proof. We begin noticing that Proposition 2.2.1 ensures, for any $n \in \mathbb{N}$, that a solution η^n of (Co$-$NCL$_{\mathrm{S}}$) belongs to $C([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$, satisfying the uniform bound (2.3.6). Thus, let $M_\eta := \sup_{n\in\mathbb{N}} \|\eta^n\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})}$. The argument follows the proof of Lemma 2.3.1.

From (2.3.5) we know

$$\|\eta_t^n - \eta_0\|_\infty \leqslant \varepsilon_n \left(\int_0^t \|\eta_s^n - \eta_0\|_\infty \mathrm{d}s + L_\omega \int_0^t \|\rho_s - \overline{\rho}_s\|_{\mathrm{TV}} \mathrm{d}s \right)\ .$$

Noting that $t \mapsto \int_0^t \|\rho_s - \overline{\rho}_s\|_{\mathrm{TV}} \mathrm{d}s$ is non-decreasing, the corresponding version of Grönwall's inequality implies,

$$\|\eta_t^n - \eta_0\|_\infty \leqslant \varepsilon_n e^{\varepsilon_n T}(2L_\omega M T) \tag{2.4.1}$$

Furthermore, following analogous calculations to the ones in Lemma 2.3.1 used to obtain (2.3.4), we now compare ρ_t^n and $\overline{\rho}_t$ and get

$$\begin{aligned}\|\rho_t^n - \overline{\rho}_t\|_{\mathrm{TV}} \leqslant\ & 2L_\Phi C_V \|\eta^n\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\infty})} \int_0^t \|\rho_s^n - \overline{\rho}_s\|_{\mathrm{TV}} \mathrm{d}s \\ & + 2L_\Phi C_V M \int_0^t \|\eta_s^n - \eta_0\|_\infty \mathrm{d}s \\ & + 2L_\Phi \|\eta_0\|_\infty L_V M \int_0^t \|\rho_s^n - \overline{\rho}_s\|_{\mathrm{TV}} \mathrm{d}s.\end{aligned}$$

Differently from Lemma 2.3.1, we do not further bound the terms under the time integrals. Then, using (2.4.1), the uniform boundedness of η^n and η_0 and an application of the same version of Grönwall's inequality as before yields

$$\|\rho_t^n - \overline{\rho}_t\|_{\mathrm{TV}} \leqslant \varepsilon_n e^{2L_\Phi (M_\eta C_V + L_V M \|\eta_0\|_\infty) T + \varepsilon_n T} (4 L_\Phi C_V L_\omega M^2 T^2)\ . \tag{2.4.2}$$

The result follows from taking the supremum with respect to $t \in [0,T]$ in (2.4.1) and (2.4.2), and letting n go to ∞. □

The second scenario we consider is when the graph evolves faster than the mass on the vertices, i.e.

$$\begin{cases} \partial_t \rho_t = -\overline{\nabla} \cdot F^\Phi[\mu, \eta_t; \rho_t, V_t[\rho_t]] \\ \varepsilon \partial_t \eta_t(x,y) = \omega_t[\rho](x,y) - \eta_t(x,y), \\ \rho_0 \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d),\ \eta_0 \in C_b(\mathbb{R}^{2d}_{\not\infty})\ . \end{cases} \tag{Co$-$NCL$_{\mathrm{F}}$}$$

System (Co$-$NCL$_{\mathrm{F}}$) can be obtained by (formally) considering that the weight function η evolves at a time scale $\tau = \frac{t}{\varepsilon}$ and then re-writing its dynamics at the time scale t of the mass on the vertices. We are interested in studying the limit of this system as $\varepsilon \to 0^+$. As one can formally see from (Co$-$NCL$_{\mathrm{F}}$), this limit would identify the weight function η as a function depending on the mass configuration, nonlocal in case ω is as in Example 2.4.1, providing a whole class of weight functions not considered so far in the literature, to the best of our knowledge.

A classical result for systems of ODEs whose orbits are contained in Euclidean space is given by the Tykhonov Theorem, cf. e.g. [BL14, Chapter 3], providing suitable conditions for the $\varepsilon \to 0^+$ limit to yield the formal limit in (Co$-$NCL$_{\mathrm{F}}$). In our case, however, we do not exploit the previous result as we have an explicit solution for η and can evaluate the limit directly.

In order to use the latter information, we require more regularity of the function ω, in particular, we further assume:

($\boldsymbol{\omega}4$) $t \mapsto \omega_t[\cdot](\cdot,\cdot) \in W^{1,1}([0,T])$;

($\boldsymbol{\omega}5$) $\operatorname{ess\,sup}_{t\in[0,T]} \sup_{\sigma\in\mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}_{\diagup}} \left|\partial_t \omega_t[\sigma](x,y)\right| \leqslant \widetilde{C}_\omega$, for some $\widetilde{C}_\omega \in (0,\infty)$,

where $W^{1,1}$ stands for the Sobolev space. The additional hypotheses above are needed in order to exploit the explicit form of η and to use integration by parts in Sobolev spaces. We note that given our assumptions (2.3.1) and (2.3.3) and ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}5$), (Co$-$NCL$_{\mathrm{F}}$) is well-posed by Lemma 2.3.1 and Theorem 2.3.1 for any $\varepsilon > 0$.

Example 2.4.1. In this case, we can consider the following slightly modified version of the example (2.1.3)

$$\omega_t[\sigma](x,y) = \int_{\mathbb{R}^d} K(t,x,y,z)\mathrm{d}\sigma(z),$$

for $\sigma \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$. We assume $K : [0,T]\times\mathbb{R}^{2d}_{\diagup}\times\mathbb{R}^d \to \mathbb{R}$ to satisfy:

- $t \mapsto K(t,\cdot,\cdot,\cdot) \in C^1([0,T])$.
- $(x,y,z) \mapsto K(\cdot,x,y,z) \in C_0(\mathbb{R}^{2d}_{\diagup}\times\mathbb{R}^d)$.
- $\partial_t K(t,x,y,z) \in C([0,T], C_0(\mathbb{R}^{2d}_{\diagup}\times\mathbb{R}^d))$.

The assumptions ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}3$) still hold, and the derivative is given by

$$\partial_t\omega_t[\sigma](x,y) = \int_{\mathbb{R}^d} \partial_t K(t,x,y,z)\mathrm{d}\sigma(z).$$

where we note that the derivative of the function $t \mapsto \int_{\mathbb{R}^d} K(t,x,y,z)\mathrm{d}\sigma(z)$ follows from an application of the Dominated Convergence Theorem. Hence, we can verify ($\boldsymbol{\omega}4$), ($\boldsymbol{\omega}5$).

Remark 2.4.1. Later on we actually need assumptions ($\boldsymbol{\omega}4$) and ($\boldsymbol{\omega}5$) for solutions of (Co$-$NCL$_{\mathrm{F}}$), that is for time-dependent measures σ. Note that for any $(t,x,y) \in [0,T]\times\mathbb{R}^{2d}_{\diagup}$, we can consider the test function $z \mapsto K(\cdot,\cdot,\cdot,z) \in C_0(\mathbb{R}^d)$. Then, for a pair (ρ,η) solving (Co$-$NCL$_{\mathrm{F}}$) we know from Proposition 2.2.1 that $\rho \in \mathrm{AC}([0,T];\mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d))$ and, for a.e. $t \in [0,T]$, we have

$$\begin{aligned}
\partial_t(\omega_t[\rho_t])(x,y) &= \partial_t \int_{\mathbb{R}^d} K(t,x,y,z)\mathrm{d}\rho_t(z) \\
&= \int_{\mathbb{R}^d} \partial_t K(t,x,y,z)\mathrm{d}\rho_t(z) \\
&\quad + \frac{1}{2}\iint_{\mathbb{R}^{2d}_{\diagup}} (K(t,x,y,v) - K(t,x,y,u))\mathrm{d}F^\Phi[\mu,\eta_t;\rho_t,V_t[\rho_t]](u,v)\ .
\end{aligned}$$

Therefore, assumption ($\boldsymbol{\omega}4$) and ($\boldsymbol{\omega}5$) are satisfied by our assumptions on K, the flux F^Φ, and the fact that $(\rho_t)_{t\in[0,T]}$ is a solution to (Co$-$NCL$_{\mathrm{F}}$).

Theorem 2.4.2. Let $(\varepsilon_n)_{n\in\mathbb{N}} \subset (0,\infty)$ such that $\varepsilon_n \to 0^+$ as $n\to\infty$, and consider a sequence of solutions $\{(\rho^n,\eta^n)\}_{n\in\mathbb{N}}$ to (Co$-$NCL$_{\mathrm{F}}$) with $\eta^n_0 \in C_b(\mathbb{R}^{2d}_{\diagup})$ and $\rho^n_0 \in \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d)$ satisfying

$\|\eta_0^n - \omega_0[\bar{\rho}_0]\|_\infty \to 0$, and $\|\omega_0[\rho_0^n] - \omega_0[\bar{\rho}_0]\|_\infty \to 0$, as $n \to \infty$. Assume V and ω satisfy (2.3.1), (2.3.3), and ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}5$), respectively. It holds

$$d_\infty((\rho^n, \eta^n), (\rho, \omega[\rho])) \to 0 \text{ as } n \to \infty \ ,$$

where, for a.e. $t \in [0,T]$, ρ solves

$$\rho_t + \int_0^t \overline{\nabla} \cdot F^\Phi[\mu, \omega_s[\rho_s]; \rho_s, V_s[\rho_s]]\mathrm{d}s = \rho_0 \ , \tag{2.4.3}$$

to be understood in duality with $C_0(\mathbb{R}^d)$.

Proof. Let us begin by considering the solution to the ODE for the weight function η^n:

$$\eta_t^n(x,y) = e^{-t/\varepsilon_n}\left(\eta_0^n(x,y) + \int_0^t \frac{e^{s/\varepsilon_n}}{\varepsilon_n}\omega_s[\rho_s^n](x,y)\mathrm{d}s\right).$$

By our assumption ($\boldsymbol{\omega}4$), an integration by parts yields

$$\begin{aligned}\eta_t^n(x,y) &= e^{-t/\varepsilon_n}\left(\eta_0^n(x,y) + \omega_t[\rho_t^n](x,y)e^{t/\varepsilon_n} - \omega_0[\rho_0^n](x,y) - \int_0^t e^{s/\varepsilon_n}\partial_s(\omega_s[\rho_s^n])(x,y)\mathrm{d}s\right)\\ &= \eta_0^n(x,y)e^{-t/\varepsilon_n} + \omega_t[\rho_t^n](x,y) - \omega_0[\rho_0^n](x,y)e^{-t/\varepsilon_n} - \int_0^t e^{-(t-s)/\varepsilon_n}\partial_s(\omega_s[\rho_s^n])(x,y)\mathrm{d}s.\end{aligned}$$

This implies

$$\begin{aligned}\|\eta_t^n - \omega_t[\rho_t]\|_\infty &\leqslant \|\eta_0^n - \omega_0[\rho_0^n]\|_\infty e^{-t/\varepsilon_n} + \|\omega_t[\rho_t^n] - \omega_t[\rho_t]\|_\infty + \int_0^t e^{-(t-s)/\varepsilon_n}\|\partial_s(\omega_s[\rho_s^n])\|_\infty\mathrm{d}s\\ &\leqslant L_\omega\|\rho_t - \rho_t^n\|_{\mathrm{TV}} + \|\eta_0^n - \omega[\rho_0^n]\|_\infty + \widetilde{C}_\omega\varepsilon_n(1 - e^{-t/\varepsilon_n})\end{aligned}$$

where in the last line we have used assumption ($\boldsymbol{\omega}5$). On the other hand, similar to the proof of Theorem 2.4.1, we can estimate

$$\begin{aligned}\|\rho_t^n - \rho_t\|_{\mathrm{TV}} \leqslant\ & 2L_\Phi C_V M_\eta \int_0^t \|\rho_s^n - \rho_s\|_{\mathrm{TV}}\mathrm{d}s\\ &+ 2L_\Phi C_V M \int_0^t \|\eta^n - \omega_s[\rho_s]\|_\infty\mathrm{d}s\\ &+ 2L_\Phi C_\omega L_V M \int_0^t \|\rho_s^n - \rho_s\|_{\mathrm{TV}}\mathrm{d}s \ ,\end{aligned}$$

where we denote by $M_\eta := \sup_{n\in\mathbb{N}}\|\eta^n\|_{\infty, C_b(\mathbb{R}^{2d})}$. Let

$$u^n(t) := \|\rho_t^n - \rho_t\|_{\mathrm{TV}} + \|\eta_t^n - \omega_t[\rho_t]\|_\infty,$$

and note that

$$u^n(t) \leqslant \|\eta_0^n - \omega_0[\rho_0^n]\|_\infty + \widetilde{C}_\omega \varepsilon_n (1 - e^{-t/\varepsilon_n}) + C \int_0^t u^n(s) \mathrm{d}s$$

for a suitable constant C independent of n including all the others. By using Grönwall's inequality

$$\begin{aligned} u^n(t) &\leqslant \|\eta_0^n - \omega_0[\rho_0^n]\|_\infty + \widetilde{C}_\omega \varepsilon_n (1 - e^{t/\varepsilon_n}) \\ &\quad + C \int_0^t (\|\eta_0^n - \omega_0[\rho_0^n]\|_\infty + \widetilde{C}_\omega \varepsilon_n (1 - e^{s/\varepsilon_n})) e^{C(t-s)} \mathrm{d}s \\ &\leqslant \left(\|\eta_0^n - \omega_0[\rho_0^n]\|_\infty + \widetilde{C}_\omega \varepsilon_n \right) \left(1 + CTe^{CT}\right). \end{aligned} \tag{2.4.4}$$

By taking the supremum over $t \in [0,T]$ and letting $n \to \infty$ yields the result. □

Remark 2.4.2. As a byproduct of Theorem 2.4.2 we obtain another existence result for solutions of (2.4.3), i.e., equations of type (2.1.1) with a weight function depending on the mass configuration. This is, indeed, an extension of [EPS23] to a different class of weight functions.

2.5 Discrete-to-continuum limit

In this section we are interested in considering problem (Co-NCL) as the limit of a sequence of discrete finite graphs with an increasing number of vertices, i.e. a sequence of graphs whose base measure is given by a sequence of atomic measures

$$\mu^n = \sum_{i=1}^n m_i^n \delta_{x_i} \ , \quad x_i \in \mathbb{R}^d, \ m_i^n \in (0,\infty) \text{ for } i = 1, \dots, n \text{ for } n \in \mathbb{N}. \tag{2.5.1}$$

Let us assume the uniform compressibility condition (2.3.1) for μ^n, i.e., the velocity field $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$ satisfies

$$\sup_{t\in[0,T]} \sup_{\rho \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)} \sup_{x \in \mathbb{R}^d} \sum_{\substack{j=1 \\ x_j \neq x}}^n |V_t[\rho](x, x_j)| \leqslant C_V, \tag{2.5.2}$$

for a constant $C_V > 0$, and the Lipschitz assumption (2.3.3), that is there is a constant $L_V \geqslant 0$ such that, for all $t \in [0,T]$ and all $\rho, \sigma \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)$,

$$\sup_{x \in \mathbb{R}^d} \sum_{\substack{j=1 \\ x_j \neq x}}^n |V_t[\rho](x, x_j) - V_t[\sigma](x, x_j)| \leqslant L_V \|\rho - \sigma\|_{\mathrm{TV}} \ . \tag{2.5.3}$$

Additionally, if ($\boldsymbol{\omega}1$)–($\boldsymbol{\omega}3$) are satisfied, the system

$$\begin{cases} \partial_t \rho_t^n = -\overline{\nabla} \cdot F^{\Phi}[\mu^n, \eta_t^n; \rho_t^n, V_t[\rho_t^n]] \\ \partial_t \eta_t^n(x,y) = \omega_t[\rho^n](x,y) - \eta_t^n(x,y), \\ \rho_0 \in \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d),\ \eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup}), \end{cases} \tag{Co$-$NCL$_{\mathrm{n}}$}$$

is well-posed by Theorem 2.3.1 for any $n \in \mathbb{N}$. Under suitable assumptions we will show that (Co-NCL) can be obtained as approximation by (Co$-$NCL$_{\mathrm{n}}$), i.e. that (Co-NCL) is a good mean field approximation for evolution problems on large finite graphs. Throughout the rest of this section we will consider the upwind interpolation for ease of presentation

$$\Phi_{\mathrm{upwind}}\,(a,b;w) = aw_+ - bw_-, \qquad \text{for } (a,b,w) \in \mathbb{R}^3.$$

Other admissible interpolations can be also considered, see Remark 2.5.2. We require stronger regularity of the velocity field $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d) \times \mathbb{R}^{2d} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d})$, namely, we assume the map $\mathbb{R}^{2d} \ni (x,y) \mapsto V[\cdot](\cdot,x,y) \in C_0(\mathbb{R}^{2d})$, in view of the weak-* convergence we use below. Note that now the velocity field takes values in the space of antisymmetric vector fields over all of $\mathbb{R}^{2d}$, which does not allow for singularities on the diagonal $\{x = y\}$. Consequently, we also replace (2.5.2) by the slightly stronger condition

$$\sup_{t\in[0,T]} \sup_{\rho\in\mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}} |V_t[\rho](x,y)| \leqslant C_V. \tag{2.5.4}$$

Remark 2.5.1. Note that our guiding example for the velocity field, namely $V_t[\rho_t](x,y) = -\overline{\nabla}(K * \rho_t)(x,y)$ satisfies these assumptions if the kernel $K \in C_0(\mathbb{R}^{2d})$.

We also require more regularity of the function η to prove the stability result with respect to μ in Theorem 2.5.1 below, i.e., $\eta \in C([0,T], C_0(\mathbb{R}^{2d}))$. Note that it is now well defined on the diagonal and vanishes at infinity. To guarantee this requirement is met, we will need the initial datum $\eta_0 \in C_0(\mathbb{R}^{2d})$ and the function $\omega : [0,T] \times \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d} \to \mathbb{R}$ to be such that the map $\mathbb{R}^{2d} \ni (x,y) \mapsto \omega.[\cdot](x,y) \in C_0(\mathbb{R}^{2d})$, as well as satisfying the rest of the assumptions in ($\boldsymbol{\omega}1$)–($\boldsymbol{\omega}3$).

Below, we present a stability result with respect to the base measure. Note that in case μ^n is a sequence of atomic measures of the form (2.5.1) the next theorem is a discrete-to-continuum limit for (Co-NCL).

Theorem 2.5.1. Fix $\Phi \equiv \Phi_{\mathrm{upwind}}$ and consider a sequence $\{\mu_n\}_{n\in\mathbb{N}} \in \mathcal{M}_{\mathrm{TV}}^+(\mathbb{R}^d)$ such that $\mu^n \xrightarrow{*} \mu \in \mathcal{M}_{\mathrm{TV}}^+(\mathbb{R}^d)$. Let $V : [0,T] \times \mathcal{M}_{\mathrm{TV}}^M(\mathbb{R}^d) \times \mathbb{R}^{2d} \to \mathcal{V}^{as}(\mathbb{R}^{2d})$ satisfy assumptions (2.3.3) and (2.5.4),

uniformly in n, and $\omega : [0,T] \times \mathcal{M}^M_{\mathrm{TV}}(\mathbb{R}^d) \times \mathbb{R}^{2d} \to \mathbb{R}^d$ satisfy ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}3$). Assume $((x,y) \mapsto V[\cdot](\cdot,x,y)) \in C_0(\mathbb{R}^{2d})$ and $((x,y) \mapsto \omega.[\cdot](x,y)) \in C_0(\mathbb{R}^{2d})$. Let us consider a sequence of solutions $\{(\rho^n,\eta^n)\}_{n\in\mathbb{N}}$ to (Co$-$NCL$_\mathrm{n}$) associated to $\{\mu_n\}$ and with $\lambda = \lambda^n = |\rho| \otimes \mu + \mu \otimes |\rho| + |\rho^n| \otimes \mu^n + \mu^n \otimes |\rho^n|$ for $n \in \mathbb{N}$ in definition 2.2.5. Let (ρ,η) be the solution to (Co-NCL) depending on μ. If $\|\rho_0^n - \rho^0\|_{\mathrm{TV}} \to 0$ and $\|\eta_0^n - \eta_0\|_\infty \to 0$ as $n \to \infty$, then

$$\lim_{n\to\infty} d_\infty((\rho^n,\eta^n),(\rho,\eta)) = 0 \ .$$

Proof. We begin with an estimate analogous to the one in the proof of Lemma 2.3.1, with an additional term that accounts for the difference in the base measure between the solutions compared (see Remark 2.3.2, that is, for any $\varphi \in C_0(\mathbb{R}^d)$ we have

$$\begin{aligned}
\left|\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_s(x) - \int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_s^n(x)\right| &\leqslant \left|\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_0(x) - \int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_0^n(x)\right| \\
&\quad + \left|\int_0^t \int_{\mathbb{R}^{2d}} \frac{\overline{\nabla}\varphi(x,y)}{2}\left(\Phi\left(\frac{\mathrm{d}\rho_s\otimes\mathrm{d}\mu}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu\otimes\mathrm{d}\rho_s}{\mathrm{d}\lambda^n},V_s[\rho_s]\right)\eta_s\right.\right. \\
&\quad \left.\left. - \Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu^n}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu^n\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n},V_s[\rho_s^n]\right)\eta_s^n\right)\mathrm{d}\lambda^n(x,y)\mathrm{d}s\right| . \\
&\leqslant \left|\int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_0(x) - \int_{\mathbb{R}^d}\varphi(x)\mathrm{d}\rho_0^n(x)\right| \\
&\quad + \left|\int_0^t \int_{\mathbb{R}^{2d}} \frac{\overline{\nabla}\varphi(x,y)}{2}\eta_s\left(\Phi\left(\frac{\mathrm{d}\rho_s\otimes\mathrm{d}\mu}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu\otimes\mathrm{d}\rho_s}{\mathrm{d}\lambda^n};V_s\left[\rho_s\right]\right)\right.\right. \\
&\quad \left.\left. - \Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n};V_s\left[\rho_s\right]\right)\right)\mathrm{d}\lambda^n\mathrm{d}s\right| \\
&\quad + \left|\int_0^t \int_{\mathbb{R}^{2d}} \frac{\overline{\nabla}\varphi(x,y)}{2}\eta_s\left(\Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n};V_s\left[\rho_s\right]\right)\right.\right. \\
&\quad \left.\left. - \Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu^n}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu^n\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n};V_s\left[\rho_s\right]\right)\right)\mathrm{d}\lambda^n\mathrm{d}s\right| \\
&\quad + \left|\int_0^t \int_{\mathbb{R}^{2d}} \frac{\overline{\nabla}\varphi(x,y)}{2}\left(\eta_s-\eta_s^n\right)\right. \\
&\quad \left. \times \Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu^n}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu^n\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n};V_s\left[\rho_s\right]\right)\mathrm{d}\lambda^n\mathrm{d}s\right| \\
&\quad + \left|\int_0^t \int_{\mathbb{R}^{2d}} \frac{\overline{\nabla}\varphi(x,y)}{2}\eta_s^n\left(\Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu^n}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu^n\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n};V_s\left[\rho_s\right]\right)\right.\right. \\
&\quad \left.\left. - \Phi\left(\frac{\mathrm{d}\rho_s^n\otimes\mathrm{d}\mu^n}{\mathrm{d}\lambda^n},\frac{\mathrm{d}\mu^n\otimes\mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n};V_s\left[\rho_s^n\right]\right)\right)\mathrm{d}\lambda^n ds\right| \\
&=: I_0 + I + II + III + IV \ .
\end{aligned}$$

We can estimate the terms I, III and IV analogously to the proof of Lemma 2.3.1 (see also the proof of Theorem 2.4.1), noting that none of those estimates depends on λ^n, so in the rest of

this proof we focus on obtaining a bound for II.

$$
\begin{aligned}
II = &\left| \int_0^t \int_{\mathbb{R}^{2d}} \frac{\overline{\nabla}\varphi(x,y)}{2} \eta_s \Bigg(\Phi\left(\frac{\mathrm{d}\rho_s^n \otimes \mathrm{d}\mu}{\mathrm{d}\lambda^n}, \frac{\mathrm{d}\mu \otimes \mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n}, V_s[\rho_s] \right) \right. \\
&\left. - \Phi\left(\frac{\mathrm{d}\rho_s^n \otimes \mathrm{d}\mu^n}{\mathrm{d}\lambda^n}, \frac{\mathrm{d}\mu^n \otimes \mathrm{d}\rho_s^n}{\mathrm{d}\lambda^n}, V_s[\rho_s] \right) \Bigg) \mathrm{d}\lambda^n(x,y)\mathrm{d}s \right| \\
= &\left| \int_0^t \int_{\mathbb{R}^{2d}} \frac{(\varphi(y)-\varphi(x))}{2} \eta_s(x,y) \Big(V_s\left[\rho_s\right]_+ (x,y) d\rho_s^n(x) d(\mu-\mu^n)(y) \right. \\
&\left. - V_s\left[\rho_s\right]_- (x,y) d(\mu-\mu^n)(x) d\rho_s^n(y) \Big) \mathrm{d}s \right| \\
\leqslant &\left| \int_0^t \int_{\mathbb{R}^{2d}} \frac{(\varphi(y)-\varphi(x))}{2} \eta_s(x,y) V_s\left[\rho_s\right]_+ (x,y) d\rho_s^n(x) d(\mu-\mu^n)(y) \mathrm{d}s \right| \\
&+ \left| \int_0^t \int_{\mathbb{R}^{2d}} \frac{(\varphi(y)-\varphi(x))}{2} \eta_s(x,y) V_s\left[\rho_s\right]_- (x,y) d(\mu-\mu^n)(x) d\rho_s^n(y) \mathrm{d}s \right| \\
\leqslant &\sup_{t\in[0,T]} \left| \int_{\mathbb{R}^{2d}} \frac{(\varphi(y)-\varphi(x))}{2} \eta_t(x,y) V_t\left[\rho_t\right]_+ (x,y) d\rho_t^n(x) d(\mu-\mu^n)(y) \right| T \\
&+ \sup_{t\in[0,T]} \left| \int_{\mathbb{R}^{2d}} \frac{(\varphi(y)-\varphi(x))}{2} \eta_t(x,y) V_t\left[\rho_t\right]_- (x,y) d(\mu-\mu^n)(x) d\rho_t^n(y) \right| T \\
=: &\, II_1^n + II_2^n.
\end{aligned}
$$

As previously mentioned, by means of analogous calculations to those in the proof of Lemma 2.3.1 and Theorem 2.4.1 we now compare ρ_t and ρ_t^n and obtain

$$
\begin{aligned}
\|\rho_t - \rho_t^n\|_{\mathrm{TV}} \leqslant &\|\rho_0 - \rho_0^n\|_{\mathrm{TV}} + 2L_\Phi C_V \|\eta\|_\infty \int_0^t \|\rho_s - \rho_s^n\|_{\mathrm{TV}} \mathrm{d}s \\
&+ 2L_\Phi C_V M \int_0^t \|\eta_s - \eta_s^n\|_\infty \mathrm{d}s \\
&+ 2L_\Phi L_V M M_\eta \int_0^t \|\rho_s - \rho_s^n\|_{\mathrm{TV}} \mathrm{d}s \\
&+ \sup_{\varphi\in C_0(\mathbb{R}^d), \|\varphi\|_\infty \leqslant 1} (II_1^n + II_2^n),
\end{aligned}
$$

where $M_\eta := \sup_{n\in\mathbb{N}} \|\eta^n\|_{\infty, C_b(\mathbb{R}^{2d})}$. On the other hand, using the explicit solution for the ODE describing the dynamics of η^n and η as well as the Lipschitz condition ($\boldsymbol{\omega}2$) yields the following estimate

$$
\|\eta_t - \eta_t^n\|_\infty \leqslant \|\eta_0 - \eta_0^n\|_\infty + \int_0^t \|\eta_s - \eta_s^n\|_\infty \mathrm{d}s + L_\omega \int_0^t \|\rho_s - \rho_s^n\|_{\mathrm{TV}} \mathrm{d}s.
$$

Setting

$$
u^n(t) := \|\rho_t - \rho_t^n\|_{\mathrm{TV}} + \|\eta_t - \eta_t^n\|_\infty \ ,
$$

an application of Grönwall's inequality yields

$$u^n(t) \leqslant \big(\|\rho_0 - \rho_0^n\|_{\mathrm{TV}} + \|\eta_0 - \eta_0^n\|_\infty + \sup_{\varphi \in C_0(\mathbb{R}^d), \|\varphi\|_\infty \leqslant 1} (II_1^n + II_2^n)\big) e^{CT} \ ,$$

for a suitable constant $C \in (0, \infty)$ depending on all the others. We now turn our attention to II_1^n and II_2^n. Since $\{\rho_t^n\}_{n\in\mathbb{N}}$ is a sequence of solutions, we have that for any $t \in [0,T]$ and any $n \in \mathbb{N}$, $\|\rho_t^n\|_{TV,\infty} \leqslant M$. Then, by [AFP00, Theorem 1.59] we can extract a weakly-* converging subsequence, denoted again, with a slight abuse of notation, by $\{\rho_t^n\}_{n\in\mathbb{N}}$ converging to some $\bar{\rho}_t \in \mathcal{M}_{\mathrm{TV}}(\mathbb{R}^d)$. Finally, note that the integrands in II_1^n and II_2^n are both in $C_0(\mathbb{R}^{2d})$ due to our assumptions. Together with the assumption $\mu^n \overset{*}{\rightharpoonup} \mu$, we have that, up to a subsequence II_1^n and II_2^n converge to 0 as $n \to \infty$. Having the convergence of the initial conditions by our assumptions, by taking the supremum over $t \in [0,T]$ we let $n \to \infty$, which yields the result up to passing to a subsequence. Using that the limit is the unique solution to (Co-NCL), we infer the convergence holds for the whole sequence. □

Remark 2.5.2. Theorem 2.5.1 can be obtained for other admissible interpolations Φ that allow to estimate $II \leqslant II_1^n + II_2^n$ as in the proof above. For instance, we can have

$$\Phi(a,b;w) = \sum_{i=1}^{M} g_i(w)(\alpha_i a + \beta_i b), \qquad \text{for } (a,b,w) \in \mathbb{R}^3,$$

where $M \in \mathbb{N}$, $\alpha_i, \beta_i \in \mathbb{R}$, g_i Lipschitz for any $i = 1, \dots, M$. In case of the upwind interpolation $M = 2$, $g_1(w) = w_+$, $\alpha_1 = 1$, $\beta_1 = 0$, $g_2(w) = w_-$, $\alpha_2 = 0$, $\beta_2 = -1$. Another example would be the arithmetic mean, for which $M = 1$, $g_1(w) = w$, $\alpha_1 = \beta_1 = 1/2$, and similarly for other suitable mean multipliers.

Chapter 3

The graph continuity equation

3.1 Introduction

In this chapter, we establish a link between (Co-NCL) and a continuity equation on Euclidean space. To do so, we first link (Co-NCL) with a nonlinear Euler equation (see (Euler Co-NCL) below), for which we study some of its properties and establish its well-posedness. We embed the solutions of (Euler Co-NCL) in the space of probability measures over $\mathbb{R} \times \mathbb{R}^d$ with second marginal μ, where μ is the base measure of a co-evolving graph. We obtain that this measure solves a particular nonlinear continuity equation whose velocity field is given by the velocity field of the equation for the mass in (Euler Co-NCL). We study the well-posedness of this equation by using the method of characteristics and then applying the Banach fixed-point theorem. We conclude the chapter by discussing a failed strategy attempting to use this equation to obtain the long-time behaviour of (Euler Co-NCL) and discuss the obstacles to obtaining a discrete-to-continuum limit in this framework. This chapter is part of a work in preparation [CEM24].

3.1.1 Euler Equation

We focus on the case when μ defines the set of all possible vertices so that it makes sense to consider $\rho \ll \mu$. In particular, for any $t \in [0,T]$, $\rho_t := r_t \mathrm{d}\mu$, for a curve $r \in C([0,T], L^2_\mu(\mathbb{R}^d))$, and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. In this setting, the nonlocal continuity equation in (Co-NCL) can be rewritten so that it can be interpreted as a nonlinear Euler equation. We observe that in this set up there is an interesting comparison between the study carried out in [EPSS21, EPS23, EM24] and the recent developments on heterogeneous interacting particle systems, e.g., in [PT24, AD24].

Indeed, first we note that, by choosing $\lambda = \mu \otimes \mu$ in Definition 2.2.5, we can write the flux as

$$\mathrm{d}F^{\Phi}[\mu, \eta; g, w](x, y) = \Phi(g(x), g(y); w(x, y))\eta(x, y)\mathrm{d}(\mu \otimes \mu)(x, y),$$

for any $g \in L^2_\mu(\mathbb{R}^d)$, $w \in \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$, and $(x, y) \in \mathbb{R}^{2d}_{\diagup}$, so that Φ is a function of the μ-density g. Furthermore, if we assume that Φ is jointly antisymmetric, that is, $\Phi(a, b; -v) = -\Phi(b, a; v)$ for any $a, b, v \in \mathbb{R}$, and η symmetric, the nonlocal divergence of $F^{\Phi}[\mu, \eta; g, V[g]]$ is given by

$$\overline{\nabla} \cdot F[\mu, \eta; g, V[g]](x) = \int_{\mathbb{R}^d \setminus \{x\}} \Phi(g(x), g(y); V[g](x, y))\eta(x, y)\mathrm{d}\mu(y) \ , \tag{3.1.1}$$

for μ-a.e. $x \in \mathbb{R}^d$. In view of this consideration, we shall consider the system

$$\begin{aligned} \partial_t r_t(x) &= -\int_{\mathbb{R}^d \setminus \{x\}} \Phi(r_t(x), r_t(y); V_s[r_t](x, y))\eta_t(x, y)\mathrm{d}\mu(y) \\ \partial_t \eta_t &= \omega_t[r] - \eta_t, \end{aligned} \tag{Euler Co-NCL}$$

for given initial data $r^0 \in L^2_\mu(\mathbb{R}^d)$ and $\eta^0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. Solutions of (Euler Co-NCL) are defined as follows.

Definition 3.1.1 (Solution to (Euler Co-NCL))**.** Given a jointly antisymmetric admissible flux interpolation Φ, a velocity field $V : [0, T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$, and a function $\omega : [0, T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$, a pair $(r, \eta) : [0, T] \to L^2_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ is a solution of (Euler Co-NCL) with initial datum $(r_0, \eta_0) \in L^2_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^d)$ if it satisfies the following properties:

1. for μ-a.e. $x \in \mathbb{R}^d$, the maps $t \mapsto \int_{\mathbb{R}^d \setminus \{x\}} \Phi(r_t(x), r_t(y); V_s[r_t](x, y))\eta_t(x, y)\mathrm{d}\mu(y)$ and $t \mapsto \omega_t[\rho_t] - \eta_t$ belong to $L^1([0, T])$;

2. for μ-a.e. $x \in \mathbb{R}^d$, $t \in [0, T]$, $(x, y) \in \mathbb{R}^{2d}_{\diagup}$, it holds that

$$r_t(x) = r_0(x) - \int_0^t \int_{\mathbb{R}^d \setminus \{x\}} \Phi(r_s(x), r_s(y); V_s[r_s](x, y))\eta_s(x, y)\mathrm{d}\mu(y)\mathrm{d}s \tag{3.1.2a}$$

$$\eta_t(x, y) = \eta_0(x, y) + \int_0^t (\omega_s[r](x, y) - \eta_s(x, y))\ \mathrm{d}s. \tag{3.1.2b}$$

3. $r \in AC([0, T], L^2_\mu(\mathbb{R}^d))$, $\eta \in AC([0, T], C_b(\mathbb{R}^{2d}_{\diagup}))$.

We recall that our guiding examples for the functions V and ω are given by

$$\omega_t[r_t](x, y) = \int_{\mathbb{R}^d} W(t, x, y, z)r(z)\mathrm{d}\mu(z), \tag{3.1.3}$$

$$V_t[r_t](x, y) = -\int_{\mathbb{R}^d} (K(y, z) - K(x, z))r(z)\mathrm{d}\mu(z), \tag{3.1.4}$$

where for $(x,y) \in \mathbb{R}^{2d}_{\diagup}$, the map $t \mapsto W(t,\cdot,\cdot,\cdot) \in L^1([0,T])$, and $W \in C_b([0,T] \times \mathbb{R}^{2d}_{\diagup} \times \mathbb{R}^d)$, $K \in C_b(\mathbb{R}^3)$ and the velocity defined in (3.1.4) is related to the nonlocal interaction energy $\mathcal{E}(\rho) = \frac{1}{2}\iint K(x,y)\mathrm{d}\rho(x)\mathrm{d}\rho(y)$.

We obtain well-posedness of (Euler Co-NCL) by exploiting the Banach fixed-point theorem in the space $C([0,T], L^2_\mu(\mathbb{R}^d)) \times C([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$ equipped with the distance

$$\tilde{d}_\infty((f^1,g^1),(f^2,g^2)) = \|f^1 - f^2\|_{\infty, L^2_\mu(\mathbb{R}^d)} + \|g^1 - g^2\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})},$$

where

$$\|f^1 - f^2\|_{\infty, L^2_\mu(\mathbb{R}^d)} := \sup_{t\in[0,T]} \|f^1_t - f^2_t\|_{L^2_\mu(\mathbb{R}^d)} \ ,$$
$$\|g^1 - g^2\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} := \sup_{t\in[0,T]} \|g^1_t - g^2_t\|_{C_b(\mathbb{R}^{2d}_{\diagup})} \ .$$

We begin by proving *a priori* properties for (Euler Co-NCL) useful to define the solution maps.

Proposition 3.1.1. Let Φ be a jointly antisymmetric admissible flux interpolation, $r_0 \in L^2_\mu(\mathbb{R}^d)$, $\eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. Assume $\omega : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ is such that the map $(x,y) \in \mathbb{R}^{2d}_{\diagup} \mapsto \omega_t[\cdot](\cdot,x,y)$ is continuous and that for a constant $C_\omega > 0$

$$\int_0^T \sup_{r\in L^2_\mu(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}_{\diagup}} |\omega_s[r](x,y)|\mathrm{d}s \leqslant C_\omega \ .$$

Suppose that $V : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$ satisfies

$$\int_0^T \sup_{r\in L^2_\mu(\mathbb{R}^d)} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_t](x,y)|^2 \mathrm{d}\mu(y)\mathrm{d}t \leqslant C_V, \tag{3.1.5}$$

for some $C_V > 0$. For a pair $(r,\eta) : [0,T] \to L^2_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ satisfying (3.1.2) the following properties hold.

1. For μ-a.e. $x \in \mathbb{R}^d$ and every $(x,y) \in \mathbb{R}^{2d}_{\diagup}$, the maps

$$t \mapsto \int_{\mathbb{R}^d\setminus\{x\}} \Phi(r_t(x), r_t(y); V_s[r_s](x,y))\eta_s(x,x)\mathrm{d}\mu(y) \in L^1([0,T]),$$
$$t \mapsto \omega_t[r](x,y) - \eta_t(x,y) \in L^1([0,T]),$$

   and $r \in L^\infty([0,T]; L^2_\mu(\mathbb{R}^d)), \eta \in L^\infty([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$.

2. $r \in AC([0,T]; L^2_\mu(\mathbb{R}^d))$ and $\eta \in AC([0,T]; C_b(\mathbb{R}^{2d}_{\diagup}))$.

Proof. We obtain $\eta \in L^\infty([0,T]; C_b(\mathbb{R}^{2d}_{\not\Delta}))$ in an analogous way to Proposition 2.2.1, with

$$\|\eta_t\|_\infty \leqslant (\|\eta_0\|_\infty + C_\omega)\, e^T \ , \tag{3.1.6}$$

and note that the integrability of $t \mapsto \omega_t[r](x,y) - \eta_t(x,y)$ directly follows. As for r, for μ-a.e. $x \in \mathbb{R}^d$, by using (i), (ii) in Definition 2.2.3, boundedness of η, and (3.1.5) we have

$$\begin{aligned}
|r_t(x)| &\leqslant |r_0(x)| + \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} |\Phi(r_s(x), r_s(y); V_s[r_s](x,y))\eta_s(x,y)|\, \mathrm{d}\mu(y)\mathrm{d}s \\
&\leqslant |r_0(x)| + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_s](x,y)|\, (|r_s(x)| + |r_s(y)|)\mathrm{d}\mu(y)\mathrm{d}s \\
&\leqslant |r_0(x)| + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \int_0^t \left(\sup_{r\in L^2_\mu(\mathbb{R}^d)} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_s](x,y)|^2\, \mathrm{d}\mu(y) \right)^{\frac{1}{2}} \\
&\quad \times \left(\int_{\mathbb{R}^d\setminus\{x\}} (|r_s(x)| + |r_s(y)|)^2 \mathrm{d}\mu(y) \right)^{\frac{1}{2}} \mathrm{d}s \\
&\leqslant |r_0(x)| + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \int_0^t \bar{v}_s^{1/2} (\|r_s\|_{L^2_\mu(\mathbb{R}^d)} + |r_s(x)|)\mathrm{d}s \ ,
\end{aligned}$$

where we used the Cauchy–Schwarz and Minkowski inequalities, $\mu(\mathbb{R}^d) = 1$ and we set

$$\bar{v}_t := \sup_{r\in L^2_\mu(\mathbb{R}^d)} \sup_{x\in\mathbb{R}^d} \left(\int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_t](x,y)|^2 \mathrm{d}\mu(y) \right) \ . \tag{3.1.7}$$

An application of Minkowski's inequality and [Fol99, Theorem 6.19 a.] yields

$$\|r_t\|_{L^2_\mu(\mathbb{R}^d)} \leqslant \|r_0\|_{L^2_\mu(\mathbb{R}^d)} + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} 2 \int_0^t \bar{v}_s^{1/2} \|r_s\|_{L^2_\mu(\mathbb{R}^d)} \mathrm{d}s \ .$$

Then, an application of Grönwall's inequality yield, for any $t \in [0,T]$,

$$\|r_t\|_{L^2_\mu(\mathbb{R}^d)} \leqslant \|r_0\|_{L^2_\mu(\mathbb{R}^d)} e^{2L_\Phi \|\eta\|_\infty C_V}, \tag{3.1.8}$$

thus $r \in L^\infty([0,T]; L^2_\mu(\mathbb{R}^d))$. The integrability of the nonlocal divergence can be proven by a further application of Gronwall inequality. More precisely, by the previous argument we know that, for μ-a.e. $x \in \mathbb{R}^d$, More precisely, by the previous argument we know that, for μ-a.e.

$x \in \mathbb{R}^d$,

$$\begin{aligned}
|r_t(x)| &\leqslant |r_0(x)| + \int_0^t \int_{\mathbb{R}^d \setminus \{x\}} |\Phi(r_t(x), r_t(y); V_s[r_s](x,y))\eta_s(x,y)| \, \mathrm{d}\mu(y)\mathrm{d}s \\
&\leqslant |r_0(x)| + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \left(\int_0^t \|r_0\|_{L^2_\mu(\mathbb{R}^d)} e^{2L_\Phi \|\eta\|_\infty C_V \bar{v}_s^{1/2}} \mathrm{d}s + \int_0^t \bar{v}_s^{1/2} |r_s(x)| \mathrm{d}s \right) \\
&\leqslant |r_0(x)| + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \left(\|r_0\|_{L^2_\mu(\mathbb{R}^d)} e^{2L_\Phi \|\eta\|_\infty C_V} \left(t \int_0^t \bar{v}_s \mathrm{d}s \right)^{1/2} + \int_0^t \bar{v}_s^{1/2} |r_s(x)| \mathrm{d}s \right) \\
&\leqslant |r_0(x)| + \bar{C}\sqrt{t} + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \int_0^t \bar{v}_s^{1/2} |r_s(x)| \mathrm{d}s,
\end{aligned}$$

where the third inequality follows from an application of the Cauchy-Schwarz inequality and we denoted $\bar{C} = L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \|r_0\|_{L^2_\mu(\mathbb{R}^d)} e^{2L_\Phi \|\eta\|_\infty C_V} C_V^{1/2}$. Then, for any $t \in [0,T]$ and μ-a.e. $x \in \mathbb{R}^d$,

$$|r_t(x)| \leqslant (|r_0(x)| + \bar{C}) e^{L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \sqrt{t} \sqrt{C_V}},$$

where we have used Grönwall's inequality and the Cauchy-Schwarz inequality again. From this we infer time-integrability of the nonlocal divergence. Part (2) is consequence of the time integrability and (3.1.2). □

We proceed with proving well-posedness of (Euler Co-NCL). Fix $r_0 \in L^2_\mu(\mathbb{R}^d)$, $\eta_0 \in C_b(\mathbb{R}^{2d}_{\not\Delta})$ for η_0 symmetric, and assume Φ is a jointly antisymmetric admissible flux interpolation as in Definition 2.2.3. Furthermore, in addition to assumptions ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}3$) we will also require the following

$$(x,y) \mapsto \omega_\cdot[\cdot](x,y) \text{ is symmetric.} \tag{$\tilde{\boldsymbol{\omega}}4$}$$

Assumption ($\tilde{\boldsymbol{\omega}}4$) together with the symmetry of η_0 are required to ensure the symmetry of η_t for $t \in [0,T]$ (see Lemma 4.3.4), which, in turn, is needed for the nonlocal divergence of (Co-NCL) to be as in (3.1.1).

With a slight abuse of notation, throughout this chapter we again denote by

$$\mathcal{AC}_T := AC([0,T]; L^2_\mu(\mathbb{R}^d)) \times AC([0,T]; C_b(\mathbb{R}^{2d}_{\not\Delta})),$$

and equip it with the distance $\tilde{d}_\infty$. We define the solution map $\mathcal{S} := (\mathcal{S}_V, \mathcal{S}_\omega) : \mathcal{AC}_T \to \mathcal{AC}_T$ as

$$\begin{aligned}
\mathcal{S}_V(r,\eta)(t)(x) &:= r_0(x) - \int_0^t \int_{\mathbb{R}^d \setminus \{x\}} \Phi(r_s(x), r_s(y); V_s[r_s](x,y))\eta_s(x,y) \mathrm{d}\mu(y)\mathrm{d}s, \quad \mu\text{-a.e. } x \in \mathbb{R}^d, \\
\mathcal{S}_\omega(r,\eta)(t)(x,y) &:= \eta_0(x,y) + \int_0^t (\omega_s[r](x,y) - \eta_s(x,y)) \, \mathrm{d}s \ ,
\end{aligned}$$

for any $t \in [0,T]$ and $(x,y) \in \mathbb{R}^{2d}_{\not\Delta}$. The a priori properties in Proposition 3.1.1 ensure this map

$\mathcal{S}$ is well defined. Below, we prove existence and uniqueness of solutions to (Euler Co-NCL) by assuming a uniform-in-time integrability assumption for the velocity field V motivated by our guiding example, see (3.1.4), although the result holds for V satisfying (3.1.5). The stronger assumption becomes important in Section 3.2, where it will be important to obtain the Lipschitzness of the characteristics.

Theorem 3.1.1. Let Φ be a jointly antisymmetric admissible flux interpolation as in Definition 2.2.3 and consider an initial datum $(r^0, \eta^0) \in L^2_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$. Assume that $V : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^d)$ satisfies (3.1.5) or

$$\sup_{t\in[0,T]} \sup_{r\in L^2_\mu(\mathbb{R}^d)} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_t](x,x')|^2 \mathrm{d}\mu(x') \leqslant C_V, \tag{3.1.9a}$$

and, for all $t \in [0,T]$ and any $g, \tilde{g} \in L^2_\mu(\mathbb{R}^d)$,

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_t[g](x,x') - V_t[\tilde{g}](x,x')|^2 \mathrm{d}\mu(x') \leqslant L_V \|g-\tilde{g}\|^2_{L^2_\mu(\mathbb{R}^d)}, \tag{3.1.9b}$$

for constants $C_V, L_V > 0$. Assume that $\omega : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ satisfies ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}3$) and ($\tilde{\boldsymbol{\omega}}4$). Then, there exists a solution (r,η) to (Euler Co-NCL) with initial data (r^0,η^0), in the sense of Definition 3.1.1 .

Proof. Let us assume (3.1.9a) since the argument is slightly more interesting for this case towards the end of the proof. For an initial datum (r^0,η^0), we consider $(r,\eta), (\tilde{r},\tilde{\eta}) \in \mathcal{AC}_T$ and, for μ-a.e. $x \in \mathbb{R}^d$, we have

$$\begin{aligned}
&|\mathcal{S}_V(r,\eta)(t) - \mathcal{S}_V(\tilde{r},\tilde{\eta})(t)|(x) \\
&\leqslant \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} \Big|\Phi(r_s(x), r_s(y); V_s[r_s](x,y))\eta_s(x,y) - \Phi(\tilde{r}_s(x),\tilde{r}_s(y); V_s[\tilde{r}](x,y))\tilde{\eta}_s(x,y)\Big| \mathrm{d}\mu(y)\mathrm{d}s \\
&\leqslant L_\Phi \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} \big|\eta_s(x,y)\big|(|r_s(x)-\tilde{r}_s(x)| + |r_s(y)-\tilde{r}_s(y)|)|V_s[r_s](x,y)|\mathrm{d}\mu(y)\mathrm{d}s \\
&\quad + L_\Phi \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} (|\tilde{r}_s(x)| + |\tilde{r}_s(y)|)\,|V_s[r_s](x,y)|\big|\eta_s(x,y) - \tilde{\eta}_s(x.y)\big|\mathrm{d}\mu(y)\mathrm{d}s \\
&\quad + L_\Phi \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} \big|\tilde{\eta}_s(x,y)\big|\,(|\tilde{r}_s(x)| + |\tilde{r}_s(y)|)\,|V_s[r_s](x,y) - V_s[\tilde{r}](x,y)|\mathrm{d}\mu(y)\mathrm{d}s \\
&=: I + II + III,
\end{aligned}$$

where we used (i), (ii) in Definition 2.2.3. I can be estimated as follows by means of Cauchy–

Schwarz and Minkowski inequalities

$$
\begin{aligned}
|I| &\leqslant L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} \left(\sup_{t\in[0,T]} \sup_{r\in L^2_\mu(\mathbb{R}^d)} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_t](x,y)|^2 \,\mathrm{d}\mu(y) \right)^{1/2} \\
&\qquad \times \int_0^t \left(\int_{\mathbb{R}^d\setminus\{x\}} (|r_s(x)-\tilde r_s(x)| + |r_s(y)-\tilde r_s(y)|)^2 \,\mathrm{d}\mu(y) \right)^{1/2} \mathrm{d}s \qquad (3.1.10) \\
&\leqslant L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V^{1/2} \int_0^t (\|r_s-\tilde r_s\|_{L^2_\mu(\mathbb{R}^d)} + |r_s(x)-\tilde r_s(x)|)\mathrm{d}s \ ,
\end{aligned}
$$

where in the second inequality we used (3.1.9a). Squaring both sides and integrating with respect to μ

$$
\begin{aligned}
\|I\|^2_{L^2_\mu(\mathbb{R}^d)} &\leqslant L_\Phi^2 \|\eta\|^2_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V \int_{\mathbb{R}^d} \left| \int_0^t (\|r_s-\tilde r_s\|_{L^2_\mu(\mathbb{R}^d)} + |r_s(x)-\tilde r_s(x)|)\mathrm{d}s \right|^2 \mathrm{d}\mu(x) \\
&\leqslant L_\Phi^2 \|\eta\|^2_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V T \int_{\mathbb{R}^d} \int_0^T (\|r_s-\tilde r_s\|_{L^2_\mu(\mathbb{R}^d)} + |r_s(x)-\tilde r_s(x)|)^2 \mathrm{d}s\mathrm{d}\mu(x) \\
&\leqslant L_\Phi^2 \|\eta\|^2_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} 2C_V T \int_0^T \|r_s-\tilde r_s\|^2_{L^2_\mu(\mathbb{R}^d)} \mathrm{d}s \\
&\leqslant L_\Phi^2 \|\eta\|^2_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} 2C_V T^2 \|r-\tilde r\|^2_{\infty, L^2_\mu(\mathbb{R}^d)},
\end{aligned}
$$

thus

$$
\|I\|_{L^2_\mu(\mathbb{R}^d)} \leqslant \sqrt{2} L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V^{1/2} \|r-\tilde r\|_{\infty, L^2_\mu(\mathbb{R}^d)} T \ .
$$

For II we obtain similarly that

$$
|II| \leqslant L_\Phi \|\eta-\tilde\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V^{1/2} \int_0^t (\|\tilde r_s\|_{L^2_\mu(\mathbb{R}^d)} + |r_s(x)|)\mathrm{d}s \ ,
$$

and thus

$$
\begin{aligned}
\|II\|_{L^2_\mu(\mathbb{R}^d)} &\leqslant L_\Phi \|\eta-\tilde\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V^{1/2} (\|r\|_{\infty, L^2_\mu(\mathbb{R}^d)} + \|\tilde r\|_{\infty, L^2_\mu(\mathbb{R}^d)}) T \\
&\leqslant 2L_\Phi \|\eta-\tilde\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} C_V^{1/2} \bar C_T T \ ,
\end{aligned}
$$

where in the last inequality $\bar C_T = \|r_0\|_{L^2_\mu(\mathbb{R}^d)} e^{2L_\Phi(\|\eta_0\|_\infty + C_\omega T)e^T C_V T}$, from (3.1.8). Finally, for III we use (3.1.9b) to have

$$
\begin{aligned}
|III| &\leqslant L_\Phi \|\tilde\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} \left(\sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d\setminus\{x\}} |V_s[r_t](x,y) - V_t[\tilde r](x,y)|^2 \mathrm{d}\mu(y) \right)^{1/2} \\
&\qquad \times \int_0^t \|\tilde r_s\|_{L^2_\mu(\mathbb{R}^d)} + |\tilde r_s(x)| \mathrm{d}s \\
&\leqslant L_\Phi \|\tilde\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\not/})} L_V^{1/2} \|r-\tilde r\|_{\infty, L^2_\mu(\mathbb{R}^d)} \int_0^t \|\tilde r_s\|_{L^2_\mu(\mathbb{R}^d)} + |\tilde r_s(x)| \mathrm{d}s.
\end{aligned}
$$

Therefore,

$$\begin{aligned}\|III\|_{L^2_\mu(\mathbb{R}^d)} &\leqslant \sqrt{2}L_\Phi\|\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}L_V^{1/2}\|r\|_{\infty,L^2_\mu(\mathbb{R}^d)}\|r-\tilde{r}\|_{\infty,L^2_\mu(\mathbb{R}^d)}T\\ &\leqslant \sqrt{2}L_\Phi\|\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}L_V^{1/2}\|r-\tilde{r}\|_{\infty,L^2_\mu(\mathbb{R}^d)}\bar{C}_T T.\end{aligned}$$

The same estimate for $\mathcal{S}_\omega$ gives

$$\|\mathcal{S}_\omega(r,\eta)-\mathcal{S}_\omega(\tilde{r},\tilde{\eta})\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})} \leqslant \|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}T + L_\omega\|r-\tilde{r}\|_{\infty,L^2_\mu(\mathbb{R}^d)}T,$$

due to ($\boldsymbol{\omega}2$) and Proposition 3.1.1, see also Lemma 2.3.1. A combination of the previous estimates implies that

$$\begin{aligned}d_\infty((r,\eta),(\tilde{r},\tilde{\eta})) &\leqslant (\sqrt{2}L_\Phi\|\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}(L_V^{1/2}\bar{C}_T + C_V^{1/2}) + L_\omega))\|r-\tilde{r}\|_{\infty,L^2_\mu(\mathbb{R}^d)}T\\ &\quad + (1+2L_\Phi C_V^{1/2}\bar{C}_T)\|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}T\\ &\leqslant \underbrace{(\sqrt{2}L_\Phi(\|\eta_0\|_\infty + C_\omega T)e^T(L_V^{1/2}\bar{C}_T + C_V^{1/2}) + L_\omega))}_{=:\alpha(T)}\|r-\tilde{r}\|_{\infty,L^2_\mu(\mathbb{R}^d)}T\\ &\quad + \underbrace{(1+2L_\Phi C_V^{1/2}\bar{C}_T)}_{=:\beta(T)}\|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}T\end{aligned}$$

where we used (3.1.6) with ($\boldsymbol{\omega}3$). In particular, it follows that

$$d_\infty(\mathcal{S}(r,\eta),\mathcal{S}(\tilde{r},\tilde{\eta})) \leqslant \max\{\alpha(T),\beta(T)\}Td_\infty((r,\eta),(\tilde{r},\tilde{\eta}))$$

so that for $T<\kappa(T):=1/\max\{\alpha(T),\beta(T)\}$ the solution map is a contraction. By the Banach fixed-point theorem we can conclude there exists a unique solution to (Euler Co-NCL) in $C_T = C([0,T];L^2_\mu(\mathbb{R}^d))\times C([0,T];C_b(\mathbb{R}^{2d}_{\not/}))$ for $T<T^*$, where T^* is such that $\kappa(T^*)^{-1}T^*=1$. This can be verified since $\alpha(T)\simeq 1+e^T+Te^T$ and $\beta(T)\simeq 1+e^T$. Absolute continuity in time follows since $\overline{\nabla}\cdot F[\mu,\eta;r,V[r]]$ and the map $(t\mapsto \omega_t[r_t]-\eta_t)$ belong to $L^1([0,T])$, as proven in Proposition 3.1.1. The solution can be extended by using a standard iteration procedure so that the time interval does not depend on any T^*, see e.g. Theorem 2.3.1.

Now we turn to the analysis of the case V satisfying (3.1.5). Following the same strategy, upon using (3.1.5) instead of (3.1.9a), we obtain the bound

$$\begin{aligned}d_\infty((r,\eta),(\tilde{r},\tilde{\eta})) &\leqslant \underbrace{(\sqrt{2}L_\Phi(\|\eta_0\|_\infty + C_\omega T)e^T(L_V^{1/2}\bar{C} + C_V^{1/2}) + L_\omega))}_{=:\alpha(T)}\|r-\tilde{r}\|_{\infty,L^2_\mu(\mathbb{R}^d)}(T+\sqrt{T})\\ &\quad + \underbrace{(1+2L_\Phi C_V^{1/2}\bar{C})}_{=:\beta}\|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\not/})}(T+\sqrt{T})\ ,\end{aligned}$$

where $\bar{C} = \|r_0\|_{L^2_\mu(\mathbb{R}^d)} e^{2L_\Phi(\|\eta_0\|_\infty + C_\omega T)e^T C_V}$ and thus

$$d_\infty(\mathcal{S}(r,\eta),\mathcal{S}(\tilde{r},\tilde{\eta})) \leqslant \max\{\alpha(T),\beta\}(T+\sqrt{T})d_\infty((r,\eta),(\tilde{r},\tilde{\eta}))\ ,$$

so that for $T + \sqrt{T} < \kappa(T) := 1/\max\{\alpha(T),\beta\}$ the solution map is a contraction. The rest of the argument follows the previous case in a completely analogous way. □

Remark 3.1.1. The assumptions on the joint antisymmetry of Φ and the symmetry of η through $(\tilde{\omega}4)$ and η_0 symmetric in Theorem 3.1.1 are needed so that (Co-NCL) can be expressed as (Euler Co-NCL), see [EPS23, Section 5]. However, these assumptions are not required to obtain well-posedness of the dynamics described by (Euler Co-NCL) itself.

Having proved the well-posedness of (Euler Co-NCL) we turn to the study of the graph continuity equation in the next section.

3.2 The graph continuity equation

Throughout this section we consider a co-evolving graph, $\mathcal{G}_t = (\mu, \eta_t[r_t])$, for (r,η) a solution to (Euler Co-NCL) in the sense of Definition 3.1.1, which exists and is unique due to Theorem 3.1.1. We observe that (Euler Co-NCL) is linked to a particular nonlinear continuity equation if we consider the evolution of the measure $\delta_{r_t} \otimes \mu$, see also [PT24, Section 3.1.3] for a similar case. More precisely, for any Borel set $A \times B \in \mathcal{B}(\mathbb{R}\times\mathbb{R}^d)$ let us consider the measure $\sigma := \delta_r \otimes \mu$ given by

$$(\delta_r \otimes \mu)(A\times B) = \int_B \delta_{r(x)}(A)\mathrm{d}\mu(x), \tag{3.2.1}$$

where (r,η) is the solution of (Euler Co-NCL) with base measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. One can check that $\delta_{r_t}\otimes\mu$ is a distributional solution (in the sense of Definition 3.2.1) of the following initial value problem

$$\begin{cases} \partial_t \sigma + \partial_\xi(\sigma \mathcal{X}[\sigma, r, \eta]) = 0, \\ \sigma_0 = \bar{\sigma} \in \mathcal{P}(\mathbb{R}\times\mathbb{R}^d), \end{cases} \tag{graph-CE}$$

where, for any $\sigma \in \mathcal{P}(\mathbb{R}\times\mathbb{R}^d)$ the velocity field is

$$\mathcal{X}[\sigma,r,\eta](t,\xi,x) := -\int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(\xi,\xi';V_t[r_t](x,x'))\eta_t(x,x')\mathrm{d}\sigma(\xi',x')\ , \tag{3.2.2}$$

for a solution (r,η) to (Euler Co-NCL). We refer to (graph-CE) as the *graph-continuity equation* since the velocity field depends on the solution of the co-evolving graph problem. We devote

the rest of this chapter to the study of its well-posedness, first. Solutions to (graph-CE) are understood in the distributional sense.

Definition 3.2.1. Let $\mathcal{G}_t(\mu, \eta_t[r_t])$ be a given co-evolving graph, where (r, η) solve (Euler Co-NCL). A curve $\sigma \in C([0,T], \mathcal{P}(\mathbb{R}\times\mathbb{R}^d))$ is a distributional solution to (graph-CE) with initial data $\sigma_0 \in \mathcal{P}(\mathbb{R}\times\mathbb{R}^d)$ if, for any $\varphi \in C_c^1([0,T]\times\mathbb{R}\times\mathbb{R}^d)$, it holds

$$\begin{aligned}&\int_0^T \int_{\mathbb{R}\times\mathbb{R}^d} \partial_t\varphi(t,\xi,x) d\sigma_t(\xi,x)\mathrm{d}t + \int_{\mathbb{R}\times\mathbb{R}^d} \varphi(0,\xi,x)\mathrm{d}\sigma_0(\xi,x) - \int_{\mathbb{R}\times\mathbb{R}^d} \varphi(T,x,\xi)\mathrm{d}\sigma_T(\xi,x)\\ &= \int_0^T \int_{\mathbb{R}\times\mathbb{R}^d}\int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \partial_\xi\varphi(t,\xi,x)\Phi(\xi,\xi';V_t[r_t](x,x'))\eta_t(x,x')\mathrm{d}\sigma_t(\xi',x')\mathrm{d}\sigma_t(\xi,x)\mathrm{d}t\ .\end{aligned} \tag{3.2.3}$$

We are interested in evolutions on graphs, whose possible vertices are given by μ, which is constant in time. For this reason, it makes sense to consider the disintegration of a solution σ_t to (graph-CE) of the form (3.2.1) with respect to the marginal μ on $\mathbb{R}^d$, i.e. $\sigma_t = \int_{\mathbb{R}^d} \sigma_{t,x}\mathrm{d}\mu(x)$. In what follows, we will denote by $\pi^1 : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}$ and $\pi^2 : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$ the canonical projections given by $\pi^1(x,x') = x$ and $\pi^2(x,x') = x'$ for any $(x,x') \in \mathbb{R}\times\mathbb{R}^d$, so that $\mu = \pi^2_\#\sigma$ for any measure σ on $\mathbb{R}\times\mathbb{R}^d$ with marginal μ on $\mathbb{R}^d$. By uniqueness μ-almost everywhere of the disintegration we consider the *disintegrated version* of (graph-CE)

$$\begin{cases} \partial_t\sigma_{t,x} + \partial_\xi(\sigma_{t,x}\mathcal{X}[\sigma,r,\eta](t,\xi,x)) = 0,\\ \sigma_{0,x} = \bar\sigma_x, \end{cases} \tag{3.2.4}$$

for μ-almost every $x \in \mathbb{R}^d$ and $\bar\sigma = \int_{\mathbb{R}^d}\bar\sigma_x\mathrm{d}\mu(x)$. For μ-a.e. $x \in \mathbb{R}^d$ we interpret the above equation in the distributional sense, i.e., for any $\varphi \in C_c^1([0,T]\times\mathbb{R}\times\mathbb{R}^d)$,

$$\begin{aligned}&\int_0^T \int_{\mathbb{R}} \partial_t\varphi(t,\xi,x) d\sigma_{t,x}(\xi)\mathrm{d}t + \int_{\mathbb{R}} \varphi(0,\xi,x)\mathrm{d}\sigma_{0,x}(\xi) - \int_{\mathbb{R}} \varphi(T,x,\xi)\mathrm{d}\sigma_{T,x}(\xi)\\ &\quad= \int_0^T \int_{\mathbb{R}}\int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \partial_\xi\varphi(t,\xi,x)\Phi(\xi,\xi';V_t[r_t](x,x'))\eta_t(x,x')\mathrm{d}\sigma_t(\xi',x')\mathrm{d}\sigma_{t,x}(\xi)\mathrm{d}t.\end{aligned} \tag{3.2.5}$$

Note that integrating (3.2.5) with respect to μ we obtain $\sigma_t = \int_{\mathbb{R}^d}\sigma_{t,x}\mathrm{d}\mu(x)$ is a distributional solution to (graph-CE). Due to the nature of our problem, we introduce the notation

$$\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d) := \{\nu \in \mathcal{P}_2(\mathbb{R}\times\mathbb{R}^d) \mid {\pi^2}_\#\nu = \mu\}\ ,$$

for measures in $\mathcal{P}_2(\mathbb{R}\times\mathbb{R}^d)$ having the marginal μ on $\mathbb{R}^d$. This means we will consider the embedding of the dynamics (Euler Co-NCL) into (graph-CE) for graphs with the same base

measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. We equip $\mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d)$ with the distance

$$L^2_\mu d_2(\nu^1, \nu^2) := \left(\int_{\mathbb{R}^d} d_2^2(\nu^1_x, \nu^2_x) \mathrm{d}\mu(x) \right)^{\frac{1}{2}} , \quad \nu^1, \nu^2 \in \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d) ,$$

which forms a complete metric space, see Proposition A.0.1 in Appendix A for the formal statement and proof. For $\sigma \in \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d)$, we denote by $m_i[\sigma_{t,x}] := \int_{\mathbb{R}} |\xi|^i \mathrm{d}\sigma_{t,x}(\xi)$ μ-a.e. $x \in \mathbb{R}^d$, and $t \in [0,T]$ the i-th moment of the disintegration of σ with respect to μ, for $i = 1, 2$, and we will denote by $m_i[\gamma_*] := \sup_{t\in[0,T]} m_i[\gamma_t]$ the bound in time for the i-th moment of a curve $\gamma \in C([0,T]; \mathcal{P}(\mathbb{R}))$. In order to prove existence and uniqueness of weak solutions we proceed by looking at the characteristics' equation.

Lemma 3.2.2. Let $V : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ be continuous in time and assume it satisfies (3.1.9a) and (3.1.9b). Let $\sigma \in C([0,T], \mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d))$ with ${\pi^1}_\# \sigma = \gamma \in C([0,T], \mathcal{P}(\mathbb{R}))$ and (r, η) be the solution of (Euler Co-NCL). Then, for μ-a.e. $x \in \mathbb{R}^d$, the ODE initial value problem

$$\begin{aligned} \frac{d}{dt} v_t(x) &= \mathcal{X}[\sigma, r, \eta](t, v_t(x), x) , \\ v_0(x) &= u \in \mathbb{R} , \end{aligned}$$

has a unique solution on $[0,T]$.

Proof. We begin by showing that our assumptions imply the map $\mathbb{R} \ni \xi \mapsto \mathcal{X}[\sigma, r, \eta](t, \xi, x)$ is globally Lipschitz, for any $(t, x) \in [0,T] \times \mathbb{R}^d$. Indeed, for any $a, b \in \mathbb{R}$ we use (i), (ii) from Definition 2.2.3, and (3.1.9a) to obtain

$$\begin{aligned} &|\mathcal{X}[\sigma, r, \eta](t, a, x) - \mathcal{X}[\sigma, r, \eta](t, b, x)| \\ &\quad \leqslant \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Big| (\Phi(a, \xi'; V_s[r_t](x, x')) - \Phi(b, \xi'; V_s[r_t](x, x'))\eta_t(x, x') \Big| \mathrm{d}\sigma_t(\xi', x') \\ &\quad \leqslant L_\Phi \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} |\eta_t(x, x')| \, |V_s[r_t](x, x')| \, |a - b| \mathrm{d}\sigma_{t,x'}(\xi') \mathrm{d}\mu(x') \\ &\quad \leqslant L_\Phi C_V^{1/2} \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} |a - b| . \end{aligned} \tag{3.2.6}$$

Let us check that $t \mapsto \mathcal{X}[\sigma, r, \eta](t, a, x)$ is continuous. Without loss of generality, for any

$(a,x) \in \mathbb{R} \times \mathbb{R}^d$ and $s < t \in [0,T]$, consider

$$
\begin{aligned}
&|\mathcal{X}[\sigma,r,\eta](t,a,x) - \mathcal{X}[\sigma,r,\eta](s,a,x)| \\
&= \left| \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x')\mathrm{d}\sigma_t(\xi',x') - \Phi(a,\xi';V_s[r_s](x,x'))\eta_s(x,x')\mathrm{d}\sigma_s(\xi',x') \right| \\
&\leqslant \left| \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x')\mathrm{d}(\sigma_t-\sigma_s)(\xi',x') \right| \\
&\qquad + \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \left| \Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x') - \Phi(a,\xi';V_s[r_s](x,x'))\eta_s(x,x') \right| \mathrm{d}\sigma_s(\xi',x') \\
&=: I + II \ .
\end{aligned}
$$

We begin by estimating I. Note that for $(\mu\otimes\mu)$-a.e. $(x,x') \in \mathbb{R}^{2d}_{\diagup}$, $a \in \mathbb{R}$, and any $t \in [0,T]$ we have

$$
\begin{aligned}
&Lip[\Phi(a,\cdot,V_s[r_t](x,x'))\eta_t(x,x')] \\
&\qquad := \sup_{\xi^1\neq\xi^2\in\mathbb{R}} \frac{|(\Phi(a,\xi^1;V_s[r_t](x,x')) - \Phi(a,\xi^2;V_s[r_t](x,x')))\eta_t(x,x')|}{|\xi^1-\xi^2|} \\
&\qquad \leqslant L_\Phi |\eta_t(x,x')||V_s[r_t](x,x')| \\
&\qquad \leqslant L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} |V_s[r_t](x,x')| \ ,
\end{aligned}
\tag{3.2.7}
$$

similarly to (3.2.6). Thus, the Lipschitz constant of the map $\xi \mapsto \Phi(a,\xi;V_s[r_t](x,x'))\eta_t(x,x')$ is bounded by $L_\Phi\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}|V_s[r_t](x,x')|$ for $(\mu\otimes\mu)$-a.e. $(x,x') \in \mathbb{R}^d\times\mathbb{R}^d$ and any $t\in[0,T]$. Then we can estimate I by

$$
\begin{aligned}
|I| &= \left| \int_{\mathbb{R}^d}\int_{\mathbb{R}} \Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x')\mathrm{d}(\sigma_{t,x}-\sigma_{s,x})(\xi')\mathrm{d}\mu(x') \right| \\
&\leqslant \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}} \Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x')\mathrm{d}(\sigma_{t,x'}-\sigma_{s,x'})(\xi') \right| \mathrm{d}\mu(x') \\
&\leqslant \int_{\mathbb{R}^d} Lip[\Phi(a,\cdot,V_s[r_t](x,x')\eta_t(x,x')]d_1(\sigma_{t,x'},\sigma_{s,x'})\mathrm{d}\mu(x') \\
&\leqslant L_\Phi\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} \int_{\mathbb{R}^d} |V_s[r_t](x,x')| d_1(\sigma_{t,x'},\sigma_{s,x'})\mathrm{d}\mu(x') \\
&\leqslant L_\Phi\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})} C_V^{1/2} L_\mu^2 d_2(\sigma_t,\sigma_s),
\end{aligned}
$$

where we used Cauchy-Schwarz inequality and (3.1.9a). By the continuity in time of σ we have $I \to 0$ as $t \to s$. Let us now focus on II. For any $a \in \mathbb{R}$ and $x \in \mathbb{R}^d$, note that

$\Phi_t - \Phi_s := \Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x') - \Phi(a,\xi';V_s[r_s](x,x'))\eta_s(x,x')$ is integrable, since

$$
\begin{aligned}
|II| &\leqslant \int_{\mathbb{R}\times\mathbb{R}^d} \left|\Phi(a,\xi';V_s[r_s](x,x'))\eta_s(x,x')\right| \mathrm{d}\sigma_s(\xi',x') \\
&\quad + \int_{\mathbb{R}\times\mathbb{R}^d} \left|\Phi(a,\xi';V_s[r_t](x,x'))\eta_t(x,x')\right| \mathrm{d}\sigma_s(\xi',x') \\
&\leqslant L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} \Bigg(|a| \int_{\mathbb{R}^d} |V_s[r_s](x,x')| \mathrm{d}\mu(x') + \int_{\mathbb{R}^d} |V_s[r_s](x,x')| \int_{\mathbb{R}} |\xi'| \mathrm{d}\sigma_{s,x'}(\xi') \mathrm{d}\mu(x') \\
&\quad + |a| \int_{\mathbb{R}^d} |V_s[r_t](x,x')| \mathrm{d}\mu(x') + \int_{\mathbb{R}^d} |V_s[r_t](x,x')| \int_{\mathbb{R}} |\xi'| \mathrm{d}\sigma_{s,x'}(\xi') \mathrm{d}\mu(x') \Bigg) \\
&= L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} \Bigg(|a| \int_{\mathbb{R}^d} |V_s[r_s](x,x') \mathrm{d}\mu(x') + \int_{\mathbb{R}^d} |V_s[r_s](x,x')| m_1[\sigma_{s,x'}] \mathrm{d}\mu(x') \\
&\quad + |a| \int_{\mathbb{R}^d} |V_s[r_t](x,x')| \mathrm{d}\mu(x') + \int_{\mathbb{R}^d} |V_s[r_t](x,x')| m_1[\sigma_{s,x'}] \mathrm{d}\mu(x') \Bigg) \\
&\leqslant 2L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} C_V^{1/2} |a| + L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} C_V^{1/2} \Bigg(\left(\int_{\mathbb{R}^d} m_1^2[\sigma_{s,x'}] \mathrm{d}\mu(x') \right)^{1/2} \\
&\quad + \left(\int_{\mathbb{R}^d} m_1^2[\sigma_{s,x'}] \mathrm{d}\mu(x') \right)^{1/2} \Bigg) \\
&= 2L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} C_V^{1/2} |a| \\
&\quad + L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} C_V^{1/2} \Bigg(\left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} |\xi'| \mathrm{d}\sigma_{s,x'}(\xi') \right)^2 \mathrm{d}\mu(x') \right)^{1/2} \\
&\quad + \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} |\xi'| \mathrm{d}\sigma_{s,x'}(\xi') \right)^2 \mathrm{d}\mu(x') \right)^{1/2} \Bigg) \\
&\leqslant 2L_\Phi \|\eta\|_\infty C_V^{1/2} |a| + 2L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} C_V^{1/2} \left(\int_{\mathbb{R}^d} m_2[\sigma_{s,x'}] \mathrm{d}\mu(x') \right)^{1/2} \\
&= 2L_\Phi \|\eta\|_\infty C_V^{1/2} |a| + 2L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\infty})} C_V^{1/2} m_2^{1/2}[\gamma_s] < \infty \ ,
\end{aligned}
$$

where in the third inequality we used Cauchy-Schwarz and (3.1.9a), and the final inequality follows from Jensen's inequality. Then, since $t \mapsto V_t$ and $t \mapsto \eta_t$ are continuous, and using $\Phi_t - \Phi_s$ as a majorizing integrable function, by the Dominated Convergence Theorem we have that $II \to 0$ as $t \to s$. The limit $s \to t$ follows from an analogous argument. The remainder of the statement follows from classical Cauchy-Lipschitz theory for ODEs. □

The previous lemma guarantees we can consider, for μ-a.e. $x \in \mathbb{R}^d$, a flow map $f_x[\sigma, r, \eta]$, dependent on σ, r, η, which is the solution to

$$
\begin{aligned}
&\frac{d}{dt} f_x[\sigma,r,\eta](t,u) = \mathcal{X}[\sigma,r,\eta](t, f_x[\sigma,r,\eta](t,u), x) \ , \\
&f_x[\sigma,r,\eta](0,u) = u \in \mathbb{R} \ ,
\end{aligned}
\tag{3.2.8}
$$

In particular, for an initial time $s > 0$, we write $f_x[\sigma, r, \eta]((s,t),u)$, meaning that $f_x[\sigma, r, \eta](s,u) = u$.

Before establishing the well-posedness of the nonlinear problem (graph-CE) we will need to study some further properties of the flow.

Lemma 3.2.3. Let $V : [0,T] \times L^2_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ be continuous in time and such that it satisfies (3.1.9a) - (3.1.9b). Let $\sigma \in C([0,T], \mathcal{P}^\mu_2(\mathbb{R}\times\mathbb{R}^d))$ with $\pi^1{}_\#\sigma = \gamma$ and (r,η) be the solution of (Euler Co-NCL). Then, the flow map $f_x[\sigma, r, \eta]$ defined in (3.2.8) satisfies the properties below.

1. Linear growth: for all $T > 0$, there exists a constant $\widetilde{C}(T) = C(T, \eta_0, m_2[\gamma]) > 0$ such that

$$|f_x[\sigma, r, \eta](t,u)| \leqslant \widetilde{C}(T)(1 + |u|),$$

   for any $t \in [0,T]$, $u \in \mathbb{R}$, μ-a.e. $x \in \mathbb{R}^d$.

2. Lipschitz in u: there exists a constant $\bar{C}(T) = C(T, \eta_0) > 0$ such that for any $t \in [0,T]$ and any $u, \ell \in \mathbb{R}$

$$|f_x[\sigma, r, \eta](t,u) - f_x[\sigma, r, \eta](t,\ell)| \leqslant e^{\bar{C}(T)}|u - \ell| \ .$$

3. Time continuity: the map $t \mapsto f_x[\sigma, r, \eta](t, \cdot)$ is continuous.

Proof. We begin by proving 1. By definition of the flow we have

$$f_x[\sigma, r, \eta](t,u) = u + \int_0^t \mathcal{X}[\sigma, r, \eta](s, f_x[\sigma, r, \eta](s,u), x)\mathrm{d}s \ .$$

By means of (i), (ii) in Definition 2.2.3,

$$\begin{aligned}
&\left|\mathcal{X}[\sigma, r, \eta](s, f_x[\sigma, r, \eta](s,u), x)\right| \\
&\qquad \leqslant \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \left|\Phi(f_x[\sigma, r, \eta](s,u), \xi'; V_s[r_s](x,x'))\eta_s(x,x')\right| \mathrm{d}\sigma_s(\xi', x') \\
&\qquad \leqslant L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} C_V^{1/2} \left(|f_x[\sigma, r, \eta](s,u)| + \left(\int_{\mathbb{R}^d} m_1^2[\sigma_{s,x'}] \mathrm{d}\mu(x') \right)^{1/2} \right) \\
&\qquad \leqslant L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} C_V^{1/2} \left(|f_x[\sigma, r, \eta](s,u)| + m_2[\gamma_s]^{1/2} \right) \\
&\qquad \leqslant L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} C_V^{1/2} \left(|f_x[\sigma, r, \eta](s,u)| + m_2[\gamma_*]^{1/2} \right) \\
&\qquad \leqslant L_\Phi \left(\|\eta_0\|_\infty + C_\omega T \right) e^T C_V^{1/2} \left(|f_x[\sigma, r, \eta](s,u)| + m_2[\gamma_*]^{1/2} \right) \\
&\qquad \leqslant \widetilde{C}(T)(1 + |f_x[\sigma, r, \eta](s,u)|) \ ,
\end{aligned}$$

for a constant $\widetilde{C}(T) = C(T, \eta_0, m_2(\gamma))$, where we used (3.1.6) in the penultimate inequality. An application of Grönwall's inequality yields

$$|f_x[\sigma, r, \eta](t, u)| \leqslant e^{\widetilde{C}(T)}(1 + |u|) \ ,$$

where we denote $\widetilde{C}(T) = \widetilde{C}(T)T$ by slight abuse of notation. To prove 2., consider $u, \ell \in \mathbb{R}$ and use (3.2.6) from Lemma 3.2.2 in order to have

$$\begin{aligned}
|f_x[\sigma, r, \eta](t, u) - f_x[\sigma, r, \eta](t, \ell)| &\leqslant |u - \ell| + \int_0^t |\mathcal{X}[\sigma, r, \eta](s, f_x[\sigma, r, \eta](s, u), x) \\
&\qquad - \mathcal{X}[\sigma, r, \eta](s, f_x[\sigma, r, \eta](s, \ell), x)| \mathrm{d}s \\
&\leqslant |u - \ell| \\
&\qquad + L_\Phi C_V^{1/2} \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} \int_0^t |f_x[\sigma, r, \eta](s, \ell) - f_x[\sigma, r, \eta](s, u)| \mathrm{d}s \ .
\end{aligned}$$

An application of Grönwall's inequality yields

$$\begin{aligned}
|f_x[\sigma, r, \eta](t, u) - f_x[\sigma, r, \eta](t, \ell)| &\leqslant \exp\left(L_\Phi C_V^{1/2} \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} T\right) |u - \ell| \ , \\
&\leqslant \exp\left(L_\Phi C_V^{1/2} \left(\|\eta_0\|_\infty + C_\omega T\right) T e^T\right) |u - \ell| \\
&\leqslant e^{\bar{C}_T} |u - \ell|
\end{aligned}$$

for a constant $\bar{C}(T) = C(T, \eta_0) > 0$, which concludes the proof of (2). The final statement follows from the definition of the flow and the time continuity of the velocity field $\mathcal{X}[\sigma, r, \eta]$ proved in Lemma 3.2.2. □

The flow map f_x depends on a probability measure σ and a co-evolving graph $\mathcal{G}_t(\mu, \eta_t[r_t])$. To conclude our argument, we need to estimate the difference of the flow maps depending on different measures $\sigma^1,\ \sigma^2$. To this end, let us consider the space $C([0, T], \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d))$ with the following metric

$$\mathcal{D}_{\mu,d}(\nu^1, \nu^2) := \sup_{t \in [0,T]} L_\mu^2 d_2(\nu_t^1, \nu_t^2) = \sup_{t \in [0,T]} \left(\int_{\mathbb{R}^d} d_2^2(\nu_{t,x'}^1, \nu_{t,x'}^2) \mathrm{d}\mu(x') \right)^{1/2} ,$$

for $\nu^1, \nu^2 \in C([0, T], \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d))$, so that $(C([0, T], \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d)), \mathcal{D}_{\mu,d})$ is a complete metric space, which follows from analogous arguments to Proposition A.0.1 in Appendix A. We obtain the following estimate.

Lemma 3.2.4. Let $\sigma^i \in C([0, T], \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d))$, with $\pi^1{}_\# \sigma^i = \gamma^i$ and (r, η) be the solution of (Euler Co-NCL) with initial condition (r_0, η_0). Let $V : [0, T] \times L_\mu^2(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ be continuous in time satisfying (3.1.9a) and (3.1.9b). Consider two flows $f_x^i[\sigma^i, r, \eta]$, for μ-a.e.

$x \in \mathbb{R}^d$, solutions to (3.2.8), for $i = 1, 2$. Then, it holds

$$|f_x^1[\sigma^1, r, \eta](t, u) - f_x^2[\sigma^2, r, \eta](t, u)|^2 \leqslant C(T) T^2 \mathcal{D}_{\mu,d}^2(\sigma^1, \sigma^2) \ , \tag{3.2.9}$$

where $C(T) = C(T, \eta_0) > 0$. Consequently, for any measure $\nu \in \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d)$, with disintegration $\nu = \int_{\mathbb{R}^d} \nu_x \mathrm{d}\mu(x)$, we have

$$\mathcal{D}_{\mu,d}(f^1[\sigma^1, r, \eta]_\# \nu, f^2[\sigma^2, r, \eta]_\# \nu) \leqslant C(T) T \mathcal{D}_{\mu,d}(\sigma_1, \sigma_2) \ . \tag{3.2.10}$$

Proof. To improve the readability we use the shorthand notation $f_x^i := f_x^i[\sigma^i, r, \eta]$, for $i = 1, 2$. Let us start by estimating the difference of the two flows:

$$\begin{aligned}
|f_x^1(t,u) - f_x^2(t,u)| =& \Big| \int_0^t \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(f_x^1(s,u), \xi'; V_s[r_s](x,x'))\eta_s(x,x')\mathrm{d}\sigma_s^1(\xi',x') \\
&\qquad - \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(f_x^2(s,u), \xi'; V_s[r_s](x,x'))\eta_s(x,x')\mathrm{d}\sigma_s^2(\xi',x')\mathrm{d}s \Big| \\
\leqslant& \int_0^t \Big| \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Big(\Phi(f_x^1(s,u), \xi'; V_s[r_s](x,x'))\eta_s(x,x') \\
&\qquad - \Phi(f_x^2(s,u), \xi'; V_s[r_s](x,x'))\eta_s(x,x') \Big) \mathrm{d}\sigma_s^1(\xi',x') \Big| \mathrm{d}s \\
&+ \int_0^t \Big| \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Phi(f_x^2(s,u), \xi'; V_s[r_s](x,x'))\eta_s(x,x')\mathrm{d}(\sigma_s^1 - \sigma_s^2)(\xi',x') \Big| \mathrm{d}s \\
=:& I + II.
\end{aligned}$$

We bound the term I using analogous calculations to the ones for (3.1.10) to obtain:

$$\begin{aligned}
|I| =& \int_0^t \Big| \int_{\mathbb{R}\times\mathbb{R}^d\setminus\{x\}} \Big(\Phi(f_x^1(s,u), \xi'; V_s[r_s](x,x')) \\
&\qquad - \Phi(f_x^2(s,u), \xi'; V_s[r_s](x,x')) \Big) \eta_s(x,x')\mathrm{d}\sigma_s^1(\xi',x') \Big| \mathrm{d}s \\
\leqslant& L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}\diagup)} \int_0^t \int_{\mathbb{R}^d} |V_s[r_s](x,x')| |f_x^1(s,u) - f_x^2(s,u)| \mathrm{d}\mu(x') \\
\leqslant& L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}\diagup)} C_V^{1/2} \int_0^t |f_x^1(s,u) - f_x^2(s,u)| \mathrm{d}s \ .
\end{aligned}$$

Whence

$$|I|^2 \leqslant L_\Phi^2 \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}\diagup)}^2 C_V \ T \int_0^t |f_x^1(s,u) - f_x^2(s,u)|^2 \mathrm{d}s \ .$$

On the other hand, for II, we have

$$
\begin{aligned}
|II| &\leqslant \int_0^t \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}} \Phi(f_x^2(s,u),\xi';V_s[r_s](x,x'))\eta_s(x,x')\mathrm{d}(\sigma^1_{s,x'}-\sigma^2_{s,x'})(\xi') \right| \mathrm{d}\mu(x')\mathrm{d}s \\
&\leqslant \int_0^t \int_{\mathbb{R}^d} Lip[\Phi(f_x^2(s,u),\cdot,V_s[r_s](x,x')\eta_s(x,x')]d_1(\sigma^1_{s,x'},\sigma^2_{s,x'})\mathrm{d}\mu(x')\mathrm{d}s \\
&\leqslant L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} \int_0^t \int_{\mathbb{R}^d} |V_s[r_s](x,x')| d_1(\sigma^1_{s,x'},\sigma^2_{s,x'})\mathrm{d}\mu(x')\mathrm{d}s \\
&\leqslant L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V^{1/2} \int_0^t \left(\int_{\mathbb{R}^d} d_1^2(\sigma^1_{s,x'},\sigma^2_{s,x'})\mathrm{d}\mu(x') \right)^{1/2} \mathrm{d}s \\
&\leqslant L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V^{1/2} \int_0^t L_\mu^2 d_2(\sigma_s^1,\sigma_s^2)\mathrm{d}s \\
&\leqslant L_\Phi \|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V^{1/2} T \mathcal{D}_{\mu,d}(\sigma^1,\sigma^2) \ ,
\end{aligned}
\tag{3.2.11}
$$

where the second and third inequality follow from the definition by duality of the 1-Wasserstein distance and (3.2.7), the fourth inequality follows from Cauchy-Schwarz. Squaring both sides and taking the supremum over time yields

$$
|II|^2 \leqslant L_\Phi^2 \|\eta\|^2_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V \mathcal{D}^2_{\mu,d}(\sigma^1,\sigma^2) T^2 \ .
$$

Thus, collecting all the terms we get that

$$
\begin{aligned}
|f_x^1(t,u)-f_x^2(t,u)|^2 \leqslant\ & L_\Phi^2 \|\eta\|^2_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V \ T \int_0^t |f_x^1(s,u)-f_x^2(s,u)|^2 \mathrm{d}s \\
& + L_\Phi^2 \|\eta\|^2_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V \mathcal{D}^2_{\mu,d}(\sigma^1,\sigma^2) T^2,
\end{aligned}
$$

and an application of Grönwall's inequality yields

$$
\begin{aligned}
|f_x^1(t,u)-f_x^2(t,u)|^2 &\leqslant \exp\{L_\Phi^2 \|\eta\|^2_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V \ T^2\} \\
&\quad \times \|\eta\|^2_{\infty,C_b(\mathbb{R}^{2d}_{\not\angle})} C_V \mathcal{D}^2_{\mu,d}(\sigma^1,\sigma^2) L_\Phi^2 T^2 \\
&\leqslant \exp\{L_\Phi^2 (\|\eta_0\|_\infty + C_\omega T)^2 e^{2T} C_V T^2\} \\
&\quad \times L_\Phi^2 (\|\eta_0\|_\infty + C_\omega T)^2 e^{2T} C_V \mathcal{D}^2_{\mu,d}(\sigma^1,\sigma^2) T^2 \\
&= C(T) \mathcal{D}^2_{\mu,d}(\sigma^1,\sigma^2) T^2 \ .
\end{aligned}
\tag{3.2.12}
$$

The first statement is therefore proven. To obtain the second estimate, consider the transference plan $\Pi_{t,x} = (f_x^1 \times f_x^2)_\# \nu_x$ and denote by $\kappa := \pi^1_\# \nu$ the first marginal of ν. For μ-a.e. $x \in \mathbb{R}^d$ we

have that

$$\begin{aligned} d_2^2(f^1_{x\,\#}\nu_x, f^2_{x\,\#}\nu_x) &\leqslant \int_{\mathbb{R}\times\mathbb{R}} |a-b|^2 \mathrm{d}\Pi_{t,x}(a,b) \\ &\leqslant \int_{\mathbb{R}} |f^1_x(t,u) - f^2_x(t,u)|^2 \mathrm{d}\nu_x(u) \\ &\leqslant C(T) T^2 \mathcal{D}^2_{\mu,d}(\sigma^1,\sigma^2), \end{aligned}$$

where the final inequality follows from (3.2.12). Integrating with respect to μ, taking the square root on both sides and the supremum in time gives

$$\mathcal{D}_{\mu,d}(f^1_{\#}\nu, f^2_{\#}\nu) \leqslant C(T) T \mathcal{D}_{\mu,d}(\sigma^1,\sigma^2)\ .$$

□

In order to prove well-posedness of (graph-CE) through its disintegrated version, we first need to study well-posedness for solutions of a linear version of this problem, namely

$$\begin{aligned} &\partial_t \tilde{\sigma}_{t,x} + \partial_\xi(\tilde{\sigma}_{t,x}\mathcal{X}[\sigma,\rho,\eta]) = 0, \\ &\tilde{\sigma}_{0,x} = \bar{\sigma}_x \in \mathcal{P}(\mathbb{R}), \end{aligned} \tag{3.2.13}$$

for μ-almost every $x \in \mathbb{R}^d$ and $\bar{\sigma} = \int_{\mathbb{R}^d} \bar{\sigma}_x \mathrm{d}\mu(x)$. We prove existence and uniqueness of solutions to (3.2.13) by duality, that is, we first consider classical solutions to the transport equation, μ-a.e. $x \in \mathbb{R}^d$:

$$\begin{cases} \partial_t \varphi + \partial_\xi \varphi \mathcal{X}[\sigma,r,\eta](t,\xi,x) = 0, \qquad t<T,\ (\xi,x) \in \mathbb{R}\times\mathbb{R}^d \\ \varphi(T,\xi,x) = \psi(\xi,x) \in C^1_c(\mathbb{R}\times\mathbb{R}^d). \end{cases} \tag{3.2.14}$$

The method of characteristics provides the following result; see [Eva10, Chapter 3] for further details.

Proposition 3.2.1. Let $\sigma \in C([0,T], \mathcal{P}^\mu_2(\mathbb{R}\times\mathbb{R}^d))$ be given. The problem (3.2.14) has a unique classical solution $\varphi(t,\xi,x) = \psi(f_x[\sigma,r,\eta]((t,T),u)) \in C^1_c([0,T]\times\mathbb{R})$, for μ-a.e. $x \in \mathbb{R}^d$, where the flow f_x is as in (3.2.8).

Therefore, for a given $\sigma \in C([0,T]; \mathcal{P}^\mu_2(\mathbb{R}\times\mathbb{R}^d))$, we can construct explicit distributional solutions for (3.2.13) as in the following theorem.

Proposition 3.2.2. Let $\tilde{\sigma}_0 \in \mathcal{P}^\mu_2(\mathbb{R}\times\mathbb{R}^d)$ such that $\tilde{\sigma}_0 = \int_{\mathbb{R}^d} \tilde{\sigma}_{0,x'} \mathrm{d}\mu(x')$ and $\sigma \in C([0,T], \mathcal{P}^\mu_2(\mathbb{R}\times\mathbb{R}^d))$. Assume $V : [0,T]\times L^2_\mu(\mathbb{R}^d)\times\mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ satisfies (3.1.9a), (3.1.9b) and that it is continuous in time. Let (r,η) be a solution to (Euler Co-NCL). Then, $\tilde{\sigma}_{t,x} = f_x[\sigma,r,\eta](t,x,\cdot)_\#\tilde{\sigma}_{0,x}$ is the unique distributional solution to (3.2.13), for μ-a.e. $x \in \mathbb{R}^d$, where f_x is defined in (3.2.8).

Proof. First, we notice that by direct computation it can be checked that $\tilde{\sigma}_{t,x} = f_x[\sigma,\rho,\eta](t,\cdot)_{\#}\tilde{\sigma}_{0,x}$ satisfies the equation in the sense of distributions, i.e., (3.2.5), for μ-a.e. $x \in \mathbb{R}^d$. Secondly, by the linearity of (3.2.13), we can show uniqueness for solutions with $\tilde{\sigma}_0 = 0$. By choosing $\varphi(t,\xi,x) = \psi(f_x[\sigma,\rho,\eta]((t,T),\xi))$ as a test function in (3.2.5), we obtain

$$\int_{\mathbb{R}} \varphi(T,\xi,x)\mathrm{d}\tilde{\sigma}_{T,x}(\xi) = \int_{\mathbb{R}} \psi(\xi,x)\mathrm{d}\tilde{\sigma}_{T,x}(\xi) = 0 \ , \ \mu\text{-a.e. } x \in \mathbb{R}^d,$$

for all $\psi \in C_c^1(\mathbb{R})$, whence $\tilde{\sigma}_{T,x} = 0$ and we obtain uniqueness. □

We are ready to prove the well-posedness of (graph-CE). For any $\nu \in \mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d)$ with disintegration $\nu = \int_{\mathbb{R}^d} \nu_x \mathrm{d}\mu(x)$, $x \in \operatorname{supp}\mu$ and the solution (r,η) of (Euler Co-NCL), we define the solution map $\mathcal{S}^{\nu,r,\eta} : C([0,T],\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d)) \to C([0,T],\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d))$ as

$$\mathcal{S}^{\nu,r,\eta}[\sigma](t) := \int_{\mathbb{R}^d} f_x[\sigma,r,\eta](t,\cdot)_{\#}\nu_x \mathrm{d}\mu(x) \ , \quad \sigma \in \mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d), \tag{3.2.15}$$

where $f_x[\sigma,r,\eta]$ is defined in (3.2.8). Note that the solution map is well defined due to the properties of the flow proved in Lemma 3.2.3. We denote the disintegration with respect to μ by $\mathcal{S}_x^{\nu,r,\eta}[\sigma](t) := f_x[\sigma,r,\eta](t,\cdot)_{\#}\nu_x$, for μ-a.e. $x \in \mathbb{R}^d$. We prove well-posedness of the nonlinear problem (graph-CE) by another fixed-point argument.

Theorem 3.2.1. Let $V : [0,T]\times L^2_{\mu}(\mathbb{R}^d)\times\mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ be continuous in time and satisfy (3.1.9a) and (3.1.9b). Let (r,η) be the solution of (Euler Co-NCL). Then, there exists a unique solution $\sigma \in C([0,T],\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d))$ to (graph-CE) with initial condition $\nu \in \mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d)$ given by $\sigma = \int_{\mathbb{R}^d}\sigma_{t,x}\mathrm{d}\mu(x)$ where $\sigma_{t,x} = f_x[\sigma,r,\eta](t,\cdot)_{\#}\nu_x$ for μ-a.e. $x \in \mathbb{R}^d$.

Proof. Let us consider the flow $f_x[\sigma,\rho,\eta]$ as defined in (3.2.8), for some $\sigma \in C([0,T],\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d))$. Note that by Proposition 3.2.2, we have that $\tilde{\sigma}_x = \mathcal{S}_x^{\nu,r,\eta}[\sigma]$ is the unique solution in $C([0,T],\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d))$ to the linear problem (3.2.13) with initial data $\nu \in \mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d)$. Thus, to obtain existence and uniqueness of solution to the nonlinear problem (graph-CE) it suffices to show the solution map $\mathcal{S}^{\nu,r,\eta}[\sigma]$ has a unique fixed-point. Using the disintegrated version of the solution map, $\mathcal{S}_x^{\nu,r,\eta}[\sigma](t)$, and Lemma 3.2.4 we have, for any $\sigma^1,\sigma^2 \in C([0,T],\mathcal{P}_2^{\mu}(\mathbb{R}\times\mathbb{R}^d))$,

$$\mathcal{D}_{\mu,d}(\mathcal{S}^{\nu,r,\eta}[\sigma^1],\mathcal{S}^{\nu,r,\eta}[\sigma^2]) \leqslant C(T)T\mathcal{D}_{\mu,d}(\sigma^1,\sigma^2),$$

where $C(T) > 0$ is the constant from (3.2.10) and its leading terms are of the form $C(T) = \exp\{\alpha T^4 e^T + \gamma T^2 e^T + \beta\}$, for $\alpha,\gamma,\beta > 0$, since by tracing back the constant in (3.2.9) we see that it is given by $\exp\{L_{\Phi}^2(\|\eta_0\|_{\infty} + C_{\omega}T)^2 e^{2T} C_V T^2\} L_{\Phi}^2(\|\eta_0\|_{\infty} + C_{\omega}T)^2 e^{2T} C_V$. Then, for $T < C(T)^{-1}$, the solution map is a contraction. By the Banach fixed-point Theorem we can conclude there

exists a unique solution to (graph-CE) in $C([0,T],\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d))$ for $T<T^*$, where T^* is such that $C_T T^*=1$. The solution can be extended using a standard iteration procedure so that the time interval does not depend on any T^*, see e.g. Theorem 2.3.1. □

Theorem 3.2.2. Assume $V:[0,T]\times L^2_\mu(\mathbb{R}^d)\times\mathbb{R}^{2d}_{\diagup}\to\mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ to be continuous in time and to satisfy (3.1.9a) and (3.1.9b). Let (r,η) be the solution to (Euler Co-NCL) with initial data $(r_0,\eta_0)\in L^2_\mu(\mathbb{R}^d)\times C_b(\mathbb{R}^{2d}_{\diagup})$ and let $\sigma^i\in C([0,T],\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d))$ be the solutions to the corresponding (graph-CE) with initial conditions $\tilde\sigma_0^i\in\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d)$ for $i=1,2$. Then we have the following stability estimate

$$L^2_\mu d_2^2(\sigma_t^1,\sigma_t^2)\leqslant 2e^{\tilde C_T+2C(T)T^2}L^2_\mu d_2^2(\tilde\sigma_0^1,\tilde\sigma_0^2)\ ,$$

where $\tilde C_T>0$ is the constant obtained in 3.2.3 2. and $C(T)$ is the constant in (3.2.12), both depending on η_0.

Proof. Let us write $f_x^i=f_x^i[\sigma^i,r,\eta]$ for $i=1,2$. Then we have that for μ-a.e. $x\in\mathbb{R}^d$ that

$$\begin{aligned}d_2^2(\sigma_{t,x}^1,\sigma_{t,x}^2)&=d_2^2(f^1[\sigma^1,\eta,r](t,x)_\#\tilde\sigma_0^1,f^2[\sigma^2,\eta,r](t,x)_\#\tilde\sigma_0^2)\\&\leqslant 2(d_2^2(f_x^1(t)_\#\tilde\sigma_0^1,f_x^2(t)_\#\tilde\sigma_0^1)+d_2^2(f_x^2(t)_\#\tilde\sigma_0^1,f_x^2(t)_\#\tilde\sigma_0^2))\\&\leqslant 2C(T)T\int_0^t L^2_\mu d_2^2(\sigma_s^1,\sigma_s^2)\mathrm{d}s+2d_2^2(f_x^2(t)_\#\tilde\sigma_0^1,f_x^2(t)_\#\tilde\sigma_0^2)\ ,\end{aligned}$$

where $C(T)$ is the constant obtained in (3.2.12), depending on η_0, where we used analogous calculations to the ones in the proof of the first statement in Lemma 3.2.4 to get the final inequality. More precisely, we do not take the supremum in time in (3.2.11) and then, in (3.2.12), we obtain

$$|f_x^1(t,u)-f_x^2(t,u)|^2\leqslant C(T)\int_0^t L^2_\mu d_2^2(\sigma_s^1,\sigma_s^2)\mathrm{d}s\ ,$$

after an application of Jensen's inequality. To bound the second term on the right-hand side we proceed as follows. Let $\Pi_{o,x}\in\Gamma(\tilde\sigma_{0,x}^1,\tilde\sigma_{0,x}^2)$ be an optimal transference plan for $\tilde\sigma_{0,x}^1$ and $\tilde\sigma_{0,x}^2$ in d_2. Then for μ-a.e. $x\in\mathbb{R}^d$ we have that

$$\begin{aligned}d_2^2(f_x^2(t)_\#\tilde\sigma_0^1,f_x^2(t)_\#\tilde\sigma_0^2)&\leqslant\int_{\mathbb{R}\times\mathbb{R}^d}\int_{\mathbb{R}\times\mathbb{R}^d}|f_x^2(t,u)-f_x^2(t,v)|^2\mathrm{d}\Pi_{o,x}(u,v)\\&\leqslant e^{\bar C_T}\int_{\mathbb{R}\times\mathbb{R}^d}\int_{\mathbb{R}\times\mathbb{R}^d}|u-v|^2\mathrm{d}\Pi_{o,x}(u,v)\\&\leqslant e^{\bar C_T}d_2^2(\tilde\sigma_{0,x}^1,\tilde\sigma_{0,x}^2)\ ,\end{aligned}$$

where the second inequality follows from Lemma 3.2.3 2. and $\bar C(T)=C(T,\eta_0)>0$. Collecting

terms, we get

$$d_2^2(f_x^2(t)_\#\tilde{\sigma}_0^1, f_x^2(t)_\#\tilde{\sigma}_0^2) \leqslant 2C(T)T\int_0^t L_\mu^2 d_2^2(\sigma_s^1,\sigma_s^2)\mathrm{d}s + 2e^{\bar{C}_T} d_2^2(\tilde{\sigma}_{0,x}^1, \tilde{\sigma}_{0,x}^2)$$

Integrating with respect to μ and an application of Grönwall's inequality yields the desired result. □

3.2.1 Comments on the asymptotic behaviour and discrete-to-continuum limit of (Euler Co-NCL)

The findings presented in this chapter were obtained to try to leverage results on the long-time behaviour of the continuity equation on $\mathbb{R}^d$ and apply them to our setting. We attempted to obtain contractivity of solutions to the graph continuity equation in the 2-Wasserstein distance. The only result in this direction we managed to obtain required variants of the following hypothesis

$$\inf_{t\in[0,T]}\inf_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d\setminus\{x\}} V_t(x,x')\eta_t(x,x')\mathrm{d}\mu(x') > 0\ . \tag{3.2.16}$$

We invite the reader to look at Proposition A.0.2 in Appendix A to see how (3.2.16) is obtained. Even for fully connected graphs, i.e. with $\eta > 0$ for all $t \geqslant 0$, this condition is not reasonable since it can only be satisfied if the velocity is positive for all vertices for any time. This implies that all vertices must "give mass" to, at least, one neighbour. However, the author could not find any example that fulfilled this condition and that was of the form $V(x,x') = -\overline{\nabla}\phi(x,x') = -(\phi(x')-\phi(x))$ for some suitable, potentially solution dependent, function ϕ. This motivated the alternative approach presented in Chapter 4 to obtain the long-time behaviour of (Euler Co-NCL).

Furthermore, as we showed in Theorem 2.5.1 in Chapter 2, in our framework, we obtain a discrete-to-continuum limit by considering a sequence of atomic base measures $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^+(\mathbb{R}^d)$ that converges to a limiting base measure $\mu \in \mathcal{M}^+(\mathbb{R}^d)$. However, given that the space $\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d)$ has a fixed second marginal μ on $\mathbb{R}^d$, we cannot consider a sequence of solutions to (graph-CE) depending on each element of the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ defined on the same space. A natural way to circumvent this issue would be to consider the full space $\mathcal{P}_2(\mathbb{R}\times\mathbb{R}^d)$ equiped with the 2-Wasserstein distance, but this would require us to assume Lipschitzness of the maps $(x,x') \mapsto V_\cdot[\cdot](x,x')$ and $(x,x') \mapsto \eta_\cdot(x,x')$, the latter being very restrictive particularly if one wants to consider sparse graphs. In other approaches in the literature, this is not an issue as the discrete-to-continuum limit is obtained by considering sequences of the analog to η in our work.

Chapter 4

Long-time behaviour

4.1 Introduction

In this section we obtain a result on the long-time behaviour for solutions to (Euler Co-NCL) in the case of a specific class of velocities. We highlight two main features. First, the velocities in this class are *pointwise*, meaning that $V_t[r](x, x') = V_t(r_t(x), r_t(x'))$ for a solution (r, η) to (Euler Co-NCL). Second, they are *monotonic*, which intuitively means that the direction of the velocity is given by the "negative mass gradient" at two vertices $(x, x') \in \mathbb{R}^{2d}_{\diagup}$, see Definition 4.2.1 for a precise definition.

As we anticipated in Chapter 1, we begin by showing well-posedness of (Euler Co-NCL) in $L^\infty_\mu(\mathbb{R}^d)$ using the Banach fixed-point theorem. Afterward, we study the long-time behaviour of the solution to (Euler Co-NCL) and, for this, we shall constrain the co-evolving graph to a compact set $K \subset \mathbb{R}^d$ and fix the interpolation to be Φ_{upwind}. Then, for initial data satisfying $\int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) = M$, we obtain that, as $t \to \infty$, the solution of (Euler Co-NCL) converges to $\frac{M}{\mu(K)}$, the uniform mass distribution on the set K. This result is natural considering the monotonicity of the velocities. We prove this result by showing that the supremum norm of r decreases monotonically in time, while the infimum of r increases monotonically in time until the dynamics stabilise at $\frac{M}{\mu(K)}$. The rest of this chapter proceeds as follows. In Section 4.2 we show the well-posedness of (Euler Co-NCL) for $r \in L^\infty_\mu(\mathbb{R}^d)$. Then, in Section 4.3 we study its long-time behaviour.

4.2 L^∞ theory for (Euler Co-NCL)

Similarly to the previous chapter, we consider the problem

$$\begin{aligned}\partial_t r_t(x) &= -\int_{\mathbb{R}^d\setminus\{x\}} \Phi(r_t(x), r_t(y); V_t[r](x,y))\eta_t(x,y)\mathrm{d}\mu(y) \ , \\ \partial_t \eta_t &= \omega_t[r] - \eta_t \ ,\end{aligned} \tag{Euler Co-NCL}$$

for given initial data $r^0 \in L^\infty_\mu(\mathbb{R}^d)$ and $\eta^0 \in C_b(\mathbb{R}^{2d}_{\not\sim})$. Note that, unlike in Chapter 3, the initial datum r^0 belongs to $L^\infty_\mu(\mathbb{R}^d)$ as we seek to prove well-posedness of this problem in this space.

We consider velocities $V : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\not\sim} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\not\sim})$ that depend on r pointwise, i.e., for $r \in L^\infty_\mu(\mathbb{R}^d)$,

$$V_t[r](x,x') = V_t(r(x), r(x')).$$

We refer to such velocities as *pointwise* and we recall the definition below.

Definition 4.2.1. We say that a pointwise velocity $V : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\not\sim} \to \mathcal{V}^{as}(\mathbb{R}^d)$ is *monotonic* if, for $r \in L^\infty_\mu(\mathbb{R}^d)$, any $t \geqslant 0$, and μ-a.e. $x, x' \in \mathbb{R}^d$ we have that:

- $r(x) > r(y)$, then $V_t^+(r(x), r(x')) > V_t^+(r(y), r(x'))$,
- $r(x) < r(y)$, then $V_t^-(r(x), r(x')) > V_t^-(r(y), r(x'))$,
- $r(x) = r(y)$, then $V_t(r(x), r(x')) = V_t(r(y), r(x'))$,
- $r(x) = r(x')$, then $V_t(r(x), r(x')) = 0$.

Let us explain the meaning of the conditions in the previous definition. The first condition can be interpreted as saying that, if the mass is greater at a vertex $x \in \mathbb{R}^d$ than $y \in \mathbb{R}^d$, then the outflow from vertex x to its neighbours is greater than the outflow from y. Conversely, if the mass at vertex x is lower than the mass at vertex y, the inflow to x from its neighbours should be greater than the inflow to y. The rest of the conditions follow the same type of reasoning. This type of velocity structure is natural for a system with fixed positive mass. Indeed, this is what we obtain in (Euler Co-NCL) if $\Phi \equiv \Phi_{upwind}$ and η_t is symmetric and positive for $t \geqslant 0$, which follows from [EPS23, Proposition 5.2] and we shall stress this later. Monotonic velocities ensure that the mass on the graph diffuses as we shall show in Section 4.3. Our guiding example for such velocities is given by

$$V_t(r_t(x), r_t(x')) = -\overline{\nabla}\alpha(r_t)(x,x') = \alpha(r_t(x)) - \alpha(r_t(x')) \ , \tag{4.2.1}$$

where $\alpha : \mathbb{R} \to \mathbb{R}$ is monotonically increasing and bounded such as $\alpha(x) = \frac{1}{1+e^{-x}}$ for $x \in \mathbb{R}^d$. We remark that this type of velocity is different from the ones considered in Chapter

2 and Chapter 3 which came from interaction energies. Focusing on the dynamics for η, $\omega : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ will now satisfy the following version of assumption ($\boldsymbol{\omega}1$)-($\tilde{\boldsymbol{\omega}}4$).

($\boldsymbol{\omega}1$) the map $(x,y) \in \mathbb{R}^{2d}_{\diagup} \mapsto \omega_t[\cdot](\cdot,x,y)$ is continuous and $(t \mapsto \omega_t[\cdot](\cdot,\cdot)) \in L^1([0,T])$;

($\boldsymbol{\omega}2$) there exists a constant $L_\omega \geqslant 0$ such that for any $f, g \in L^\infty_\mu(\mathbb{R}^d)$

$$\sup_{t\in[0,T]} \sup_{(x,y)\in\mathbb{R}^{2d}_{\diagup}} |\omega_t[f](x,y) - \omega_t[g](x,y)| \leqslant L_\omega \|f-g\|_{L^\infty_\mu(\mathbb{R}^d)};$$

($\boldsymbol{\omega}3$) ω is bounded, that is, there exists a constant $\tilde{C}_{\omega,T} > 0$ such that

$$\sup_{t\in[0,T]} \sup_{r\in L^\infty_\mu(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}_{\diagup}} \left|\omega_t[r](x,y)\right| \leqslant \tilde{C}_{\omega,T}.$$

($\boldsymbol{\omega}4$) The map $(x,y) \mapsto \omega_\cdot[\cdot](x,y)$ is symmetric.

Note that we have specified the dependence of the constant $\tilde{C}_{\omega,T}$ on the length of the time interval $[0,T]$. This is because we will show well-posedness of (Euler Co-NCL) for any $T > 0$ and, in Section 4.3, we will consider solutions of the problem over the interval $[0,\infty)$. Furthermore, although assumption ($\boldsymbol{\omega}4$) is not needed to obtain the well-posedness, it is important both to obtain the form (3.1.1) for the divergence of the flux as well as to be able to apply the result in [EPS23, Proposition 5.2]. Our guiding example for the function ω is as in (3.1.3). Solutions of (Euler Co-NCL) are defined as follows.

Definition 4.2.2 (Solution to (Euler Co-NCL))**.** Given a jointly antisymmetric admissible flux interpolation Φ, a pointwise velocity field $V : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$, and a function $\omega : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$, a pair $(r,\eta) : [0,T] \to L^\infty_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ is a solution (Euler Co-NCL) with initial datum $(r_0,\eta_0) \in L^\infty_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ if it satisfies the following properties:

1. for μ-a.e. $x \in \mathbb{R}^d$ the maps $t \mapsto \int_{\mathbb{R}^d\setminus\{x\}} \Phi(r_t(x), r_t(y); V_t[r](x,y))\eta_t(x,y)\mathrm{d}\mu(y)$ and $t \mapsto \omega_t[\rho_t] - \eta_t$ belong to $L^1([0,T])$;

2. for μ-a.e. $x \in \mathbb{R}^d$, $t \in [0,T]$, $(x,y) \in \mathbb{R}^{2d}_{\diagup}$, it holds

$$r_t(x) = r_0(x) - \int_0^t \int_{\mathbb{R}^d\setminus\{x\}} \Phi(r_s(x), r_s(y); V_s[r](x,y))\eta_s(x,y)\mathrm{d}\mu(y)\mathrm{d}s\ , \tag{4.2.2a}$$

$$\eta_t(x,y) = \eta_0(x,y) + \int_0^t (\omega_s[r](x,y) - \eta_s(x,y))\,\mathrm{d}s\ . \tag{4.2.2b}$$

3. $r \in AC([0,T], L^\infty_\mu(\mathbb{R}^d))$, $\eta \in AC([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$.

We obtain well-posedness of (Euler Co-NCL) by exploiting the Banach fixed-point theorem in the space $C([0,T], L^\infty_\mu(\mathbb{R}^d)) \times C([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$, equipped with the distance

$$\tilde{d}_\infty((f^1,g^1),(f^2,g^2)) = \|f^1 - f^2\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} + \|g^1 - g^2\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})},$$

where

$$\|f^1 - f^2\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} := \sup_{t\in[0,T]} \|f^1_t - f^2_t\|_{L^\infty_\mu(\mathbb{R}^d)} \ ,$$
$$\|g^1 - g^2\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} := \sup_{t\in[0,T]} \|g^1_t - g^2_t\|_{C_b(\mathbb{R}^{2d}_{\diagup})} \ .$$

As in the previous chapter, we prove *a priori* properties for (Euler Co-NCL) to define the solution maps later on.

Proposition 4.2.1. Let Φ be a jointly antisymmetric admissible flux interpolation, $r_0 \in L^\infty_\mu(\mathbb{R}^d)$, $\eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})$. Assume $\omega : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ is such that the map $(x,y) \in \mathbb{R}^{2d}_{\diagup} \mapsto \omega_t[\cdot](x,y)$ is continuous and that, for a constant $C_{\omega,T} > 0$,

$$\int_0^T \sup_{r\in L^\infty_\mu(\mathbb{R}^d)} \sup_{(x,y)\in\mathbb{R}^{2d}_{\diagup}} |\omega_s[r](x,y)| \mathrm{d}s \leqslant C_{\omega,T} \ . \tag{4.2.3}$$

Suppose that $V : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{\mathrm{as}}(\mathbb{R}^{2d}_{\diagup})$ is pointwise and satisfies

$$\sup_{t\in[0,T]} \sup_{r\in L^\infty_\mu(\mathbb{R}^d)} \sup_{(x,x')\in\mathbb{R}^{2d}_{\diagup}} |V_t(r_t(x), r_t(x'))| \leqslant C_{V,T} \ , \tag{4.2.4}$$

for some $C_{V,T} > 0$. For a pair $(r,\eta) : [0,T] \to L^\infty_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ satisfying (4.2.2a) and (4.2.2b), respectively, the following properties hold.

1. For μ-a.e. $x \in \mathbb{R}^d$ and every $(x,y) \in \mathbb{R}^{2d}_{\diagup}$, the maps

$$t \mapsto \int_{\mathbb{R}^d\setminus\{x\}} \Phi(r_t(x), r_t(y); V_s(r_s(x), r_s(x')))\eta_s(x,x)\mathrm{d}\mu(y) \in L^1([0,T]) \ ,$$
$$t \mapsto \omega_t[r](x,y) - \eta_t(x,y) \in L^1([0,T]) \ ,$$

   and $r \in L^\infty([0,T]; L^\infty_\mu(\mathbb{R}^d)), \eta \in L^\infty([0,T], C_b(\mathbb{R}^{2d}_{\diagup}))$.

2. $r \in AC([0,T]; L^\infty_\mu(\mathbb{R}^d))$ and $\eta \in AC([0,T]; C_b(\mathbb{R}^{2d}_{\diagup}))$.

Proof. We obtain $\eta \in L^\infty([0,T]; C_b(\mathbb{R}^{2d}_{\diagup}))$ in an analogous way to Proposition 3.1.1, with

$$\|\eta_t\|_\infty \leqslant (\|\eta_0\|_\infty + C_{\omega,T})\, e^T \quad , \tag{4.2.5}$$

and this implies the integrability of the map $t \mapsto \omega_t[r](x,y) - \eta_t(x,y)$ for $(x,y) \in \mathbb{R}^{2d}_{\diagup}$ as well as $\eta \in AC([0,T]; C_b(\mathbb{R}^{2d}_{\diagup}))$. As for r, for μ-a.e. $x \in \mathbb{R}^d$, by using (i), (ii) in Definition 2.2.3, boundedness of η, and (4.2.4) we have

$$\begin{aligned}
|r_t(x)| &\leqslant |r_0(x)| + \int_0^t \int_{\mathbb{R}^d \setminus \{x\}} |\Phi(r_s(x), r_s(y); V_s(r_s(x), r_s(y)))\eta_s(x,y)| \, \mathrm{d}\mu(y)\mathrm{d}s \\
&\leqslant |r_0(x)| + L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} \int_0^t \int_{\mathbb{R}^d \setminus \{x\}} |V_s(r_s(x), r_s(y))| \, (|r_s(x)| + |r_s(y)|) \mathrm{d}\mu(y)\mathrm{d}s \\
&\leqslant \|r_0\|_{L^\infty_\mu(\mathbb{R}^d)} + 2L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} C_{V,T} \int_0^t \|r_s\|_{L^\infty_\mu(\mathbb{R}^d)} \mathrm{d}s \, .
\end{aligned}$$

Then, an application of Grönwall's inequality yields for any $t \in [0,T]$,

$$\|r_t\|_{L^\infty_\mu(\mathbb{R}^d)} \leqslant \|r_0\|_{L^\infty_\mu(\mathbb{R}^d)} e^{2L_\Phi \|\eta\|_{\infty, C_b(\mathbb{R}^{2d}_{\diagup})} C_{V,T} T}, \tag{4.2.6}$$

thus $r \in L^\infty([0,T]; L^\infty_\mu(\mathbb{R}^d))$. The integrability of the nonlocal divergence follows directly from the previous bound and, thus, $r \in AC([0,T]; L^\infty_\mu(\mathbb{R}^d))$. □

4.2.1 Well-posedness

We proceed with proving well-posedness of (Euler Co-NCL) on an arbitrary time interval $[0,T]$, for $T > 0$. Fix $r_0 \in L^\infty_\mu(\mathbb{R}^d)$, $\eta_0 \in C_b(\mathbb{R}^{2d}_{\diagup})$, and assume Φ is a jointly antisymmetric admissible flux interpolation as in Definition 2.2.3. Denote by

$$\bar{\mathcal{AC}}_{\mathcal{T}} := AC([0,T]; L^\infty_\mu(\mathbb{R}^d)) \times AC([0,T]; C_b(\mathbb{R}^{2d}_{\diagup})),$$

and equip it with the distance $\bar{d}_\infty$. We define the solution map $\mathcal{S} := (\mathcal{S}_V, \mathcal{S}_\omega) : \bar{\mathcal{AC}}_{\mathcal{T}} \to \bar{\mathcal{AC}}_{\mathcal{T}}$ as

$$\begin{aligned}
\mathcal{S}_V(r,\eta)(t)(x) &:= r_0(x) - \int_0^t \int_{\mathbb{R}^d \setminus \{x\}} \Phi(r_s(x), r_s(y); V_s[r](x,y))\eta_s(x,y)\mathrm{d}\mu(y)\mathrm{d}s, \quad \mu\text{-a.e. } x \in \mathbb{R}^d, \\
\mathcal{S}_\omega(r,\eta)(t)(x,y) &:= \eta_0(x,y) + \int_0^t (\omega_s[r](x,y) - \eta_s(x,y)) \, \mathrm{d}s \, ,
\end{aligned}$$

for any $t \in [0,T]$ and $(x,y) \in \mathbb{R}^{2d}_{\diagup}$. The a priori properties in Proposition 4.2.1 ensure this map $\mathcal{S}$ is well defined. Below, we prove existence and uniqueness of solutions to (Euler Co-NCL).

Theorem 4.2.1. Let Φ be a jointly antisymmetric admissible flux interpolation as in Definition 2.2.3 and consider an initial datum $(r^0, \eta^0) \in L^\infty_\mu(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$. Assume that $V : [0,T] \times L^\infty_\mu(\mathbb{R}^d) \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ satisfies (4.2.4) and, for any $g, \tilde{g} \in L^\infty_\mu(\mathbb{R}^d)$,

$$\sup_{t\in[0,T]}\sup_{(x,x')\in\mathbb{R}^{2d}_{\diagup}}|V_t(g_t(x),g_t(x'))-V_t(\tilde{g}_t(x),\tilde{g}_t(x'))|\leqslant L_{V,T}\|g-\tilde{g}\|_{L^\infty_\mu(\mathbb{R}^d)}, \tag{4.2.7}$$

for a constant $L_{V,T}>0$. Assume that $\omega:[0,T]\times L^\infty_\mu(\mathbb{R}^d)\times\mathbb{R}^{2d}_{\diagup}\to\mathbb{R}$ satisfies ($\boldsymbol{\omega}$1)-($\boldsymbol{\omega}$4). Then, there exists a solution (r,η) to (Euler Co-NCL) with initial data (r^0,η^0), in the sense of Definition 4.2.2 .

Proof. For an initial datum (r^0,η^0), we consider $(r,\eta),(\tilde{r},\tilde{\eta})\in\mathcal{AC}_T$ and, for μ-a.e. $x\in\mathbb{R}^d$, we have

$$\begin{aligned}
&|\mathcal{S}_V(r,\eta)(t)-\mathcal{S}_V(\tilde{r},\tilde{\eta})(t)|(x)\\
&\leqslant\int_0^t\int_{\mathbb{R}^d\setminus\{x\}}\Big|\Phi(r_s(x),r_s(y);V_s[r](x,y))\eta_s(x,y)-\Phi(\tilde{r}_s(x),\tilde{r}_s(y);V_s[\tilde{r}](x,y))\tilde{\eta}_s(x,y)\Big|\mathrm{d}\mu(y)\mathrm{d}s\\
&\leqslant L_\Phi\int_0^t\int_{\mathbb{R}^d\setminus\{x\}}\big|\eta_s(x,y)\big|(|r_s(x)-\tilde{r}_s(x)|+|r_s(y)-\tilde{r}_s(y)|)|V_s[r](x,y)|\mathrm{d}\mu(y)\mathrm{d}s\\
&\quad+L_\Phi\int_0^t\int_{\mathbb{R}^d\setminus\{x\}}\left(|\tilde{r}_s(x)|+|\tilde{r}_s(y)|\right)|V_s[r](x,y)|\big|\eta_s(x,y)-\tilde{\eta}_s(x.y)\big|\mathrm{d}\mu(y)\mathrm{d}s\\
&\quad+L_\Phi\int_0^t\int_{\mathbb{R}^d\setminus\{x\}}\big|\tilde{\eta}_s(x,y)\big|\left(|\tilde{r}_s(x)|+|\tilde{r}_s(y)|\right)|V_s[r](x,y)-V_s[\tilde{r}](x,y)|\mathrm{d}\mu(y)\mathrm{d}s\\
&=:I+II+III,
\end{aligned}$$

where we used (i), (ii) in Definition 2.2.3. I can be estimated as follows using (4.2.4)

$$\begin{aligned}
|I|&\leqslant L_\Phi\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}C_{V,T}\int_0^t\left(|r_s(x)-\tilde{r}_s(x)|+\int_{\mathbb{R}^d\setminus\{x\}}|r_s(y)-\tilde{r}_s(y)|\mathrm{d}\mu(y)\right)\mathrm{d}s\\
&\leqslant 2L_\Phi\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}C_{V,T}\|r-\tilde{r}\|_{\infty,L^\infty_\mu(\mathbb{R}^d)}\,T\\
&\leqslant 2L_\Phi\left(\|\eta_0\|_\infty+C_{\omega,T}\right)e^TC_{V,T}\|r-\tilde{r}\|_{\infty,L^\infty_\mu(\mathbb{R}^d)}\,T\ ,
\end{aligned}$$

whence,

$$\|I\|_{\infty,L^\infty_\mu(\mathbb{R}^d)}\leqslant 2L_\Phi\left(\|\eta_0\|_\infty+C_{\omega,T}\right)e^TC_{V,T}\|r-\tilde{r}\|_{\infty,L^\infty_\mu(\mathbb{R}^d)}\,T\ .$$

For II we obtain similarly that

$$|II|\leqslant 2L_\Phi\|\tilde{r}\|_{\infty,L^\infty_\mu(\mathbb{R}^d)}C_{V,T}\|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}T\ ,$$

and thus

$$\begin{aligned}
\|II\|_{\infty,L^\infty_\mu(\mathbb{R}^d)}&\leqslant 2L_\Phi\|r_0\|_{L^\infty_\mu(\mathbb{R}^d)}e^{2L_\Phi\|\eta\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}C_{V,T}T}C_{V,T}\|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}T\\
&\leqslant 2L_\Phi\|r_0\|_{L^\infty_\mu(\mathbb{R}^d)}e^{2L_\Phi\left(\|\eta_0\|_\infty+C_{\omega,T}\right)e^TC_{V,T}T}C_{V,T}\|\eta-\tilde{\eta}\|_{\infty,C_b(\mathbb{R}^{2d}_{\diagup})}T\ .
\end{aligned}$$

Finally, for III we use (4.2.7) to obtain

$$|III| \leqslant 2L_\Phi \|\tilde{\eta}\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \|\tilde{r}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} L_{V,T} \|r - \tilde{r}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} T \ ,$$

Therefore,

$$\begin{aligned}
\|III\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} &\leqslant 2L_\Phi \left(\|\eta_0\|_\infty + C_{\omega,T}\right) \\
&\quad \times \quad \|\tilde{r}_0\|_{L^\infty_\mu(\mathbb{R}^d)} e^{(1+2L_\Phi \|\tilde{\eta}\|_\infty C_{V,T})T} L_{V,T} \|r - \tilde{r}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} T \\
&\leqslant 2L_\Phi \left(\|\eta_0\|_\infty + C_{\omega,T}\right) \\
&\qquad \times \|\tilde{r}_0\|_{L^\infty_\mu(\mathbb{R}^d)} e^{(1+2L_\Phi \left(\|\eta_0\|_\infty + C_{\omega,T}\right) e^T C_{V,T})T} L_{V,T} \|r - \tilde{r}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} T \ .
\end{aligned}$$

For $\mathcal{S}_\omega$, analogous estimates to the ones in Theorem 2.3.1 gives

$$\|\mathcal{S}_\omega(r,\eta) - \mathcal{S}_\omega(\tilde{r},\tilde{\eta})\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} \leqslant \|\eta - \tilde{\eta}\|_{\infty, C_b(\mathbb{R}^{2d}_{\not\Delta})} T + L_{\omega,T} \|r - \tilde{r}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} T \ .$$

Denoting by

$$\begin{aligned}
C^1_T &:= \left(\|\eta_0\|_\infty + C_{\omega,T}\right) e^T \\
C^2_T &:= \|r_0\|_{L^\infty_\mu(\mathbb{R}^d)} e^{2L_\Phi \left(\|\eta_0\|_\infty + C_{\omega,T}\right) e^T C_{V,T} T} \ , \\
C^3_T &:= \left(\|\eta_0\|_\infty + C_{\omega,T}\right) \|\tilde{r}_0\|_{L^\infty_\mu(\mathbb{R}^d)} e^{(1+2L_\Phi \left(\|\eta_0\|_\infty + C_{\omega,T}\right) e^T C_{V,T})T}
\end{aligned}$$

a combination of the previous estimates implies that

$$\begin{aligned}
\bar{d_\infty}((r,\eta),(\tilde{r},\tilde{\eta})) \leqslant & \underbrace{\left(2L_\Phi C^1_T C_{V,T} + 2L_\Phi C^3_T L_{V,T} + L_{\omega,T}\right)}_{:=\alpha(T)} \|r - \tilde{r}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)})T \\
& + \underbrace{\left(2L_\Phi C^2_T C_{V,T} + 1\right)}_{:=\beta(T)} \|\eta - \tilde{\eta}\|_{\infty, L^\infty_\mu(\mathbb{R}^d)} T
\end{aligned}$$

where we used (3.1.6) with ($\boldsymbol{\omega}3$). In particular, it follows that

$$\bar{d_\infty}(\mathcal{S}(r,\eta), \mathcal{S}(\tilde{r},\tilde{\eta})) \leqslant \max\{\alpha(T), \beta(T)\} T \bar{d_\infty}((r,\eta),(\tilde{r},\tilde{\eta})) \ ,$$

so that for $T < \kappa(T) := 1/\max\{\alpha(T), \beta(T)\}$ the solution map is a contraction. By the Banach fixed-point theorem we can conclude there exists a unique solution to (Euler Co-NCL) in $C_T = C([0,T]; L^\infty_\mu(\mathbb{R}^d)) \times C([0,T]; C_b(\mathbb{R}^{2d}_{\not\Delta}))$ for $T < T^*$, where T^* is such that $\kappa(T^*)^{-1} T^* = 1$. This can be verified since $\alpha(T) \simeq 1 + e^T + e^{Te^T}$ and $\beta(T) \simeq 1 + e^{Te^T}$. Absolute continuity in time follows since $\overline{\nabla} \cdot F[\mu,\eta; r, V[r]]$ and the map $(t \mapsto \omega_t[r_t] - \eta_t)$ belong to $L^1([0,T])$, as proved in Proposition 4.2.1. The solution can be extended by using a standard iteration procedure so that the time interval does not depend on any T^*, as in Theorem 2.3.1. □

4.3 Long-time behaviour

In this section, we study the long-time behaviour of (Euler Co-NCL). We assume that the vertices of the graph belong to a compact space, i.e. $\operatorname{supp}\mu = K \subset \mathbb{R}^d$ for K compact. Analogously to the previous section, we use the following notation $K^2_{\not=} = \{(x,y) \in K^2 \mid x \neq y\}$. We also need the following technical results, where the first three are a direct adaptation of [BDS22, Theorem 2.1], [BDS22, Theorem 2.2] and [BDS22, Lemma 2.3], respectively (see also [BL73] and [Dan67] for the original reference of Theorem 4.3.1 and Theorem 4.3.2, respectively).

Theorem 4.3.1 (Scorza-Dragoni)**.** Let $\Omega \subset \mathbb{R}^d$ be a locally compact space and $f : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ be such that $x \in \Omega \mapsto f(t,x) \in \mathbb{R}$ is μ-measurable for each $t \geqslant 0$, and $t \in \mathbb{R}_+ \mapsto f(t,x) \in \mathbb{R}$ is continuous for μ-almost every $x \in \Omega$. Then, for every $\varepsilon > 0$, there exists a compact set $\Omega_\varepsilon \subset \Omega$ satisfying $\mu(\Omega \backslash \Omega_\varepsilon) < \varepsilon$ and such that the restricted map $f : \mathbb{R}_+ \times \Omega_\varepsilon \to \mathbb{R}$ is continuous.

Theorem 4.3.2 (Danskin)**.** Let $\Omega \subset \mathbb{R}^d$ be a compact set and $f : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ be a continuous function such that $t \in \mathbb{R}_+ \mapsto f(t,x) \in \mathbb{R}$ is differentiable for all $x \in \Omega$. Then, the application $g : t \in \mathbb{R}_+ \mapsto \max_{x\in\Omega} f(t,x) \in \mathbb{R}$ is differentiable $\mathscr{L}^1$-almost everywhere, with

$$\frac{d}{dt} g(t) = \max_{x \in \overline{\Omega}(t)} \partial_t f(t,x)$$

for $\mathscr{L}^1$-almost every $t \geqslant 0$, where we introduced the notation $\overline{\Omega}(t) := \arg\max_{x\in\Omega} f(t,x)$.

Lemma 4.3.1 (Interior estimates for supremums)**.** Let $\Omega \subset \mathbb{R}^d$ be a compact set and $f \in L^\infty_\mu(\Omega, \mathbb{R}^d)$. Then, for every $\delta > 0$, there exists $\varepsilon > 0$ such that

$$\operatorname*{ess\,sup}_{x\in\Omega} f(x) - \delta \leqslant \operatorname*{ess\,sup}_{x\in\Omega^\varepsilon} f(x) \leqslant \operatorname*{ess\,sup}_{x\in\Omega} f(x), \tag{4.3.1}$$

whenever $\Omega_\varepsilon \subset \Omega$ is a measurable set satisfying $\mu(\Omega\backslash\Omega_\varepsilon) < \varepsilon$. In particular, it holds that

$$\operatorname*{ess\,sup}_{x\in\Omega^\varepsilon} f(x) \xrightarrow{\varepsilon\to 0^+} \operatorname*{ess\,sup}_{x\in\Omega} f(x),$$

for each family of sets $(\Omega_\varepsilon)_{\varepsilon>0} \subset \Omega$ satisfying these properties.

Proof. The second inequality in (4.3.1) is satisfied for any subset of Ω. For the first inequality, by definition of the essential supremum of a set, for every $\delta > 0$ there exists a set $\Omega_\delta \subset \Omega$ of positive measure such that

$$f(x) > \operatorname*{ess\,sup}_{x\in\Omega} f(x) - \delta$$

for μ-a.e. $x \in \Omega_\delta$. Then, by choosing $\varepsilon > 0$ satisfying $\varepsilon < \mu(\Omega_\delta)$ and considering any closed set $\Omega_\varepsilon \subset \Omega$ with $\mu(\Omega\backslash\Omega_\varepsilon) < \varepsilon$, it necessarily holds that $\mu(\Omega_\delta \cap \Omega_\varepsilon) > 0$. Indeed assuming the

converse, by contradiction, one would have

$$\mu(\Omega_\delta \cup \Omega_\varepsilon) = \mu(\Omega_\delta) + \mu(\Omega_\varepsilon) > \mu(\Omega) \ ,$$

which is a contradiction since both sets are subsets of Ω. Hence, by construction we have

$$f(x) > \operatorname*{ess\,sup}_{x\in\Omega} f(x) - \delta \ ,$$

for μ-a.e. $x \in \Omega_\delta \cap \Omega_\varepsilon$, which concludes the proof. □

We will also resort to the following generalised version of Grönwall's inequality

Lemma 4.3.2. Let $\phi : \mathbb{R} \to \mathbb{R}$ be a locally-Lipschitz function. Assume that $f \in C(\mathbb{R}^+, \mathbb{R})$ solves the inequality $f(t) \leqslant f(0) + \int_0^t \phi(f(s))\mathrm{d}s$ with $f(0) = a \in \mathbb{R}$ and $g \in C(\mathbb{R}^+, \mathbb{R})$ solves the equation $g(t) = g(0) + \int_0^t \phi(g(s))\mathrm{d}s$ with $f(0) \leqslant g(0)$. Then for $t \geqslant 0$

$$f(t) \leqslant g(t) \ .$$

Proof. Let $h(t) := f(t) - g(t)$ and, for a contradiction, assume there is $b > 0$ such that $h(b) > 0$. By continuity of h, Bolzano's theorem implies that there must be a $c \in (0, b)$ such that $h(t) \geqslant 0$ for $t \in (c, b]$ and $h(c) = 0$. Then, for $t \in (c, b]$, we have

$$h(t) = h(c) + \int_c^t \phi(f(s)) - \phi(g(s))\mathrm{d}s \leqslant h(c) + L\int_c^t (f(s) - g(s))\mathrm{d}s \leqslant h(c) + L\int_c^t h(s)\mathrm{d}s \ \ ,$$

where, for $t \in (c, b]$, $L > 0$ is the local-Lipschitz constant of ϕ. Then, by Grönwall's inequality

$$h(b) \leqslant h(c)e^{L(b-c)} = 0 \ ,$$

which gives the contradiction. □

We will also require the following result guaranteeing the opposite differential inequality, which can be proven in a similar way.

Lemma 4.3.3. Let $\phi : \mathbb{R} \to \mathbb{R}$ be a locally-Lipschitz function. Assume that $f \in C(\mathbb{R}^+, \mathbb{R})$ solves the inequality $f(t) \geqslant f(0) + \int_0^t \phi(f(s))\mathrm{d}s$ with $f(0) = a \in \mathbb{R}$ and $g \in C(\mathbb{R}^+, \mathbb{R})$ solves the equation $g(t) = g(0) + \int_0^t \phi(g(s))\mathrm{d}s$ with $f(0) \geqslant g(0)$. Then for $t \geqslant 0$

$$f(t) \geqslant g(t) \ .$$

Throughout this section, we consider (Euler Co-NCL) on the time domain $[0, \infty)$. The

well-posedness follows upon adapting assumptions ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}3$), (4.2.3), (4.2.4) and (4.2.7) to hold on $[0,\infty)$, where we now drop the T subscript of the corresponding constant of each assumption. We will also require the following stronger assumption on the regularity of the velocity.

Assumption 1. Assume that $V : [0,\infty) \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathcal{V}^{as}(K^2_{\diagup})$ is continuous in time and is of the form

$$V(r_t(x), r_t(x')) = \alpha(r_t(x)) - \alpha(r_t(x'))$$

where $\alpha \in C^1_b(\mathbb{R}^d)$ and it is monotonically increasing, i.e.

$$\alpha'(x) = \frac{d}{dx}\alpha(x) > c \ ,$$

for some $c > 0$, for $x \in K$. We will denote by

$$\alpha'_* := \inf_{x\in K} \alpha'(x).$$

To obtain our long-time behaviour result we need to impose some further conditions on η. First, let us recall that we can obtain an explicit representation for η given by

$$\eta_t(x,y) = e^{-t}\eta_0(x,y) + \int_0^t e^{-(t-s)}\omega_s[r](x,y)\mathrm{d}s \ . \tag{4.3.2}$$

We refer the reader to Lemma A.0.1 in Appendix A for a derivation of (4.3.2). Using (4.3.2), we can obtain positivity preservation for the edge-weight function η as well as a sharper bound on its supremum norm which we present in the following two results.

Lemma 4.3.4. Let $(r,\eta) \in AC([0,\infty), L^\infty_\mu(\mathbb{R}^d)) \times AC([0,\infty), C_b(K^2_{\diagup}))$ be the solution of (Euler Co-NCL) with initial datum $(r_0,\eta_0) \in (L^\infty_\mu(\mathbb{R}^d), C_b(K^2_{\diagup}))$. If in addition to assumptions ($\boldsymbol{\omega}1$)-($\boldsymbol{\omega}3$) we impose

$$\omega_* := \inf_{t\in[0,\infty)} \inf_{(x,x')\in K^2_{\diagup}} \inf_{r\in L^\infty_\mu(\mathbb{R}^d)\setminus\{0\}} \omega_t[r](x,x') > 0 \ , \tag{4.3.3}$$

and $\eta_0(x,x') > 0$, then $\eta_t > 0$ for $t \in [0,\infty)$. Furthermore, if $\eta^0 \in C_b(K^2_{\diagup})$ is symmetric, and ω satisfies assumption ($\boldsymbol{\omega}4$) for any $t \in [0,\infty)$ and $r \in L^\infty_\mu(\mathbb{R}^d)$, η_t is symmetric for $t \in [0,\infty)$.

Proof. Using the representation for η in (4.3.2) we have for any $(x,x') \in K^2_{\diagup}$ and $t \geqslant 0$

$$\begin{aligned}\eta_t(x,x') &= e^{-t}\left(\eta_0(x,x') + \int_0^t e^s\omega_s\left[r\right](x,x')\mathrm{d}s\right)\\ &\geqslant \inf_{(x,x')\in K^2_{\diagup}} \eta_0(x,x')e^{-t} + \omega_*(1-e^{-t}) > 0\end{aligned}$$

by (4.3.3) and our assumption on η_0. The claim about symmetry follows directly from the representation (4.3.2). □

Lemma 4.3.5. Let $(r,\eta) \in AC([0,T], L^\infty_\mu(\mathbb{R}^d)) \times AC([0,T], C_b(K^2_{\diagup}))$ solve (Euler Co-NCL) with initial datum $(r_0,\eta_0) \in (L^\infty_\mu(\mathbb{R}^d), C_b(K^2_{\diagup}))$ where ω satisfies ($\boldsymbol{\omega}$1)-($\boldsymbol{\omega}$4) and the corresponding version of (4.2.3). Then

$$\|\eta_t\|_{\infty, C_b(K^2_{\diagup})} \leqslant \|\eta_0\|_{\infty, C_b(K^2_{\diagup})} + C_\omega \ . \tag{4.3.4}$$

Proof. The proof follows directly from the representation (4.3.2). □

In what follows we will use use the notation

$$\eta_* = \inf_{t\in[0,\infty)} \inf_{(x,x')\in K^2_{\diagup}} \eta_t(x,x') \ ,$$

and we have $\eta_* > 0$ in view of Lemma 4.3.4. With these assumptions and notations we have the following theorem.

Theorem 4.3.3. Fix $\Phi \equiv \Phi_{upwind}$. Let $(r,\eta) \in AC(\mathbb{R}^+, L^\infty_\mu(K)) \times AC(\mathbb{R}^+, C_b(K^2_{\diagup}))$ be a solution of (Euler Co-NCL) with a pointwise monotonic velocity field $V : \mathbb{R}^+ \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathcal{V}^{as}(K^2_{\diagup})$ satisfying (4.2.4), (4.2.7), and Assumption 1. Assume $\omega : [0,\infty) \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathbb{R}$ satisfies ($\boldsymbol{\omega}$1)-($\boldsymbol{\omega}$4), (4.2.3) and (4.3.3). Assume $r_0 \geqslant 0$ and $\int_K r_0(x)\mathrm{d}\mu(x) = M$ and that $\eta^0 \in C_b(K^2_{\diagup})$ is symmetric and $\eta_0 > 0$. Then we have

$$\|r_t\|_{L^\infty_\mu(K)} \leqslant \frac{\|r_0\|_{L^\infty_\mu(K)} \alpha'_* \eta_* M e^{\alpha'_* \eta_* M t}}{\alpha'_* \eta_* \left[\|r_0\|_{L^\infty_\mu(K)} \mu(K)(e^{\alpha'_* \eta_* M t} - 1) + M \right]} \ . \tag{4.3.5}$$

Proof. We begin by noting that, given our assumptions on ω and η^0, we have that η_t is positive and symmetric for $t \geqslant 0$. Thus, since we are restricting to the upwind interpolation, by [EPS23, Proposition 5.2], (Euler Co-NCL) is non-negativity preserving on $t \geqslant 0$. Next, fix $\varepsilon > 0$ and observe that by Theorem 4.3.1 there exists a compact set $K^\varepsilon \subset K$ such that the restriction $r : \mathbb{R}^+ \times K^\varepsilon \to \mathbb{R}$ is continuous and $\mu(K \backslash K^\varepsilon) < \varepsilon$. For all $t \geqslant 0$, we define the family of restricted supremum norms

$$L_\varepsilon(t) := \max_{x\in K^\varepsilon} |r_t(x)| \ .$$

Since (r,η) is the solution of (Euler Co-NCL), we have $r \in AC([0,\infty), L^\infty_\mu(\mathbb{R}^d))$. As any absolutely continuous curve can be reparametrized in time to be Lipschitz continuous (see [San15, Box 5.1]), the mapping $L_\varepsilon(t)$ is Lipschitz, being the pointwise maximum of a family of equi-Lipschitz continuous functions. In particular, it is differentiable $\mathscr{L}^1$ almost everywhere by

Rademacher's theorem. Let us denote by $\mathbb{V}_\varepsilon(t) := \arg\max_{x\in K^\varepsilon} r_t(x)$. Then, by Theorem 4.3.2 we have

$$\partial_t L_\varepsilon(t) = \max_{x\in\mathbb{V}_\varepsilon(t)} \partial_t r_t(x) \ .$$

Thus, taking an arbitrary vertex $x \in \mathbb{V}_\varepsilon(t)$, we have

$$\begin{aligned}
\partial_t r_t(x) = & - r_t(x) \int_{K^\varepsilon\setminus\{x\}} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& - r_t(x) \int_{K\setminus K^\varepsilon} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K^\varepsilon\setminus\{x\}} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K\setminus K^\varepsilon} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \ .
\end{aligned} \tag{4.3.6}$$

For any $x \in \mathbb{V}_\varepsilon(t)$, it holds

$$\int_{K^\varepsilon\setminus\{x\}} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') = 0 \ ,$$

since $V_t^-(r_t(x), r_t(x')) = \big(\alpha(r_t(x')) - \alpha(r_t(x))\big)_+ = 0$, $x \in \mathbb{V}_\varepsilon(t)$, $x' \in K^\varepsilon\setminus\{x\}$ and α is monotonically increasing. Furthermore, since $\int_{K\setminus K^\varepsilon} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \geqslant 0$, by $r_t \geqslant 0$ we infer

$$\begin{aligned}
\partial_t r_t(x) \leqslant & - r_t(x) \int_{K^\varepsilon\setminus\{x\}} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K\setminus K^\varepsilon} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \ .
\end{aligned}$$

In particular, recalling that for any $x \in \mathbb{V}_\varepsilon(t)$ we have $r_t(x) = L_\varepsilon(t)$ we can rewrite

$$\begin{aligned}
\partial_t L_\varepsilon(t) \leqslant & - L_\varepsilon(t) \int_{K^\varepsilon\setminus\{x\}} V_t^+(L_\varepsilon(t), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K\setminus K^\varepsilon} r_t(x') V_t^-(L_\varepsilon(t), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
\leqslant & - L_\varepsilon(t)\eta_* \int_{K^\varepsilon\setminus\{x\}} V_t^+(L_\varepsilon^+(t), r_t(x'))\mathrm{d}\mu(x') \\
& + C_V \|r_t\|_{L^\infty_\mu(K)} \|\eta\|_{\infty, C_b(K^2_{\diagup})} \varepsilon \ ,
\end{aligned}$$

where we used the boundedness of V and η and that $\mu(K\setminus K^\varepsilon) < \varepsilon$. The monotonicity of V, for $x \in \mathbb{V}_\varepsilon(t)$ and μ-a.e. $x' \in K^\varepsilon$,

$$V_t^+(L_\varepsilon(t), r_t(x')) = V_t(L_\varepsilon(t), r_t(x')) \ .$$

Furthermore, for velocities satisfying Assumption 1, the mean-value theorem gives

$$\begin{aligned}\alpha(L_\varepsilon(t)) - \alpha(r_t(x')) &= \alpha'(c_{t,x})(L_\varepsilon(t) - r_t(x')) \\ &\geqslant \alpha'_*(L_\varepsilon(t) - r_t(x'))\ ,\end{aligned}$$

for some $c_{t,x} \in [L_\varepsilon(t), r_t(x')]$, recalling that $\alpha'_* = \inf_{x\in K} \alpha'(x)$. Hence,

$$\begin{aligned}\partial_t L_\varepsilon(t) \leqslant\ & - L_\varepsilon(t)\eta_*\alpha'_* \int_{K^\varepsilon\setminus\{x\}} (L_\varepsilon(t) - r_t(x'))\mathrm{d}\mu(x') \\ & + C_V\|r_t\|_{L^\infty_\mu(K)}\|\eta\|_{\infty,C_b(K^2_{\diagup})}\varepsilon\ , \\ =\ & - L^2_\varepsilon(t)\eta_*\alpha'_*\mu(K^\varepsilon) + \eta_*\alpha'_* L_\varepsilon(t)\int_{K^\varepsilon\setminus\{x\}} r_t(x')\mathrm{d}\mu(x') \\ & + C_V\|r_t\|_{L^\infty_\mu(K)}\|\eta\|_{\infty,C_b(K^2_{\diagup})}\varepsilon\ ,\end{aligned}$$

whence, integrating with respect to time yields

$$\begin{aligned}L_\varepsilon(t) \leqslant L_\varepsilon(0) &- \int_0^t \left[L^2_\varepsilon(s)\eta_*\alpha'_*\mu(K^\varepsilon) - \eta_*\alpha'_* L_\varepsilon(s)\int_{K^\varepsilon\setminus\{x\}} r_s(x')\mathrm{d}\mu(x')\right]\mathrm{d}s \\ &+ \int_0^t C_V\|r_s\|_{L^\infty_\mu(K)}\|\eta\|_{\infty,C_b(K^2_{\diagup})}\varepsilon\mathrm{d}s\ .\end{aligned}$$

Letting $\varepsilon \to 0^+$, Lemma 4.3.1 implies

$$\|r_t\|_{L^\infty_\mu(\mathbb{R}^d)} \leqslant \|r_0\|_{L^\infty_\mu(\mathbb{R}^d)} - \int_0^t \left[\|r_s\|^2_{L^\infty_\mu(\mathbb{R}^d)}\eta_*\alpha'_*\mu(K) - \eta_*\alpha'_*\|r_s\|_{L^\infty_\mu(\mathbb{R}^d)}M\right]\mathrm{d}s\ , \tag{4.3.7}$$

where we have used that $\int_K r_t(x)\mathrm{d}\mu(x') = M$, by the conservation of mass property (see Lemma A.0.2). Noting that (4.3.7) is an inequality version of a Bernoulli ODE (see Lemma A.0.3), using Lemma 4.3.2 we infer

$$\|r_t\|_{L^\infty_\mu(K)} \leqslant \frac{\|r_0\|_{L^\infty_\mu(K)}\alpha'_*\eta_* M e^{\alpha'_*\eta_* Mt}}{\alpha'_*\eta_*\left[\|r_0\|_{L^\infty_\mu(K)}\mu(K)(e^{\alpha'_*\eta_* Mt} - 1) + M\right]}\ ,$$

which gives the result. □

Next, we find a lower bound for $r_{t,*} := \inf_{x\in K} r_t(x)$. Note that since r is lower bounded and K is compact, the infimum and the minimum coincide.

Theorem 4.3.4. Fix $\Phi \equiv \Phi_{upwind}$. Let $(r,\eta) \in AC(\mathbb{R}^+, L^\infty_\mu(K)) \times AC(\mathbb{R}^+, C_b(K^2_{\diagup}))$ be a solution of (Euler Co-NCL) with a pointwise monotonic velocity field $V : \mathbb{R}^+ \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathcal{V}^{as}(K^2_{\diagup})$ satisfying (4.2.4), (4.2.7), and Assumption 1. Let $\omega : [0,\infty) \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathbb{R}$ satisfy ($\boldsymbol{\omega}$1)-($\boldsymbol{\omega}$4) and (4.2.3). Assume $r_0 \geqslant 0$ and $\int_K r_0(x)\mathrm{d}\mu(x) = M$ and $\eta^0 \in C_b(K^2_{\diagup})$ is positive and symmetric.

Then we have

$$r_{t,*} \geqslant \frac{r_{0,*}\eta_*\alpha'_* M e^{\eta_*\alpha'_* Mt}}{\eta_*\alpha'_*(M + (e^{\eta_*\alpha'_* Mt} - 1)\mu(K) r_{0,*})} \,. \tag{4.3.8}$$

Proof. As in Theorem 4.3.3, by our assumptions on ω, r^0, η^0 and the use of the upwind interpolation, we have non-negativity preservation of r for $t \geqslant 0$ by [EPS23, Proposition 5.2]. Furthermore, let us consider the Lipschitz continuous reparametrization of the absolutely continuous curve r_t. Denote $r_t^- := -r_t$ and fix $\varepsilon > 0$. By Theorem 4.3.1 there exists a compact set $K^\varepsilon \subset K$ such the restriction $r^- : [0,T] \times K^\varepsilon \to \mathbb{R}$ is continuous and $\mu(K \backslash K^\varepsilon) < \varepsilon$ for some $\varepsilon > 0$. For all $t \geqslant 0$ define the family of restricted maxima on K^ε, i.e.

$$L_\varepsilon^-(t) := \max_{x \in K^\varepsilon} r_t^-(x) \,.$$

Since $L_\varepsilon^-(\cdot)$ is the pointwise maximum of equi-Lipschitz continuous functions Theorem 4.3.2 implies

$$\partial_t L_\varepsilon^-(t) = \partial_t \max_{x \in \mathbb{V}_\varepsilon^-(t)} r_t^-(x) = \max_{x \in \mathbb{V}_\varepsilon^-(t)} \partial_t r_t^-(x) = \max_{x \in \mathbb{V}_\varepsilon^-(t)} -\partial_t r_t(x) \,, \tag{4.3.9}$$

where $\mathbb{V}_\varepsilon^-(t) = \arg\max_{x \in K^\varepsilon} r_t^-(x)$. Then, for any vertex $x \in \mathbb{V}_\varepsilon^-(t)$ it holds that

$$\begin{aligned}
\partial_t r_t^-(x) = {} & r_t(x) \int_{K^\varepsilon \backslash \{x\}} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + r_t(x) \int_{K \backslash K^\varepsilon} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& - \int_{K^\varepsilon \backslash \{x\}} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& - \int_{K \backslash K^\varepsilon} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \,.
\end{aligned}$$

Note that $r_{t,*}^\varepsilon := \min_{x \in K^\varepsilon} r_t(x) = -\max_{x \in K^\varepsilon} r_t^-(x)$. Then, in view of (4.3.9) we have for $x \in \mathbb{V}_\varepsilon^-(t)$ that

$$\begin{aligned}
\partial_t r_t(x) = {} & - r_t(x) \int_{K^\varepsilon \backslash \{x\}} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& - r_t(x) \int_{K \backslash K^\varepsilon} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K^\varepsilon \backslash \{x\}} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K \backslash K^\varepsilon} r_t(x') V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x')
\end{aligned} \tag{4.3.10}$$

which is the evolution of $r_t(x)$ for $x \in \arg\min_{x\in K^\varepsilon} r_t(x)$, since for any function f such that the maximum and the minimum exist, $\arg\max -f = \arg\min f$. Furthermore, this implies that

$$V_t^+(r_t(x), r_t(x')) = \Big(\alpha(r_t(x)) - \alpha(r_t(x'))\Big)_+ = 0 \ , \tag{4.3.11}$$

for $x \in \mathbb{V}_\varepsilon^-(t)$ and $x' \in K^\varepsilon\backslash\{x\}$, since $\mathbb{V}_\varepsilon^-(t) = \arg\max_{x\in K^\varepsilon} -r_t(x) = \arg\min_{x\in K^\varepsilon} r_t(x)$ and α is monotonic. In particular, the first term vanishes. Furthermore, noting that the last term is non-negative, we can drop it to obtain the following lower bound

$$\begin{aligned}
\partial_t r_t(x) \geqslant & -r_t(x)\int_{K\backslash K^\varepsilon} V_t^+(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
& + \int_{K^\varepsilon\backslash\{x\}} r_t(x')V_t^-(r_t(x), r_t(x'))\eta_t(x,x')\mathrm{d}\mu(x') \\
\geqslant & -\|r_t\|_{L^\infty_\mu(K)} C_V \|\eta\|_{\infty, C_b(K^2_{\diagup})}\varepsilon \\
& + \int_{K^\varepsilon\backslash\{x\}} r_t(x')\Big(\alpha(r_t(x')) - \alpha(r_t(x))\Big)_+ \eta_t(x,x')\mathrm{d}\mu(x') \\
= & -\|r_t\|_{L^\infty_\mu(K)} C_V \|\eta\|_{\infty, C_b(K^2_{\diagup})}\varepsilon \\
& + \int_{K^\varepsilon\backslash\{x\}} r_t(x')(\alpha(r_t(x')) - \alpha(r_t(x)))\eta_t(x,x')\mathrm{d}\mu(x') \\
\geqslant & -\|r_t\|_{L^\infty_\mu(K)} C_V \|\eta\|_{\infty, C_b(K^2_{\diagup})}\varepsilon \\
& + \eta_*\alpha'_*\int_{K^\varepsilon\backslash\{x\}} r_t(x')(r_t(x') - r_t(x))\mathrm{d}\mu(x')
\end{aligned}$$

where in the second and third inequality we used that $x \in \mathbb{V}_\varepsilon^-(t) = \arg\min_{x\in K^\varepsilon} r_t(x)$, and in the third inequality we also used the monotonicity of the velocity. Since for $x \in \mathbb{V}_\varepsilon^-(t)$, $r_t(x) = r^\varepsilon_{t,*}$ we have,

$$\begin{aligned}
\partial_t r^\varepsilon_{t,*} = & -\|r_t\|_{L^\infty_\mu(K)} C_V \|\eta\|_{\infty, C_b(K^2_{\diagup})}\varepsilon \\
& + \eta_*\alpha'_*\int_{K^\varepsilon} r_t(x')(r_t(x') - r^\varepsilon_{t,*})\mathrm{d}\mu(x') \\
\geqslant & -\|r_t\|_{L^\infty_\mu(K)} C_V \|\eta\|_{\infty, C_b(K^2_{\diagup})}\varepsilon \\
& + \eta_*\alpha'_*\left(r^\varepsilon_{t,*}\int_{K^\varepsilon} r_t(x')\mathrm{d}\mu(x') - (r^\varepsilon_{t,*})^2\mu(K^\varepsilon)\right) \ .
\end{aligned}$$

Integrating in time, sending $\varepsilon \to 0^+$ and using Lemma 4.3.1 we obtain

$$r_{t,*} \geqslant r_{0,*} + \int_0^t \eta_*\alpha'_*(r_{s,*}M - (r_{s,*})^2\mu(K))\mathrm{d}s, \tag{4.3.12}$$

where we used that $\int_K r_t(x)\mathrm{d}\mu(x) = M$ and the preservation of mass property from Lemma

A.0.2. Noting that (4.3.12) resembles a Bernoulli ODE, Lemma 4.3.3 and Lemma A.0.3 imply

$$r_{t,*} \geqslant \frac{r_{0,*}\eta_*\alpha'_* M e^{\eta_*\alpha'_* Mt}}{\eta_*\alpha'_*(M + (e^{\eta_*\alpha'_* Mt} - 1)\mu(K)r_{0,*})} \ .$$

□

Now we are ready to obtain the long-time behaviour of (Euler Co-NCL) for pointwise monotonic velocities.

Theorem 4.3.5. Fix $\Phi \equiv \Phi_{upwind}$. Let $(r, \eta) \in AC(\mathbb{R}^+, L^\infty_\mu(K)) \times AC(\mathbb{R}^+, C_b(K^2_{\diagup}))$ be a solution of (Euler Co-NCL) with a pointwise monotonic velocity field $V : \mathbb{R}^+ \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathcal{V}^{as}(K)$ satisfying (4.2.4), (4.2.7) and Assumption 1. Let $\omega : [0, \infty) \times L^\infty_\mu(K) \times K^2_{\diagup} \to \mathbb{R}$ satisfy ($\boldsymbol{\omega}$1)-($\boldsymbol{\omega}$4) and (4.2.3). Assume $r_0 \geqslant 0$ and $\int_K r_0(x)\mathrm{d}\mu(x) = M$ and $\eta^0 \in C_b(K^2_{\diagup})$ is positive and symmetric. Then we have that for μ-a.e. $x \in \mathbb{R}^d$

$$\lim_{t\to\infty} r_t(x) = \frac{M}{\mu(K)} \ ,$$

the uniform distribution of mass over K.

Proof. For μ-a.e. $x \in \mathbb{R}^d$ and $t \geqslant 0$ we have by (4.3.5) and (4.3.8), that

$$\begin{aligned}
\frac{r_{0,*}\eta_*\alpha'_* M e^{\eta_*\alpha'_* Mt}}{\eta_*\alpha'_*(M + (e^{\eta_*\alpha'_* Mt} - 1)\mu(K)r_{0,*})} &\leqslant r_{t,*} \\
&\leqslant r_t(x) \\
&\leqslant \|r_t\|_{L^\infty_\mu(K)} \\
&\leqslant \frac{\|r_0\|_{L^\infty_\mu(K)}\alpha'_*\eta_* M e^{\alpha'_*\eta_* Mt}}{\alpha'_*\eta_*\left[\|r_0\|_{L^\infty_\mu(K)}\mu(K)(e^{\alpha'_*\eta_* Mt} - 1) + M\right]} \ ,
\end{aligned}$$

and letting $t \to \infty$ yields the result. □

Appendix A

Auxiliary results

The next result is important for the application of the Banach fixed-point theorem in Section 3.2 of Chapter 3. The proof closely follows [KVM18, Lemma 2.1]

Proposition A.0.1. $(\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d), L_\mu^2 d_2)$ is a complete metric space.

Proof. $L_\mu^2 d_2$ is a metric since, d_2^2 is a metric. Next, we prove completeness of $(\mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d), L_\mu^2 d_2)$. To this end, let $\{\sigma^n\}_{n\in\mathbb{N}}$ be a Cauchy sequence. Then, we can find an increasing sequence of indices n_k such that

$$L_\mu^2 d_2(\sigma^{n_k}, \sigma^{n_{k+1}}) = \int_{\mathbb{R}^d} d_2^2(\sigma^{n_k}, \sigma^{n_{k+1}})\mathrm{d}\mu(x) \leqslant \frac{1}{2^{k+1}} \quad \text{for } k = 1, 2, \dots \ .$$

By Levi's convergence theorem we infer that

$$\sum_{k=1}^{\infty} d_2^2(\sigma_x^{n_k}, \sigma_x^{n_{k+1}})$$

converges for μ-a.e. $x \in \mathbb{R}^d$ to a function $W \in L_\mu^1(\mathbb{R}^d)$ and we have

$$\sum_{k=1}^{\infty} \int_{\mathbb{R}^d} d_2^2(\sigma_x^{n_k}, \sigma_x^{n_{k+1}})\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} W(x)\mathrm{d}\mu(x) < \infty \ . \tag{A.0.1}$$

Then, for any $i < j \in \mathbb{N}$,

$$d_2^2(\sigma_x^{n_i}, \sigma_x^{n_j}) \leqslant \sum_{k=i}^{j-1} d_2^2(\sigma_x^{n_k}, \sigma_x^{n_{k+1}})$$

and, since the series $\sum_{k=1}^{\infty} d_2^2(\sigma_x^{n_k}, \sigma_x^{n_{k+1}})$ is convergent, the Cauchy criterion for convergence of series yields that the sequence $\{\sigma_x^{n_k}\}$ is Cauchy for μ-a.e. $x \in \mathbb{R}^d$. By completness of $(\mathcal{P}_2(\mathbb{R}), d_2^2)$, there exists $\sigma_x = \lim_{n\to\infty} \sigma_x^n$ and $\sigma_x \in \mathcal{P}_2(\mathbb{R})$ for μ-a.e. $x \in \mathbb{R}^d$. We define the natural limit candidate $\bar{\sigma} = \int \sigma_x \mathrm{d}\mu(x)$ and note that $\bar{\sigma} \in \mathcal{P}_2^\mu(\mathbb{R}\times\mathbb{R}^d)$.

Let $\{\sigma^{n_i}\}$ be a subsequence of $\{\sigma^n\}$. For any $i < j$ we have

$$\begin{aligned} L_\mu^2 d_2(\sigma^{n_i}, \sigma^{n_j}) &= \int_{\mathbb{R}^d} d_2^2(\sigma_x^{n_i}, \sigma_x^{n_j}) \mathrm{d}\mu(x) \leqslant \int_{\mathbb{R}^d} \sum_{k=i}^{j-1} d_2^2(\sigma_x^{n_k}, \sigma_x^{n_{k+1}}) \mathrm{d}\mu(x) \\ &= \sum_{k=i}^{j-1} \int_{\mathbb{R}^d} d_2^2(\sigma_x^{n_k}, \sigma_x^{n_{k+1}}) \mathrm{d}\mu(x) \leqslant \sum_{k=i}^{j-1} \frac{1}{2^{k+1}} = \frac{1}{2^{i+1}} \left(\sum_{k=0}^{j-i-1} \frac{1}{2^k} \right) \\ &< \frac{1}{2^{i+1}} \left(\sum_{k=0}^{\infty} \frac{1}{2^k} \right) = \frac{1}{2^i} \ . \end{aligned}$$

Thus, $\{\sigma^{n_i}\}$ is Cauchy. Then, by (A.0.1) we have that for any $i < j$, $d_2^2(\sigma_x^{n_i}, \sigma_x^{n_j}) \leqslant W(x)$ for μ-a.e. $x \in \mathbb{R}^d$ and we know that $W \in L_\mu^1(\mathbb{R}^d)$. The Dominated Convergence Theorem implies

$$\begin{aligned} L_\mu^2 d_2(\sigma^{n_i}, \bar{\sigma}) &= \int_{\mathbb{R}^d} d_2^2(\sigma_x^{n_i}, \sigma_x) \mathrm{d}\mu(x) = \int_{\mathbb{R}^d} \lim_{j \to \infty} d_2^2(\sigma_x^{n_i}, \sigma_x^{n_j}) \mathrm{d}\mu(x) \\ &= \lim_{j \to \infty} \int_{\mathbb{R}^d} d_2^2(\sigma_x^{n_i}, \sigma_x^{n_j}) \mathrm{d}\mu(x) < \frac{1}{2^i} \to 0 \ , \end{aligned}$$

as $i \to \infty$. Since, any subsequence of $\{\sigma^n\}$ converges to $\bar{\sigma}$, we have that the whole sequence converges to $\bar{\sigma}$ and $\mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d)$ is complete. □

We present the proposition giving rise to the condition (3.2.16) discussed in Section 3.2.1 in Chapter 3. As we shall see, even in the simpler case of $\Phi \equiv \Phi_{upwind}$ and V and ω not depending on $r \in L_\mu^2(\mathbb{R}^d)$, so that the system (Euler Co-NCL) decouples, i.e. the graph is no longer co-evolving, we could not find a way around this condition to obtain a contraction in $(\mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d), L_\mu^2 d_2)$.

Proposition A.0.2. Fix $\Phi \equiv \Phi_{upwind}$. Let $\sigma^i \in C([0,T], \mathcal{P}_2^\mu(\mathbb{R} \times \mathbb{R}^d))$ be μ-monokinetic solutions of (graph-CE), with initial data $\sigma_0^i = \delta_{r_0^i(\cdot)} \otimes \mu$, where (r^i, η) are solutions to (Euler Co-NCL) starting from $(r_0^i, \bar{\eta}) \in L_\mu^2(\mathbb{R}^d) \times C_b(\mathbb{R}^{2d}_{\diagup})$ for $i = 1, 2$. Let $V : [0,T] \times \mathbb{R}^{2d}_{\diagup} \to \mathcal{V}^{as}(\mathbb{R}^{2d}_{\diagup})$ be an antisymmetric velocity field which is continuous in time and satisfies (3.1.9a) and (3.1.9b), and $\omega : [0,T] \times \mathbb{R}^{2d}_{\diagup} \to \mathbb{R}$ satisfy ($\boldsymbol{\omega}1$)-($\tilde{\boldsymbol{\omega}}4$). Consider two solutions (r^i, η) to (Euler Co-NCL) with initial condition $(r_0^i, \bar{\eta}) \in (L_\mu^2(\mathbb{R}^d), C_b(\mathbb{R}^{2d}_{\diagup}))$. Assume,

$$\lambda := \inf_{t \in [0,T]} \inf_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} V_t(x, x') \eta_t(x, x') \mathrm{d}\mu(x') > 0 \ , \tag{A.0.2}$$

Then,

$$L_\mu^2 d_2(\sigma_t^1, \sigma_t^2) \leqslant e^{-\lambda t} L_\mu^2 d_2(\sigma_0^1, \sigma_0^2) \ .$$

Proof. Let $\sigma_{0,x}^i = \delta_{r_t^i(x)} \in \mathcal{P}_2(\mathbb{R})$ be the disintegration of σ^i with respect to μ for $i = 1, 2$. Then, for μ-a.e. $x \in \mathbb{R}^d$, $\Pi_{0,x} = \delta_{r_t^1(x)} \otimes \delta_{r_t^2(x)}$ is the optimal transference plan between $\sigma_{0,x}^1$ and $\sigma_{0,x}^2$,

so that

$$L_\mu^2 d_2(\sigma_t^1,\sigma_t^2) = \int_{\mathbb{R}^d} d_2^2(\sigma_{t,x}^1,\sigma_{t,x}^2)\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} |r_t^1(x)-r_t^2(x)|^2\mathrm{d}\mu(x)\ .$$

Note that since $r^1, r^2 \in L^\infty([0,T], L_\mu^2(\mathbb{R}^d))$ we have by the Dominated Convergence Theorem

$$\begin{aligned}
\frac{d}{dt}L_\mu^2 d_2(\sigma_t^1,\sigma_t^2) =& \int_{\mathbb{R}^d}\frac{d}{dt}|r_t^1(x)-r_t^2(x)|^2\mathrm{d}\mu(x)\\
=& -2\int_{\mathbb{R}^d}(r_t^1(x)-r_t^2(x))\\
&\times\Bigg(-r_t^1(x)\int_{\mathbb{R}^d\setminus\{x\}}V_t^+(x,x')\eta(x,x')\mathrm{d}\mu(x') + \int_{\mathbb{R}^d} r_t^1(x')V_t^-(x,x')\eta(x,x')\mathrm{d}\mu(x')\\
&+ r_t^2(x)\int_{\mathbb{R}^d\setminus\{x\}}V_t^+(x,x')\eta(x,x')\mathrm{d}\mu(x') - \int_{\mathbb{R}^d} r_t^2(x')V_t^-(x,x')\eta(x,x')\mathrm{d}\mu(x')\Bigg)\\
=& -2\int_{\mathbb{R}^d}(r_t^1(x)-r_t^2(x))^2\int_{\mathbb{R}^d\setminus\{x\}}V_t^+(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
&+2\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\setminus\{x\}}(r_t^1(x)-r_t^2(x))(r_t^1(x')-r_t^2(x'))V_t^-(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
\leqslant& -2\int_{\mathbb{R}^d}(r_t^1(x)-r_t^2(x))^2\int_{\mathbb{R}^d\setminus\{x\}}V_t^+(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
&+\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\setminus\{x\}}(r_t^1(x)-r_t^2(x))^2+(r_t^1(x')-r_t^2(x'))^2V_t^-(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
=& -2\int_{\mathbb{R}^d}(r_t^1(x)-r_t^2(x))^2\int_{\mathbb{R}^d\setminus\{x\}}V_t^+(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
&+\int_{\mathbb{R}^d}(r_t^1(x)-r_t^2(x))^2\int_{\mathbb{R}^d\setminus\{x\}}V_t^-(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
&+\int_{\mathbb{R}^d}(r_t^1(x')-r_t^2(x'))^2\int_{\mathbb{R}^d\setminus\{x\}}V_t^+(x',x)\eta(x',x)\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
=& -\int_{\mathbb{R}^d}(r_t^1(x)-r_t^2(x))^2\int_{\mathbb{R}^d\setminus\{x\}}V_t(x,x')\eta(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\\
\leqslant& -\lambda L_\mu^2 d_2(\sigma_t^1,\sigma_t^2)\ ,
\end{aligned}$$

where in the third equality we have used the antisymmetry of V and the simmetry of η. An application of Grönwall's inequality yields the proof. □

For the reader's convenience, we show the following application of the variation of constants formula to obtain an explicit representation for η where (r,η) is a solution to (Co-NCL). We present it for the particular case of $L_\mu^\infty(\mathbb{R}^d)$ solutions of (Euler Co-NCL), but it holds in the more general setting of Chapter 2.

Lemma A.0.1. Let $(r,\eta) \in AC([0,\infty), L_\mu^\infty(\mathbb{R}^d)) \times AC([0,\infty), C_b(\mathbb{R}^{2d}_{\not\Delta}))$ be the solution of (Euler Co-NCL) with initial datum $(r_0,\eta_0) \in (L_\mu^\infty(\mathbb{R}^d), C_b(\mathbb{R}^{2d}_{\not\Delta}))$. Then, for any $(x,y) \in \mathbb{R}^{2d}_{\not\Delta}$ and $t \geqslant 0$, η satisfies

$$\eta_t(x,y) = e^{-t}\left(\eta_0(x,y) + \int_0^t e^s \omega_s\left[r\right](x,y)\mathrm{d}s\right) \ . \tag{A.0.3}$$

Proof. Define

$$\xi_t := \int_0^t e^s \omega_s[r](x,y)\mathrm{d}s \ ,$$

and note that by the Fundamental theorem of calculus we have

$$\frac{d}{dt}\xi_t = e^t \omega_t[r](x,y) \ . \tag{A.0.4}$$

On the other hand, note that

$$\begin{aligned}
\frac{d}{dt}(e^t\eta_t) &= e^t\eta_t + e^t\partial_t\eta_t && \text{(A.0.5)}\\
&= e^t\eta_t + e^t(\omega_t[r](x,y) - \eta_t(x,y)) && \text{(A.0.6)}\\
&= e^t\omega_t[r](x,y) && \text{(A.0.7)}
\end{aligned}$$

where in the second equality we used that η solves (Co-NCL). Comparing (A.0.4) and (A.0.5) we have

$$\frac{d}{dt}\xi_t = \frac{d}{dt}(e^t\eta_t) \ ,$$

and integrating with respect to time gives the result. □

Next, we present the following result showing the conservation of mass for (Euler Co-NCL).

Lemma A.0.2. Let $(r,\eta) \in AC([0,\infty), L^\infty_\mu(\mathbb{R}^d)) \times AC([0,\infty), C_b(\mathbb{R}^{2d}_{\not\Delta}))$ be the solution of (Euler Co-NCL) with $\Phi \equiv \Phi_{upwind}$ and assume that η_0 is symmetric and ($\boldsymbol{\omega}4$) holds. Then for $t \geqslant 0$

$$\int_{\mathbb{R}^d} r_t(x)\mathrm{d}\mu(x) = \int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) \ .$$

Proof. The symmetry of η_0 and ($\boldsymbol{\omega}4$) imply, by Lemma 4.3.4, that the map $(x,x') \mapsto \eta_t(x,x')$ is symmetric for $t \geqslant 0$. Then, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} r_t(x)\mathrm{d}\mu(x) = &\int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) - \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} r_t(x)(V_t[r](x,x'))_+\eta_t(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\mathrm{d}s\\
&+ \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} r_t(x')(V_t[r](x,x'))_-\eta_t(x,x')\mathrm{d}\mu(x')\mathrm{d}\mu(x)\mathrm{d}s\\
= &\int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) - \int_0^t \iint_{\mathbb{R}^{2d}_{\not\Delta}} r_t(x)(V_t[r](x,x'))_+\eta_t(x,x')\mathrm{d}\mu(x)\mathrm{d}\mu(x')\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&+ \int_0^t \iint\limits_{\mathbb{R}^{2d}} r_t(x)(V_t[r](x',x))_- \eta_t(x',x)\mathrm{d}\mu(x)\mathrm{d}\mu(x')\mathrm{d}s \\
&= \int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) - \int_0^t \iint\limits_{\mathbb{R}^{2d}} r_t(x)(V_t[r](x,x'))_+ \eta_t(x,x')\mathrm{d}\mu(x)\mathrm{d}\mu(x')\mathrm{d}s \\
&\quad + \int_0^t \iint\limits_{\mathbb{R}^{2d}} r_t(x)(V_t[r](x,x'))_+ \eta_t(x,x')\mathrm{d}\mu(x)\mathrm{d}\mu(x')\mathrm{d}s \\
&= \int_{\mathbb{R}^d} r_0(x)\mathrm{d}\mu(x) \ ,
\end{aligned}
$$

where in the second equality we swapped $x \to x'$ in the third term, in the third equality, we used the antisymmetry of the velocity and the symmetry of η, and the final equality follows by an application of Fubini's theorem in the last term. □

Finally, we present the following lemma on the explicit solution of Bernoulli ODEs for the reader's convenienince.

Lemma A.0.3. For $K, M > 0$, the ODE

$$
y'(t) = -Ky^2(t) + My(t) \ , \tag{A.0.8}
$$

with initial condition $y(0) = c$ is a Bernoulli ODE with explicit solution

$$
y(t) = \frac{e^{Mt}Mc}{M + (e^{Mt} - 1)Kc} \ .
$$

Proof. Assume $y(t) \neq 0$, otherwise the solution is trivial. Then, we have

$$
\frac{y'(t)}{y^2(t)} = -K + My^{-1}(t) \ .
$$

Denote by $u(t) := -y^{-1}(t)$, and note that, with this substitution, the equation reads

$$
u'(t) = -K - Mu(t) \ ,
$$

which is a linear ODE in u. Solving the ODE yields

$$
u(t) = -\frac{e^{-Mt}\left(Ke^{Mt} - K - Mu_0\right)}{M} \ .
$$

Substituting back to $y(t)$ gives the result.

□

Part II

Inverse Problems for Aggregation-Diffusion Equations

Notation

We list the notation we shall use throughout Part II.

- $\|\cdot\|$ to denote the Euclidean norm in $\mathbb{R}^d$ or the Frobenius norm when treating matrices.
- $\|\cdot\|_p$ denotes the p norm for a vector. When $p = 0$, it means the number of nonzero entries in a vector.
- Let $B \in \mathbb{R}^{p\times q}$ be a matrix and $B^\top$ its transpose.
- The pseudo-inverse of B is represented as B^+. For an index set $I \subset \{1, \cdots, p\}$, the submatrix of B formed by selecting row indices in I is denoted by B_I and belongs to $\mathbb{R}^{|I|\times q}$, where $|I|$ represents the cardinality of the set I.
- Let $c \in \mathbb{R}^p$ denote a vector, then $c(I) \in \mathbb{R}^{|I|}$ is the restriction of c on I.
- For integers m, n, p, we use a Matlab notation $m : p : n$ to represent the array with values starting at m, augmenting by p, and ending at or before n.
- $W(\boldsymbol{x}) := \Phi(|\boldsymbol{x}|)$ is an interaction potential.
- $\mathcal{W}^{k,p}(\mathbb{R}^d)$ is the Sobolev space with k derivatives in $L^p(\mathbb{R}^d)$.
- $\rho(t, x)\,\mathrm{d}\boldsymbol{x}$ is the density of the solution of (2.1.1).
- For any metric space X, $\mathcal{P}_p(X)$ denotes the space of probability measures with finite p-th moments over X.
- For two measures $\nu,\ \xi \in \mathcal{P}_2(X)$, we denote by $\Pi := \{\gamma \in \mathcal{P}_2(X \times X) \mid {\pi^1}_\#\gamma = \nu,\ {\pi^2}_\#\gamma = \xi\}$, the space of couplings between ν and ξ.
- For $\nu, \xi \in \mathcal{P}_p(X)$ we denote by $d_p(\nu, \xi) := \inf\left\{\int_{X\times X} |x - y|^p \mathrm{d}\gamma(x, y) \mid \gamma \in \Pi(\nu, \xi)\right\}^{1/p}$, the p-Wasserstein distance between ν and ξ for $p \in [1, \infty)$.
- $\phi,\ \Phi$ is the true interaction kernel and potential.
- $\psi,\ \Psi$ is estimated interaction kernel and potential.
- $\mathcal{E}_\infty(\Psi)$ is the error functional, see (2.2.5).
- $\mathcal{E}_{n,M,L}(\Psi)$ is the discretized error functional.
- $\gamma \in C([0, T], X)$ for any metric space X, we will denote the evaluation of the curve at some time $t \in [0, T]$ interchangeably as γ_t and $\gamma(t)$.

Chapter 1

Introduction

1.1 The aggregation-diffusion equation

This part of the thesis is focused on the study of the inverse problem of learning the interaction potential of the following PDE

$$\begin{cases} \partial_t \rho = \nabla \cdot [\rho \nabla (H'(\rho) + V + W * \rho)], \\ \rho(x, 0) = \rho_0(x), \end{cases} \tag{1.1.1}$$

where $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ is a density,

$$H(\rho) := \int_{\mathbb{R}^d} h(\rho(x)) \mathrm{d}x \tag{1.1.2}$$

is an internal energy with first variation H', $V : \mathbb{R}^d \to \mathbb{R}$ is a confinement potential, and $W : \mathbb{R}^d \to \mathbb{R}$ is an interaction potential. Equation (1.1.1) is called the nonlinear nonlocal aggregation-diffusion equation and has been extensively studied, see [GC24, CCY19, BCGC24] and the references therein, although this list is not exhaustive. As we shall discuss, it encompasses a series of important problems as particular cases depending on which of H', W or V appear in (1.1.1) as well as their particular form.

There are at least two ways of interpreting (1.1.1). On the one hand, we can note that (1.1.1) has the structure of a continuity equation $\partial_t \rho + \nabla \cdot (\rho u) = 0$ where the velocity field u is given by the gradient of the first variation of the free energy functional, i.e.

$$u = \nabla \mathcal{F}'(\rho) \ ,$$

where $\mathcal{F}(\rho) := H(\rho) + \mathcal{V}(\rho) + \mathcal{W}(\rho)$,

$$\mathcal{V}(\rho) := \int_{\mathbb{R}^d} V(x)\rho(x)\mathrm{d}x$$

is the potential energy, and

$$\mathcal{W}(\rho) := \frac{1}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} W(x-y)\rho(x)\rho(y)\mathrm{d}x\mathrm{d}y$$

is the interaction energy. Then, under suitable conditions on the continuity of $\mathcal{F}$, it can be shown that (1.1.1) is a gradient flow of $\mathcal{F}$ in $\mathcal{P}_2(\mathbb{R}^d)$ with respect to the 2-Wasserstein [AGS05, San15].

On the other hand, different versions of (1.1.1) can be interpreted as the approximation of the average behavior of a large system of interacting particles moving in space [Oel90, BCnC11]. In this context, the confinement potential represents a preferred direction of movement for the particles, determined, for example, by some property of the medium. Indeed, in the simplest case, when particles do not interact, we can consider the following system of N particles X_t^i with equal mass satisfying

$$\frac{d}{dt}X_t^i = v(X_t^i) \quad \text{for } i = 1,\dots,\ N. \tag{1.1.3}$$

Then, one can consider the *empirical measure* μ^N associated to this system, given by

$$\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}\ ,$$

where δ_x is a dirac-mass at x. Following, e.g. [GC24], we have that, for a test function $\varphi \in C_c^1((0,\infty)\times\mathbb{R}^d)$

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^d}\varphi(t,x)\mathrm{d}\rho_t^N(x) &= \frac{1}{N}\sum_{i=1}^N \frac{d}{dt}\varphi(t,X_t^i)\\ &= \frac{1}{N}\sum_{i=1}^N\left(\partial_t\varphi(t,X_t^i) + \nabla\varphi(t,X_t^i)\frac{dX_t^i}{dt}\right)\\ &= \int_{\mathbb{R}^d}\left(\partial_t\varphi(t,x) + \nabla\varphi(t,x)v(x)\right)\mathrm{d}\rho_t^N(x)\ .\end{aligned} \tag{1.1.4}$$

Integrating in time, after an integration by parts, we note that we recover

$$\partial_t\rho^N + \nabla\cdot(\rho^N v) = 0\ ,$$

in a distributional sense. This means that (1.1.4) is the distributional form of the continuity

equation. Thus when $H = \mathcal{W} = 0$, (1.1.1) is a linear continuity equation.

As its name suggests, the interaction potential captures how agents affect each other. For example, particles can repel or attract each other depending on their proximity. Thus, when only $H = 0$, and we have that the vector field in (1.1.3) is given by $v(X_t^i) = -\frac{1}{N}\sum_{j=1}^N \nabla W(X_t^i - X_t^j) - V(X_t^i)$, then, system (1.1.3) becomes an interacting particle system. Furthermore, if the initial conditions are such that $\rho_0^N \to \rho^0$ as $N \to \infty$ in the 2-Wasserstein distance, then, using stability estimates first introduced by Dobrushin [Dob79], it can be shown that $\rho_t^N \to \rho_t$ in 2-Wasserstein distance as $N \to \infty$. This approximation procedure is called the *mean-field limit*. The name comes from the scaling $1/N$, which is the appropriate scaling to obtain a relevant limit in growing ensembles of exchangeable, or indistinguishable interacting particles. The scaling ensures that each particle experiences an interaction force of order 1. We refer the reader to [Jab14, Gol16] and the references therein for a general overview on the mean-field limit. Hence, when $H = 0$, (1.1.1) becomes the *aggregation-equation* (sometimes also referred to as *nonlocal interaction equation*)

$$\partial_t \rho = \nabla \cdot (\rho \nabla W * \rho) \ , \tag{1.1.5}$$

where W is typically a symmetric interaction potential [CDF+12, Rao12]. As we mentioned before, (1.1.5) has been widely used to give continuum-descriptions of microscopic models in the study of cell populations [BV06, GL95, CHS18, CMS+19], ion channel transport and animal populations as the number of agents goes to infinity [BCnC11, CCH14b, TBL06, KCB+13].

The internal energy determines whether the system exhibits diffusion and, in case it does, whether it is linear or nonlinear. The case of linear diffusion is obtained when $H(\rho) = \kappa\rho(\log\rho - 1)$, where κ is the diffusion constant. In particular, if $\mathcal{V} = \mathcal{W} = 0$ and $H(\rho) = \kappa\rho(\log\rho - 1)$, then we have linear diffusion and (1.1.1) becomes the heat equation. At the level of particles, it is well known that the heat equation corresponds to the evolution of the law of the stochastic differential equation (SDE)

$$\mathrm{d}X_t = \sqrt{2\kappa}\mathrm{d}B_t$$

where $B_t \in \mathbb{R}^d$ is a Brownian motion. If we include an interaction potential W in the previous SDE and consider N particles, we can obtain a class of stochastic IPSs of the form

$$\mathrm{d}X_t^{i,N} = -\nabla W(X_t^{i,N}, \mu^N)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t^i \ , i = 1, \dots, N \ , \tag{1.1.6}$$

where $(B_t^i)_{t\in[0,T]} \in \mathbb{R}^d$ for $i = 1, \dots, N$ are independent Brownian motions and the drift coefficient W depends on all the particles through the empirical measure $\mu^N = \sum_{i=1}^N \frac{1}{n}\delta_{X_t^{i,N}}$. Under suitable conditions, it can be shown that a version of the law of large numbers applies and, as $N \to \infty$,

(1.1.6) converges to

$$\mathrm{d}X_t^i = -\nabla W(X_t^i, \mu) + \sqrt{2\kappa}\mathrm{d}B_t^i \ , \ i = 1, \dots N \ , \tag{1.1.7}$$

in L^2, uniformly in time and in the labels of the particles. In (1.1.7), μ is the law of X_t^i for any $i = 1, \dots, N$. Furthermore, note that the system has decoupled and it is composed by N independent copies of the same SDE. The result showing that the law of the system (1.1.6) converges to the law of (1.1.7) is called *propagation of chaos*. We point the reader to [Szn91, Car16, JW17] for more details.

Using Itô's formula, one can obtain that the law of (1.1.7) solves a version of (1.1.1) with linear diffusion and an interaction potential W. The resulting PDE is called the *McKean-Vlasov* equation [Car16, Lac16, CGPS20, FS21, CGCYZ23, HJ23] and has been applied in finance, neuroscience and machine learning. There are two further important particular cases of the McKean-Vlasov equation. First, when the particle system (1.1.6) represents the phases of coupled oscillators on the torus, $V := -\omega$, the frequency of the system, and the interaction potential is given by $W(x) = \sin(x)$, then (1.1.6) is referred to as the *Kuramoto model* after the work of Kuramoto [Kur75] on systems of coupled oscillators [GC24]. Second, if $\mathcal{W} = 0$, but the confinement potential and $V \neq 0$ does not vanish, then the McKean-Vlasov equation becomes the *Fokker-Planck* equation which has been applied in finance [Sor01, NS11], opinion formation [Tos06, FPTT17, GPY17] among other applications in kinetic theory.

Another important class of models arises when the interaction potential $W = W_N$, where W_N is the so-called *Newtonian potential* given by

$$W_N(x) = \begin{cases} -\frac{1}{2\pi}\log|x| & \text{if } d = 2, \\ \frac{1}{d(d-2)\omega_d}|x|^{2-d} & \text{if } d > 2 \ , \end{cases}$$

where ω_d is a dimension dependent constant. This is the Keller-Segel model of chemotaxis [KS70, CLM14, BDP06, Pat53, CHMV18, CHVY19]. We note that there are versions of the Keller-Segel model with non-linear diffusion, with the typical choice $H(\rho) = \kappa\frac{\rho^m}{m-1}$ for $m > 1$, see for example [CC06].

Nonlinear diffusion typically arises in situations in which the rate of diffusion depends on the concentration of the matter being diffused. When $\mathcal{V} = \mathcal{W} = 0$ and H' is non-decreasing, (1.1.1) becomes the (general) *porous medium equation* [VÓ7, CT00, Ott01], known as the *porous medium equation* when $H(\rho) = \kappa\frac{\rho^m}{m-1}$ for $m > 1$. We conclude by mentioning that for this latter choice of H one can also obtain particle approximations, although the methods become more involved, see for example [DRR22, CEW24].

1.2 The role of the interaction potential

The interaction potential W is chosen to obtain a suitable model for the problem under study. Typically, W is radially symmetric, i.e. $W(x) = w(|x|)$, for some function w, capturing the idea that the interaction between particles mainly depends on their distance. A common choice is $W = |x|^a/a - |x|^b/b$ for $a > b > 0$, see for example [BCP97, CMV03, Tos04]. Furthermore, the sign of the derivative of W characterises whether the potential is attractive, $\frac{d}{dx}W(x) > 0$, or repulsive $\frac{d}{dx}W(x) \leqslant 0$. This is particularly relevant when (1.1.1) is used to model large populations of interacting agents.

For example, studies on fish-schools [BEB+09] or birds [HCH10] suggest that interaction potentials should capture three distinct regions of influence: first a region of *repulsion*, where the agents have a tendency to move away from individuals in their near proximity, next a region of *orientation*, where individuals attempt to move in the same direction as the group and, finally, a region of *attraction* in which the individuals perceive they are too far away from the group and try to reach it [Rey87]. A widely used potential with these characteristics is the attractive-repulsive Morse potential $W(x) = -C_A e^{-|x|/\ell_A} + C_R e^{-|x|/\ell_R}$, where C_A, C_R are the strengths of attraction and repulsion, while ℓ_A/ℓ_R are the lengths of attraction and repulsion, respectively. The relevant choices for the attraction and repulsion strengths and lengths in applications satisfy $C_R/C_A > 1$ and $\ell_R/\ell_A < 1$, which leads to long-range attraction and short-range repulsion, see [LTB09, BT11, CHM14].

The functional form and properties of the interaction potential are also relevant because there are a variety of models based on (1.1.1) that attempt to reproduce different forms of *spontaneous pattern formation* [vBUKB12, BCLR13, ABCvB14, CMS+19]. Also referred to as *self-organization*, this concept refers to the observation of intricate global patterns caused by simple interaction rules [LTB09, BT11, BCLR13, CHM14]. As previously mentioned, equation (1.1.1) arises as the continuum description of kinematic particle models in which inertial effects are disregarded and the particles' velocities are assumed to adjust instantaneously [TB04]. We also note that there are important classes of models, often referred to as *second-order* or *dynamic* models, which incorporate Newton's second law to determine the velocities. This class of models are particularly well suited to capture self-organization phenomena. A prominent example is the Cucker-Smale model introduced in the papers [CS07a, CS07b], see also the review [CFTV10]. Namely, in [CHL17] many variants of the Cucker-Smale model were shown to lead to *flocking* under different assumption, and in [ABCvB14] they were shown to produce the formation of mills, a vortex like motion typically observed in fish schools. Analogously, in the context of opinion dynamics, an important question is whether consensus is reached [Has21, BST22] or whether opinion clusters are formed [JM14].

1.3 The inverse problem for learning interaction potentials

As the preceding discussion illustrates, the interaction potential has far reaching consequences for the properties and the behaviour of equation (1.1.1). Thus, a strand of the literature is concerned with the following problem: is it possible to infer W from observations of trajectories of solutions to (1.1.1)? This inverse problem has received widespread attention and has been approached in, at least, two ways. On the one hand, many works tackle the inverse problem using data from trajectories of the IPS corresponding to the version of (1.1.1) under consideration. Authors following this approach have considered the inverse problem in multiple settings such as particle systems without noise [LZTM19] and with noise [LMT22], second order systems [MTZM23], systems with multiple agent types [LMT21] and more recently, systems on graphs [ASM22, LWLM24]. Furthermore, the problem has been considered in a parametric [Kas90, GCL24] as well as non-parametric setting [YCY22, DMH22]. A common feature in most of these works is the use of a quadratic error functional measuring the discrepancy between the data and the dynamics generated by the candidate interaction potential. Thus, for example, in the case of a first-order system of the form

$$\dot{X}_t^i = \frac{1}{N}\sum_{j=1}^{N} \phi(\|X_t^j - X_t^i\|)(X_t^j - X_t^i) \ , \tag{1.3.1}$$

where $X_t^i \in \mathbb{R}^d$ is the position of particle $i = 1, \dots, N$ at time $t \in [0, T]$ and $\phi = \frac{d}{dx}W$ is the interaction kernel, the error functional can be given by

$$\mathcal{E}(\varphi) = \frac{1}{N}\sum_{i=1}^{N} \left\| \dot{X}_t^i - \frac{1}{N}\sum_{j=1}^{N} \nabla\varphi(\|X_t^j - X_t^i\|)(X_t^j - X_t^i) \right\|^2 \ , \tag{1.3.2}$$

where φ is the candidate interaction potential. Another point in common in most of the previous works is the use radial interaction potentials leading to IPSs of the form (1.3.1). An advantage of these systems is that the interaction depends solely on the pairwise distances between particles. This effectively reduces the dimension of the problem from dN to N and circumvents the curse of dimensionality to obtain learning rates that are useful in practical applications.

On the other hand, researchers have approached the problem by directly using data from solutions of (1.1.1), although this approach has received less attention to the best of the author's knowledge, see for example [HKL+22, LL22, BFHM17, CLP24]. In this setting, the data is given by observations of the solution of (1.1.1) on a discrete space-time mesh, i.e. $\{\rho(t_\ell, x_m)\}_{\ell=1,m=1}^{L,M}$ where L is the number of time steps and M is the size of the spatial grid. As in the particle

system approach, a quadratic error functional such as

$$\tilde{\mathcal{E}}(\varphi) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \|\nabla\varphi * \rho - \nabla W * \rho\|^2 \rho(t,x)\mathrm{d}x\mathrm{d}t \tag{1.3.3}$$

is used, where W is the ground truth interaction potential, φ is the candidate interaction potential and ρ is the solution of (1.1.1). Thus, as in the particle system case, the quadratic functional penalises deviations of the estimator from the ground truth. Although the form of (1.3.3) has a clear interpretation, it is not useful in practical applications because it depends on the unknown ground truth W. However, after some algebra and using (1.1.1) we can obtain the functional

$$\mathcal{E}(\varphi) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\|\nabla\varphi * \rho\|^2 \rho + 2\varphi * \rho\, \partial_t\rho + 2\nabla\varphi * \rho \cdot (\rho\nabla(H'(\rho) + V))\Big]\mathrm{d}x\mathrm{d}t \tag{1.3.4}$$

where $\tilde{\mathcal{E}}(\varphi) = \mathcal{E}(\varphi) + \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \|\nabla W * \varphi\|^2 \rho \, \mathrm{d}x\mathrm{d}t$. In this setting, the problem has been mainly approached in a parametric way, i.e. by minimising (1.3.3) over a finite dimensional hypothesis space $\mathcal{H}$.

A further point in common between the IPS and PDE approaches is the need of a *coercitivity condition* to guarantee the identifiability of the true interaction potential, i.e. a condition that guarantees the uniqueness of the minimizer of (1.3.2) or (1.3.3), see for example [LZTM19, MTZM23, LMT21, TTZ23]. Indeed, the potential ill-posedness of statistical inverse problems is well known [AGS11] and this case is not different. Intuitively, the ill-posedness stems from the difficulty of learning the interaction kernel, since the researcher does not have access to direct observations of ϕ, but only to the averages given in the right hand side of (1.3.1). If one tried to use (1.3.1) to recover ϕ, it becomes apparent that it is underdetermined since the number of known quantities on the left hand side is dN which is always bounded by $N(N-1)/2$, the number of unknown quantities on the right hand side [FKRT24].

To address this issue, most of the aforementioned works introduce different versions of a measure ν of the pairwise distances between the particles. For example, in [LMT21], the authors introduce the measure

$$\nu_T(\mathrm{d}r) := \frac{1}{\binom{N}{2}T}\int_0^T \left[\sum_{i,j=1,i<j}^{N} \mathbb{E}[\delta_{r_{ij}(t)}(\mathrm{d}r)]\right]\mathrm{d}t \ ,$$

where $r_{ij}(t) = \|X_t^i - X_t^j\|$ and the expectation is over the initial distribution and the potential noise in the system in the case of a stochastic IPS (an analogous measure is used in the PDE case). Note that it is natural to assume that ϕ lies in a hypothesis space $\mathcal{H} \subseteq L^2(\nu_T)$. Then, the

coercitivity condition

$$c_{\mathcal{H}}\|\varphi\|^2_{L^2(\nu_T)} \leqslant \frac{1}{2\sigma^2 NT}\int_0^T \sum_{i=1}^N \mathbb{E}\left[\|\frac{1}{N}\sum_{j=1}^N \varphi(r_{ij}(t))r_{ij}(t)\|^2\right] \mathrm{d}t \tag{1.3.5}$$

(see [LMT22, Definition 3.1]) ensures the identifiability of the true interaction potential by using (1.3.2) or (1.3.3), see [LZTM19].

In [LLM+21], it is shown that the coercitivy condition is sufficient for identifiability and it is necessary as the number of particles $N \to \infty$. It is also shown that the coercitivity constant $c_{\mathcal{H}}$ vanishes as the dimension of the hypothesis space $\mathcal{H}$ goes to infinity. This implies that even if the coercitivity condition holds, additional regularization is needed in this case as becomes apparent in [LL23c] where L^2 regularization is used.

1.4 The contributions of this thesis

In Chapter 2 we present our work in [CERMT24] on the inverse problem of learning the interaction potential of (1.1.1). We tackle the problem using data from the solution of (1.1.1)

$$\{\rho(t_\ell, x_m) + \epsilon_m^\ell\}_{m=-M,\ell=1}^{M,L} ,$$

with observation noise $\{\varepsilon_m^\ell\}_{m=-M,\ell=1}^{M,L} \sim N(0,\sigma^2)$, where $\{x_m\}_{m=-M}^{M}$ is a uniform mesh of $\Omega = [-M, M]$ and $t_\ell = \ell\Delta t \subset [0, T]$ for $\ell = 1, \ldots, L$. Due to the ill-posedness of the problem described in the previous section, we propose to regularize it by promoting sparsity, i.e. using an L^1 type regularization. The approach is motivated by the sparsity of many popular interaction potentials when using basis functions belonging to a suitable function class.

Thus, for finite dimensional hypothesis spaces $\mathcal{H} = span\{\Psi_1, \ldots, \Psi_n\}$ with basis functions Ψ_i, we can write our candidate interaction potential, Ψ, as $\Psi = \sum_{i=1}^n c_i\Psi_i$ for some coefficients $c_i \in \mathbb{R}$. In this case, as will become clear later, the error functional in (1.3.4) can be expressed as the quadratic form

$$c^\top Ac - 2b^\top c ,$$

where

$$A(i,j) = \frac{1}{T}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla\Psi_i(y)\cdot\nabla\Psi_j(z)\Big[\int_0^T\int_{\mathbb{R}^d} \rho(t, x-y)\rho(t, x-z)\rho(t, x)\,\mathrm{d}x\,\mathrm{d}t\Big]\,\mathrm{d}y\,\mathrm{d}z , \tag{1.4.1}$$

$$b(i) = -\frac{1}{T}\int_0^T\int_{\mathbb{R}^d}\Big[\partial_t\rho(\Psi_i * \rho) + (\nabla\Psi_i * \rho)\cdot F(\rho, x)\Big]\,\mathrm{d}x\,\mathrm{d}t \tag{1.4.2}$$

for $i, j = 1, \dots, n$. Then we can formulate our regularized learning problem as

$$\begin{aligned} &\text{minimize}_{\, c \in \mathbb{R}^n} \|c\|_1 \ , \\ &\text{subject to } Ac = b \ . \end{aligned} \tag{1.4.3}$$

The above formulation is known as the Basis Pursuit (BP) problem in the literature on compressed sensing [WM22]. Most of the existing algorithms designed to solve the BP problem require the columns of the matrix A to have low correlation and to satisfy the Restricted Isometry Property (RIP), which states that A should act almost as an isometry on sparse vectors [Can08]. This property is closely related to the coercitivity condition described in the previous section.

We solve the BP problem using the PartInv algorithm introduced in [CDN14]. This algorithm is particularly well suited for our purposes because, as we show in our numerical examples, the columns of the matrix A are highly correlated. As we explained in the previous section, A is closely related to averages of the interaction potential weighted by the pairwise distances between particles of an IPS so this is to be expected. The PartInv algorithm is better suited to deal with matrices with highly correlated columns than other methods. Although it performs better than many competitors, the PartInv algorithm may still run into difficulties when identifying the correct support of the coefficient vector c, i.e. it may fail to correctly identify the entries of c that are different from zero. This is more likely when the discretisation errors and the noise σ are large. Therefore, we enhance the PartInv procedure by proposing a support prunning algorithm.

We obtain explicit bounds on the numerical error of computing the matrix A and the vector b in the noiseless and noisy regimes in Proposition 2.4.1 and Proposition 2.4.2 together with Remark 2.4.2, respectively. We summarize the results in the following Proposition

Proposition 1.4.1. Let $A_{n,M,L}$ and $b_{n,M,L}$ be the discretisations of the matrices A and b in (1.4.1). Under suitable assumptions on the regularity of the solution ρ of (1.1.1), $H'(\rho)$, V and the basis functions ψ, the discretisation error of $A_{n,M,L}$ and $b_{n,M,L}$ are bounded by

$$|A(i,j) - A_{n,M,L}(i,j)| \leqslant \alpha(\Delta t + \Delta x) \ , \tag{1.4.4}$$

$$|b(i) - b_{n,M,L}(i)| \leqslant \beta(\Delta t + \Delta x) \ , \tag{1.4.5}$$

for any $i, j = 1, \dots, n$, where α and β are positive constants that depend on the regularity of the previously mentioned functions and Δx, Δt are the sizes of the space and time mesh, respectively. Furthermore, in the presence of noise with volatility σ, we denote by $\tilde{\rho}(t_\ell, x_m) = \rho(t_\ell, x_m) + \varepsilon_m^\ell$ the perturbed solution data, and $\widetilde{A}, \widetilde{b}$ the corresponding versions of the matrices A, b depending

on $\widetilde{\rho}$. We denote by $\|\cdot\|_{L^2(\varepsilon)}$ the natural L^2 norm with respect to the noise. Then, we have

$$\|A - \widetilde{A}_{n,M,L}\|_{L^2(\varepsilon)} \leqslant \alpha n(\Delta t + \Delta x) + nC(\sigma\sqrt{\Delta t \Delta x} + \sigma^2 \Delta x) \ , \tag{1.4.6}$$

where C is a constant depending on the regularity of ψ. Finally, if $H(\rho) = \frac{\rho^2}{2}$, we have

$$\|b - \widetilde{b}_{n,M,L}\|_{L^2(\varepsilon)} \leqslant \beta\sqrt{n}(\Delta t + \Delta x) + \sqrt{n}C\sigma^2(\Delta x^{-1} + \Delta x \Delta t^{-1}) \ . \tag{1.4.7}$$

The previous estimate shows the errors introduced by the presence of noise in the discretisation of the matrix b for a fixed space-time mesh size $\Delta x, \Delta t$. We also obtain novel stability estimates in terms of the 2-Wasserstein distance of the solution to the true system and the learned one in terms of the error functional (1.3.3). The results for this can be found in Proposition 2.3.1, Proposition 2.3.3 and Proposition 2.3.4 for the cases of no diffusion, McKean-Vlasov type dynamics with solution dependent diffusion and non-linear diffusion, respectively. We summarise these results in the following proposition as the spirit of the statement does not vary from case to case, only the methods used in the proof.

Proposition 1.4.2. Let ρ be a smooth solution to (1.1.1) and $\widehat{\rho}$ be another solution of the analogous equation depending on an interaction potential $\widehat{W}$. Then, under suitable conditions on the internal potential H and the regularity of the interaction potentials $W, \widehat{W}$, we have the following stability estimate for any $t \in [0, T]$

$$d_2^2(\rho_t, \widehat{\rho}_t) \leqslant C(t) \left(d_2^2(\widehat{\rho}_0, \rho_0) + 2\tilde{\mathcal{E}}(\widehat{W}) \right) \ ,$$

where d_2 denotes the 2-Wasserstein distance and $C(t) > 0$ is a time-dependent constant.

We also mention that in the case of no diffusion and under suitable assumptions on the hypothesis space for the interaction potential, the previous proposition allows us to obtain a Γ-limit type result for sequences of minimizers of an empirical version of the error functional. We conclude by presenting several numerical examples in one and two dimensions, c.f. Section 2.5.

1.5 Perspectives and future work

As we mentioned in Section 1.3, the investigation of the inverse problem using data from the solution of (1.1.1) has received less attention. Although our numerical examples in Chapter 2 show that the Basis Pursuit approach performs better than competitor methods in multiple

scenarios, it would be interesting to obtain theoretical guarantees for the learning error rate of the interaction potential in this case.

In ongoing work at the time of submission of this thesis, we study the inverse problem of learning both the confinement and interaction potential from McKean-Vlasov equations. We aim to obtain conditions that give theoretical guarantees for the recovery of the confinement potential and interaction potential, separately. We also obtain stability estimates controlling the discrepancy of the ground truth solution and the solution depending on the learned interaction potential in terms of their Kullback-Leibler divergence. We are also working towards extending these results to second-order dynamics.

Finally, an additional direction for future research is to integrate both parts of this thesis, i.e. investigating the inverse problem of learning the velocity field of (Co-NCL) from data of its solutions.

Chapter 2

Sparse identification of nonlocal interaction kernels in nonlinear gradient flow equations via partial inversion

2.1 Introduction

We investigate the estimation of interaction potentials for a broad spectrum of nonlocal equations with gradient flow structure [CCH15, CCY19]. These equations can be written as

$$\begin{cases} \partial_t \mu = \nabla \cdot [\mu \nabla (H'(\mu) + V(x) + W * \mu(x))] \ , \quad x \in \mathbb{R}^d \ , t > 0 \ , \\ \mu(x,0) = \mu_0(x) \ , \end{cases} \tag{2.1.1}$$

where $\mu(t,x) \geqslant 0$ denotes a probability measure; $H(\mu)$ denotes the density of internal energy; $V(x)$ is a confinement potential, and $W(x)$ is an interaction potential governing the nonlocal interaction rules.

Advancements in data acquisition technologies, such as digital imaging [BCC+08] and GPS tracking [NÁBV10, LLEK10, TKI+13], have made it possible to collect density evolution data for large ensembles of particles whose evolution can be represented by different versions of (2.1.1), as discussed in Chapter 1. This leads to an intriguing question: is it possible to deduce the interaction rules from such data? Effective algorithms aligning Equation (2.1.1) with these observational data are essential. We propose a variational approach to estimate the interaction potential from observed solution data, that accounts for both discretization errors and observation

errors, as described by

$$\{\rho(t_\ell, x_m) + \epsilon_m^\ell\}_{m=-M,\ell=1}^{M,L} , \tag{2.1.2}$$

where ρ is the smooth density of μ in the sense explained in Section 2.2; (t_ℓ, x_m) represents a uniform mesh in the domain, and $\{\epsilon_m^\ell\}$ is the discrete added noise. Specifically, the solution is sought through solving a quadratic minimization problem:

$$\widehat{W} \in \arg\min_{\Psi \in \mathcal{H}} \tilde{\mathcal{E}}_\infty(\Psi) ,$$

with

$$\tilde{\mathcal{E}}_\infty(\Psi) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \|\nabla\Psi * \rho - \nabla W * \rho\|^2 \rho(t,x)\mathrm{d}x\,\mathrm{d}t .$$

Here, $W \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$ and Ψ is the candidate interaction potential in the hypothesis space $\mathcal{H}$. Due to the ill-posedness of the inverse problem [LL23a, LMT21, LLM+21, TTZ23], the solutions may not be unique or can not be stably recovered from the noisy data. We propose to regularize the inverse problem by promoting sparsity, motivated by the insight that many interaction potentials are simple functions sparse with respect to certain basis functions.

From an algorithmic perspective, our variational functional is composed of two key elements: a quadratic data fidelity term, that performs interaction force matching, and a sparsity-promoting regularizer. This formulation aligns with addressing a Basis Pursuit (BP) problem [WM22], commonly encountered in the realm of compressed sensing. While numerous state-of-the-art algorithms exist for solving BP problems, finding an algorithm that is specifically tailored and effective for a given setting remains a significant challenge.

Popular sparsity promoting techniques include the Sequentially Thresholded least squares (SINDy) [RBPK17] and variants of LASSO algorithms [KLL21, RABK19], as well as iterative greedy algorithms such as subspace pursuit [HKL+20, HKL+23] and advanced gradient descent algorithms solving L^1 minimization [Sch17]. These methods typically address the inverse problem by posing it as an optimization problem. Frequently, the strong form of a PDE is employed as the data fidelity term in the loss functional within these works. A novel approach utilizing the weak form of the PDE [MB21] has exhibited superior robustness to noise, mitigating its impact when approximating derivatives. While these methods can be applied to a wide specturm of problems, a nontrivial effort is needed to devise effective algorithms tailored to specific types of partial differential equations.

These methods have also been applied to learning problems in particle systems. For instance, [HKL+22] explored the identification of non-local potentials in aggregation equations by addressing a regularized L^1 minimization problem through PDE residuals, employing operator splitting

techniques. In [MB22], the authors employed a weak SINDy approach to discern mean-field overdamped equations from particle-level data. Finally, we note that our identification problem bears significant resemblance to the deconvolution problem [BEW19] encountered in image processing. In the latter, the objective is to recover the image from corrupted data samples, which are the result of convolving the image with a known kernel. In our context, we are concerned with solving a deconvolution problem constrained by a PDE, which introduces unique challenges. For example, the unknown coefficient is nonlinearly dependent on the observational data. Consequently, traditional algorithms from image processing cannot be directly applied, demanding innovative approaches to navigate the complexities introduced by the PDE constraints.

One of our main contributions in this chapter is that we propose the PartInv (Partial Inversion) algorithm to solve the BP problem arising in our context. This algorithm excels at handling highly coherent columns of the regression matrix (where the coherence of a matrix is the maximum absolute correlation of its columns), a phenomenon frequently observed empirically across numerous physical examples, and in particular in the ones considered in Section 2.5. Its effectiveness is further enhanced by incorporating support pruning (see Section 2.4.4), which integrates the residual data loss with time evolution error analysis. We have intensively tested our algorithm on both one and two dimensional examples, and the results demonstrate its remarkable effectiveness and superiority over alternative methodologies. Our work builds upon and extends the findings of [LL22], which primarily focused on aggregation equations with linear diffusion and noise-free solution data in one dimension. We have also made contributions by integrating a distinct regularization technique and by extending our study to more complex scenarios, including those involving nonlinear diffusion terms and noisy data.

On the other hand, we also establish new stability estimates for (2.1.1), controlling the 2-Wasserstein distance between the solution generated using $\widehat{W}$ and the solution generated with W in (2.1.1) in terms of the error functional $\tilde{\mathcal{E}}_\infty$. This analysis, which is new and not present in relevant papers [LL22, BFHM17], shows that our error functional is effective in matching the trajectories obtained by the learning procedure to the ground truth. This stands in contrast with other error functionals used in partial differential equation (PDE) discovery, such as those based on the strong or weak form of the PDEs, where no such interpretability exists. In particular, in the case of no diffusion, we show that the target functional $\tilde{\mathcal{E}}_\infty$ can be interpreted as the Γ-limit of a sequence of analogous error functionals $\tilde{\mathcal{E}}_N$ (see (2.3.4)) which depend on sequences of approximating particle systems. In doing so, we are able to sharpen [BFHM17, Theorem 1.1] by showing that the minimizer of $\tilde{\mathcal{E}}_\infty$ arising as the limit of a sequence of minimizers to $\tilde{\mathcal{E}}_N$ is the interaction potential driving the dynamics of the particle system in the $N \to \infty$ limit.

Finally, we conduct a comprehensive error analysis for the estimators. This analysis builds upon and extends the methodologies outlined in [LL22]. Our extension applies these methods to scenarios that include nonlinear and noisy solution data. This broader approach enables a more versatile application of the estimators, catering to a wider range of real-world conditions where noise and nonlinearity are common challenges.

The rest of this chapter is organised as follows. In Section 2.2, we introduce the notation and the mathematical set up of the inverse problem considered. In Section 2.3, we present the stability estimates in terms of the error functional $\tilde{\mathcal{E}}_\infty$ as well as the Γ-limit result in the case of no diffusion. In Section 2.4, we present the bounds on the numerical discretization errors incurred in the implementations of the solution method. We present numerical examples illustrating our results in Section 2.5. In Section 2.6 we present some conclusions and future perspectives. We include most proofs and auxiliary results in the appendices.

2.2 A regularised variational approach via basis pursuit

In this section, we describe the proposed method to identify the interaction potential from a single set of continuous-time trajectory data. This approach entails addressing a variational problem which is comprised of a data-fidelity term for interaction force approximation and an ℓ^1 regularization term to promote sparsity. While acquiring continuous-time observational data is not feasible in real-world scenarios, the theoretical framework provided here forms the cornerstone for the computational estimators we later propose for discrete data.

2.2.1 The error functional

Let $\mu : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be a solution of the PDE (2.1.1) in which W is the target interaction potential to be learned. For any $t \in [0,T]$, assume $\mu(t,x) = \mu_t(x)$ has a smooth density $\rho : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ with respect to the Lebesgue measure, i.e. $\mathrm{d}\mu_t = \rho(t,x)\mathrm{d}x$, and it decays fast enough as $|x| \to \infty$ for all $t \in [0,T]$. As the Equation (2.1.1) is linear in W, it can be written as

$$\partial_t \rho = \nabla \cdot (\rho \, L_\rho W + F(\rho, x)) \, , \tag{2.2.1}$$

where $L_\rho W := \nabla W * \rho$ and $F(\rho, x) := \rho \nabla (H'(\rho) + V(x))$. Let us assume for simplicity that $W \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$. It is obvious that the target interaction potential satisfies

$$W \in \underset{\Psi \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)}{\arg\min} \tilde{\mathcal{E}}_\infty(\Psi) \, , \tag{2.2.2}$$

with

$$\tilde{\mathcal{E}}_\infty(\Psi) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \|L_\rho \Psi - L_\rho W\|^2 \rho(t,x)\mathrm{d}x\,\mathrm{d}t \ . \tag{2.2.3}$$

By integrating by parts, using the fast decay at infinity of the density and the PDE (2.1.1), we deduce

$$\begin{aligned}\tilde{\mathcal{E}}_\infty(\Psi) &= \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \left[\|L_\rho \Psi\|^2 + \|L_\rho W\|^2 - 2\langle L_\rho \Psi, L_\rho W\rangle\right] \rho(t,x)\mathrm{d}x\,\mathrm{d}t \\ &= \mathcal{E}_\infty(\Psi) + \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \|L_\rho W\|^2 \rho(t,x)\mathrm{d}x\,\mathrm{d}t \ ,\end{aligned}$$

where

$$\begin{aligned}\mathcal{E}_\infty(\Psi) :=& \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\|L_\rho \Psi\|^2 \rho(t,x) + 2(\partial_t \rho - \nabla\cdot F(\rho,x))\Psi * \rho\Big]\mathrm{d}x\,\mathrm{d}t \\ =& \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\|L_\rho \Psi\|^2 \rho(t,x) + 2\Psi * \rho\,\partial_t \rho + 2\nabla\Psi * \rho \cdot F(\rho,x)\Big]\mathrm{d}x\,\mathrm{d}t \ .\end{aligned} \tag{2.2.4}$$

Given a finite dimensional subspace $\mathcal{H} = \mathrm{span}\{\Psi_i\}_{i=1}^n \subset \mathcal{W}^{2,\infty}(\mathbb{R}^d)$, we propose to approximate W by minimizers of the following functional:

$$\begin{aligned}\widehat{W} &\in \arg\min_{\Psi\in\mathcal{H}} \mathcal{E}_\infty(\Psi) \ , \\ \mathcal{E}_\infty(\Psi) :&= \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\|L_\rho \Psi\|^2 \rho(t,x) + 2\Psi * \rho\,\partial_t \rho + 2\nabla\Psi * \rho \cdot F(\rho,x)\Big]\mathrm{d}x\,\mathrm{d}t \ ,\end{aligned} \tag{2.2.5}$$

where $\widehat{W}$ is the identified potential by our method. We note that the error functional $\mathcal{E}_\infty$ promotes the matching of the interaction force with the ground truth. In fact, the previous computation shows that

$$\arg\min_{\Psi\in\mathcal{H}} \mathcal{E}_\infty(\Psi) = \arg\min_{\Psi\in\mathcal{H}} \tilde{\mathcal{E}}_\infty(\Psi) \ . \tag{2.2.6}$$

In Section 2.3 we present stability estimates showing that the 2-Wasserstein distance between solutions of (2.1.1) corresponding to the ground truth interaction potential and the learned one can be controlled by the functional $\widetilde{\mathcal{E}}_\infty$. Namely, we present results of the following type.

Proposition 2.2.1. Let $\mu, \widehat{\mu} \in C([0,T], \mathcal{P}^2(\mathbb{R}^d))$ be solutions of (2.1.1) with the interaction potential and external potential $(W,V), (\widehat{W},\widehat{V})$ respectively. Then, under suitable regularity conditions on the velocity fields of μ and $\widehat{\mu}$, we have the following stability estimate

$$d_2^2(\mu(t),\widehat{\mu}(t)) \leqslant C\left(\tilde{\mathcal{E}}_\infty(\widehat{W}) + \int_0^t \|\nabla V - \nabla\widehat{V}\|^2_{L^2(\mu_s)}\mathrm{d}s + d_2^2(\mu(0),\widehat{\mu}(0))\right) \ ,$$

where $C>0$ is a suitable constant depending only on the final time T and the Lipschitz constant of the functions involved, d_2 denotes the 2-Wasserstein distance and $\|f\|^2_{L^2(\mu_t)} := \int_{\mathbb{R}^d} f^2(x)\mathrm{d}\mu_t(x)$.

In our learning problem, we have $\widehat{V} = V$ and assume W is the only unknown. The statement in Proposition 2.2.1 shows that the minimization of $\widetilde{\mathcal{E}}_\infty$ by (2.2.6) leads to minimization in the difference between trajectories of solution to (2.1.1) corresponding to the ground truth interaction potential W and the estimated interaction potential $\widehat{W}$. In particular, whenever we have an estimator such that $\tilde{\mathcal{E}}_\infty(\widehat{W}) = 0$, then $\hat{\mu} = \mu$ on $[0, T]$. However, the error functional $\tilde{\mathcal{E}}_\infty$ depends on the *unknown* ground truth through the term $L_\rho W$, so it is not feasible in any practical computational scheme. Minimizing $\mathcal{E}_\infty$ enables practical implementation using only the data $\rho(t, x)$.

2.2.2 The vanilla least squares solutions

In this study, our focus will be on identifying radial interaction potentials W, but our computational framework can be extended to general potentials as well. From now on, we will always restrict to this class of potentials.

Let us assume that $\mathcal{H}$ is a linear subspace generated by the basis of radial functions $\mathrm{span}\{\Psi_i\}_{i=1}^n$. Let us introduce the notation $\nabla\Psi_i(x) = \psi_i(|x|)\frac{x}{|x|}$. Since $\mathcal{E}_\infty(\cdot)$ is a quadratic functional, then we can rewrite the minimization problem by means of a simple matrix representation. To simplify the notation, we omit the time dependence of the solution ρ in the rest of the section. We first introduce the following bilinear form

$$\begin{aligned}
\langle \Psi_i, \Psi_j \rangle_G &= \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \langle L_\rho \Psi_i, L_\rho \Psi_j \rangle \rho \,\mathrm{d}x\,\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla\Psi_i(y)\cdot\nabla\Psi_j(z)\rho(t,x-y)\rho(t,x-z)\,\mathrm{d}y\,\mathrm{d}z \Big] \rho \,\mathrm{d}x\,\mathrm{d}t \\
&= \frac{1}{T}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \nabla\Psi_i(y)\cdot\nabla\Psi_j(z) \Big[\int_0^T \int_{\mathbb{R}^d} \rho(t,x-y)\rho(t,x-z)\rho(t,x)\,\mathrm{d}x\,\mathrm{d}t \Big] \mathrm{d}y\,\mathrm{d}z \\
&= \frac{1}{T}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \psi_i(|y|)\psi_j(|z|) I_G(y,z) G(y,z)\,\mathrm{d}y\,\mathrm{d}z \ , \qquad (2.2.7)
\end{aligned}$$

where $I_G(y,z) = \frac{y}{|y|}\cdot\frac{z}{|z|}$ with

$$G(y,z) = \int_0^T \int_{\mathbb{R}^d} \rho(t,x-y)\rho(t,x-z)\rho(t,x)\,\mathrm{d}x\,\mathrm{d}t \ , \qquad (2.2.8)$$

for all $y, z \in \mathbb{R}^d$. Let $\Psi = \sum_{i=1}^n c_i\Psi_i$, for $i = 1, \cdots, n$, and define

$$A_{ij} = \langle \Psi_i, \Psi_j \rangle_G \ , \qquad (2.2.9)$$

$$b_i = -\frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \Big[\partial_t\rho(\Psi_i * \rho) + (\nabla\Psi_i * \rho)\cdot F(\rho, x) \Big] \mathrm{d}x\,\mathrm{d}t \ . \qquad (2.2.10)$$

Notice that $b_i = \langle \Psi_i, W \rangle_G$ when we assume that ρ is the exact solution to (2.1.1). Then we can write the error functional as

$$\mathcal{E}_\infty(c) = c^\top A c - 2b^\top c \ . \tag{2.2.11}$$

By first-order optimality, the optimal solutions satisfy the normal equation

$$Ac = b \ . \tag{2.2.12}$$

In the context of inverse problems, the system introduced in (2.2.12) often displays inherent challenges due to its ill-posed nature. Specifically, even in the scenario where W is an element of the Hilbert space $\mathcal{H}$, uniqueness of the solution is not guaranteed. Moreover, the task of matrix inversion, particularly of A, is fraught with numerical instabilities, often enhanced by perturbations such as discrete-time data and observational noise. For an extended discussion on this topic, the reader may consult [LL23a].

Identifying effective regularization methods to stabilize the recovery process remains a critical challenge. Our numerical studies reveal that the regularized least squares estimator, derived using the pseudoinverse, did not perform satisfactorily. Lang et al. [LL22] explored a Tikhonov regularization for aggregation equations with linear diffusion. In this case, the standard least squares estimators, as per Equation (2.2.12), works as maximum likelihood estimators. However, this property does not extend to cases of nonlinear diffusion, necessitating the exploration of alternative regularization approaches.

2.2.3 ℓ^1 regularization via Basis Pursuit

We note that in many prototypical examples, the true interaction kernels are typically simple functions and are often sparse with respect to a set of given basis functions, such as polynomials. Building upon this prior knowledge, we propose estimating the interaction kernel by tackling the following BP problem:

$$\begin{aligned} &\text{minimize}_{\ c\in\mathbb{R}^n} \|c\|_1 \ , \\ &\text{subject to } Ac = b \ . \end{aligned} \tag{2.2.13}$$

This approach seeks to minimize the ℓ_1-norm of the coefficient vector c within the real vector space $\mathbb{R}^n$, subject to the constraint that the product of matrix A and vector c equals the vector b.

BP problems have been actively studied in the area of compressive sensing [WM22]. Many state-of-the-art algorithms such as the CoSaMP algorithm [NT09] and the closely related subspace pursuit algorithm [DM09] are designed to address the BP problem. These iterative

greedy algorithms are particularly acclaimed for their superior recovery in the noisy data regime, combined with their rapid computational efficiency. However, their guaranteed performance often depends on certain properties of the sensing matrix A. In particular, the coherence of A should be small enough (ideally, smaller than $\frac{1}{2\|c\|_0} \leqslant \frac{1}{2}$), meaning that the maximal correlation between the normalized columns of A is small (see Proposition 3.2 in [WM22]) and A should act almost as an isometry on the set of sparse vectors, a property known as the Restricted Isometry Property (RIP) [CRT06, CT06].

Such characteristics are typically satisfied by random matrices. In our context, the sensing matrix A is determined by the inherent physics of the problem and, consequently, is non-random. Our numerical experiments revealed that in all examples, the matrix A possesses highly coherent columns, so that the incoherence parameter is very close to 1 (see Figure 2.6). Consequently, it fails to satisfy the desired RIP. Interestingly, similar challenges with the sensing matrix have been observed in super-resolution problems in imaging. How to perform sparse recovery with coherent sensing matrix is still an on-going challenge in the signal and image processing community. Only a few works focus on addressing this issue and the algorithms are heuristic and short on theoretical justifications [CENR11, FL12, CDN14].

We propose to use the PartInv Algorithm 1, a modification of the CoSaMP algorithm [NT09] to solve the BP with a coherent sensing matrix. PartInv was originally proposed in [CDN14] and showed better performance than existing greedy methods for random matrices, and is especially suitable for matrices that have subsets of highly correlated columns. Compared with CoSaMP, the only difference lies in line 3 of Algorithm 1, where $A^{\top}_{I^{(k)}}$ is replaced by the pseudo-inverse $A^{+}_{I^{(k)}}$, where, for an index set $I \subset \{1, \cdots, n\}$, the submatrix of B formed by selecting row indices in I is denoted by B_I and belongs to $\mathbb{R}^{|I| \times n}$, where $|I|$ represents the cardinality of the set I. This step can reduce the error propagation due to the coherent columns and one can refer to [CDN14] for more details. Moreover, it enjoys partial theoretical justification. More precisely, [CDN14, Theorem 3.1] provides a sufficient condition that uses a weaker condition than RIP and incoherence bounds to prove the success of the algorithm on sparse recovery. In what follows we call a vector $v \in \mathbb{R}^n$ s-sparse, for $s = 1, \ldots, n$, if it has s entries equal to 0 and $I_c \subseteq \{1, \ldots, n\}$ denotes the complement of the index set I.

2.3 Stability estimates and Γ-convergence

As we anticipated in the Introduction, in this section we present stability estimates for the 2-Wasserstein distance between solutions of (2.1.1) depending on the ground truth interaction potential and an interaction potential estimated with techniques such as Basis Pursuit (see

Algorithm 1 Given $Ac = b$ where the ground truth is s-sparse, return the best K-sparse approximation $\hat{c}$.

Require: A, b, K (an upper bound on sparsity s)

$\tilde{c} \leftarrow A^\top b$; $I^{(0)} \leftarrow$ indices of the K-largest magnitudes of $\tilde{c}$; $k \leftarrow 0$
while tolerance threshold not met **do**
  $\tilde{c}_{I^{(k)}} \leftarrow A^+_{I^{(k)}} b$
  $r \leftarrow b - A_{I^{(k)}} \tilde{c}_{I^{(k)}}$
  $J^{(k)} \leftarrow I^{(k)}_c$
  $\tilde{c}_{J^{(k)}} \leftarrow A^\top_{J^{(k)}} r$
  $I^{(k+1)} \leftarrow$ indices of K-largest magnitude components of $\tilde{c}$
  $k \leftarrow k + 1$
end while
Return $\hat{c} = A^+_{I^{(k)}} b$

Section 2.2). We present results for the cases considering: no diffusion, Mc-Kean Vlasov dynamics with solution-dependent diffusion and general nonlinear diffusion. Although the results are analogous, the techniques involved in the proofs differ in each case.

2.3.1 No diffusion, the aggregation equation case

Let us begin by considering the following interacting particle system

$$\dot{X}^i_t = -\frac{1}{N} \sum_{i \neq j} \nabla W(X^i_t - X^j_t) - \nabla V(X^i_t) \ , \quad i = 1, \dots, N \ , \tag{2.3.1}$$

for particles $(X^i)_{i=1}^N \in C([0,T], \mathbb{R}^d)$, an interaction potential $W \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$ and $V \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$, a confinement potential with supp $W \subset \Omega$ and supp $V \subset \Omega$ for a compact set $\Omega \subset \mathbb{R}^d$. Note that (2.3.1) is an ODE system driven by a velocity field analogous to the one in (2.1.1), where we have set $H \equiv 0$, i.e. there is no diffusion.

Under our assumptions on the interaction potential W and the confinement potential V, the system (2.3.1) is well posed by traditional Cauchy-Lipschitz results for ODEs. It can be shown that, as the number of particles $N \to \infty$, the sequence of empirical measures $\mu^N_t := \frac{1}{N} \sum_{i=1}^N \delta_{x_i(t)}$ of the solutions to (2.3.1) converges in the 2-Wasserstein distance to a probability measure $\mu_t \in \mathcal{P}_2(\mathbb{R}^d)$, where $\mathcal{P}_2(\mathbb{R}^d)$ denotes the space of probability measures with finite second moments for any $t \in [0,T]$. In turn, the curve $\mu \in C([0,T], \mathcal{P}_2(\mathbb{R}^d))$ solves the following PDE giving a continuum description of the system (2.3.1)

$$\partial_t \mu = \nabla \cdot (\mu(\nabla W * \mu + \nabla V)) \ . \tag{2.3.2}$$

We refer the reader to [Gol16] for more details. This is the mean-field PDE associated to the

dynamical system (2.3.1) and it is also referred to as the *aggregation equation*. We recall that for a curve $\nu \in C([0,T],\mathcal{P}_2(\mathbb{R}^d))$, we denote the L^2-norm with respect to this curve as

$$\|f\|^2_{L^2(\nu_t)} := \int_{\mathbb{R}^d} |f(x)|^2 \mathrm{d}\nu_t(x) \ ,$$

and by $d_2 : \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ the 2-Wasserstein distance defined as

$$d_2(\mu,\xi) := \min\left\{ \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \mathrm{d}\gamma(x,y) : \gamma \in \Pi(\mu,\xi) \right\}^{\frac{1}{2}} \ ,$$

where $\mu, \xi \in \mathcal{P}_2(\mathbb{R}^d)$, $\Pi(\mu,\xi) := \{\gamma \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \mid \pi^1_\# \gamma = \mu,\ \pi^2_\# \gamma = \xi\}$ is the set of transference plans between the measures μ and ξ and, for any $\nu \in \mathcal{P}(\mathbb{R}^d)$, and measurable set $A \subseteq \mathbb{R}^d$, $\pi^1_\# \nu := \nu((\pi^1)^{-1}(A))$ is the push-forward measure by the projection map to the first coordinate given by $\pi^1(x,y) = x$ and analogously for y. In what follows, we will consider W to be the ground truth interaction potential and we will write the error functional $\tilde{\mathcal{E}}_\infty$ in (2.2.6) in terms of a curve of measures $\mu \in C([0,T],\mathcal{P}_2(\mathbb{R}^d))$ solving (2.3.2) as

$$\tilde{\mathcal{E}}_\infty(\widehat{W}) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} |\nabla W * \mu_t - \nabla \widehat{W} * \widehat{\mu}_t|^2 \mathrm{d}\mu_t(x)\,\mathrm{d}t \ , \tag{2.3.3}$$

for any $\widehat{W} \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$. Now we are ready to present our first Dobrushin-type stability result.

Proposition 2.3.1. Let $W, \widehat{W}, V$, and $\widehat{V}$ belong to $\mathcal{W}^{2,\infty}(\mathbb{R}^d)$ with supp W, supp $V \subseteq \Omega$, where $\Omega \subset \mathbb{R}^d$ is a compact set. For initial data $\mu_0,\ \widehat{\mu}_0 \in \mathcal{P}_2(\mathbb{R}^d)$, let $\mu, \widehat{\mu} \in C([0,T],\mathcal{P}_2(\mathbb{R}^d))$ be solutions to the aggregation equation (2.3.2) with velocity fields $\nabla W * \mu + \nabla V,\ \nabla \widehat{W} * \widehat{\mu} + \nabla \widehat{V}$, respectively. Then, for any $t \in [0,T]$, we have

$$d_2^2(\mu_t,\widehat{\mu}_t) \leqslant C_1 \tilde{\mathcal{E}}_\infty(\widehat{W}) + C_2 \int_0^t \|\nabla V - \nabla \widehat{V}\|^2_{L^2(\mu_s)} \mathrm{d}s + C_3 d_2^2(\mu_0,\widehat{\mu}_0) \ ,$$

where C_1, C_2 and C_3 are non-negative constants depending on T and $L_W, L_{\widehat{W}}, L_V, L_{\widehat{V}}$, the Lipschitz constants of ∇W, $\nabla \widehat{W}$, ∇V and $\nabla \widehat{V}$, respectively.

Proof. We refer the reader to Appendix A.0.1 for the proof of Proposition 2.3.1. □

Remark 2.3.1. Note that the estimate in the previous proposition allows both the interaction and the confinement potentials to differ in (2.3.2). Thus, with minor modifications to the error functional $\tilde{\mathcal{E}}_\infty$ to account for the difference between V and $\widehat{V}$, we could control the 2-Wasserstein distance between μ and $\widehat{\mu}$ in terms of that new error functional and the difference in the initial data. This would be particularly relevant if the goal was to infer both the interaction as well as

the confinement potential from trajectories of the PDE. We leave this for future work and, in what follows, the reader can set $V = \widehat{V}$ in the result above.

Remark 2.3.2. We note furthermore that an analogous result holds in which the L^2 norms depend on the estimated solution $\widehat{\mu}$, instead of the ground truth solution μ. Thus, we could rewrite the estimate in Proposition 2.3.1 to depend on the minimum of the two norms in each case, but opted for the present specification for the sake of clarity.

2.3.1.1 Mean-field dynamics and Γ-convergence

In what follows, without loss of generality, we will assume $V \equiv 0$. Using the stability estimate from Proposition 2.3.1, in this section we present a Γ-limit result as well as the sharpening of [BFHM17, Theorem 1.1]. We remark that in [BFHM17] the setting is slightly different to ours, since the error functional considered by the authors depends on the interaction kernel ∇W, instead of the potential W. For the reader's convenience, we begin by recalling the notation in [BFHM17] adapted to our setting. For a compact set $K \subset \mathbb{R}^d$ let

$$X_{M,K} := \{b \in \mathcal{W}^{2,\infty}(K) : \|b\|_\infty + \|\nabla b\|_\infty + \|\nabla^2 b\|_\infty \leqslant M\} \ ,$$

and $(A^N)_{N\in\mathbb{N}}$ be a family of closed subsets of $X_{M,K}$ with the uniform approximating property, i.e. for any $b \in X_{M,K}$ there exists a sequence $(b^N)_{N\in\mathbb{N}}$ converging uniformly to b on K, such that $b^N \in A^N$ for every $N \in \mathbb{N}$. The authors considered a sequence $(\widehat{W}^N)_{N\in\mathbb{N}} \in A^N$ of minimizers of the following functional

$$\tilde{\mathcal{E}}_N(\widetilde{W}) := \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \left|(\nabla\widetilde{W} - \nabla W) * \mu_s^N(x)\right|^2 \mathrm{d}\mu_s^N(x)\mathrm{d}s \ , \tag{2.3.4}$$

where, as before, $W \in X_{M,K}$ is the true interaction potential of a system like (2.3.1) and $\mu^N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}$ is the empirical measure associated to (2.3.1) with estimated interaction potential $\widehat{W}^N$. In [BFHM17, Theorem 1.1], it is shown that if $(\widehat{W}^N)_{N\in\mathbb{N}} \in A^N$ is a sequence of minimizers of $\tilde{\mathcal{E}}_N$, this sequence has a uniformly converging subsequence to a function $\overline{W} \in X_{M,K}$. Furthermore, $\overline{W}$ is a minimizer of the limiting functional $\tilde{\mathcal{E}}_\infty(\widetilde{W})$ in (2.3.3) where μ_t is the solution of the mean-field PDE arising as the limit of the sequence $(\mu_t^N)_{N\in\mathbb{N}}$ in the d_2 distance.

Using the estimate from Proposition 2.3.1, we will improve the result in [BFHM17, Theorem 1.1] by additionally showing that the mean-field limit of the interacting particle system

$$\dot{X}_t^i = -\frac{1}{N}\sum_{i\neq j} \nabla\widehat{W}^N(X_t^i - X_t^j) \ , \tag{2.3.5}$$

is given by

$$\partial_t \widehat{\mu} = \nabla \cdot (\widehat{\mu}(\nabla \overline{W} * \widehat{\mu})) \ ,$$

i.e. the velocity field of the limiting PDE depends on $\overline{W}$, the limit of the minimizing sequence of the functional (2.3.4). Then, under the additional *coercivity condition*, i.e. that there exists a constant $c_T > 0$ such that

$$c_T \frac{1}{T} \int_0^T \|(\nabla \overline{W} - \nabla W) * \mu_t(x)\|^2_{L^2(\mu_t)} \mathrm{d}t \leqslant \tilde{\mathcal{E}}_\infty(\overline{W})$$

we get that $W = \overline{W}$ in $L^2(\mu)$ and thus $\widehat{\mu} = \mu$ in $(\mathcal{P}_2(\mathbb{R}^d), d_2)$, for any $t \in [0, T]$. We make these remarks precise in the following proposition.

Proposition 2.3.2. Let $W \in X_{M,K}$ be the true interaction potential governing the particle system (2.3.1) and $(A^N)_{N\in\mathbb{N}} \subset X_{M,K}$ be a family with the uniform approximating property. Consider a sequence of minimizers $(\widehat{W}^N)_{N\in\mathbb{N}} \in (A^N)_{N\in N}$ of the functional (2.3.4) with limit $\overline{W} \in X_{M,K}$. Let $\widehat{\mu}(0) \in \mathcal{P}_2(\mathbb{R}^d)$ with compact support be given, and $(\widehat{\mu}^N(0))_{N\in\mathbb{N}}$ be a sequence of empirical measures

$$\widehat{\mu}^N(0) = \frac{1}{N} \sum_{i=1}^N \delta_{X_0^i} \ , i = 1, \dots, N \ ,$$

such that $\lim_{N\to\infty} d_2(\widehat{\mu}^N(0), \widehat{\mu}(0)) = 0$. Let $(X_t^i)_{i=1}^N \in \mathbb{R}^d$ be the unique solution to the particle system

$$\dot{X}_t^i = -\frac{1}{N} \sum_{i\neq j} \nabla \widehat{W}^N(X_t^i - X_t^j), \quad X_0^i = x_{0,i} \quad \text{for } i = 1, \dots, N \ . \tag{2.3.6}$$

Then, the mean-field limit of the system (2.3.6) is given by

$$\partial_t \widehat{\mu} = \nabla \cdot (\widehat{\mu}(\nabla \overline{W} * \widehat{\mu})) \ .$$

Furthermore, if the coercitivity condition holds and the true system has initial condition $\mu(0) = \widehat{\mu}(0) \in \mathcal{P}_2(\mathbb{R}^d)$, we have that $W = \overline{W}$ in $L^2(\mu)$.

Proof. By Remark 2.3.2, we can apply our stability estimate from Proposition 2.3.1 with each norm depending on $\widehat{\mu}^N$, thus obtaining

$$d_2^2(\widehat{\mu}_t^N, \widehat{\mu}_t) \leqslant C_1 \tilde{\mathcal{E}}_N(\widehat{W}^N) + C_2 d_2^2(\widehat{\mu}_0^N, \widehat{\mu}_0) \ ,$$

where

$$\partial_t \widehat{\mu} = \nabla \cdot (\widehat{\mu}(\nabla \overline{W} * \widehat{\mu})) \ .$$

Note that we have the following bound for the first term for any $t \in [0,T]$

$$\tilde{\mathcal{E}}_N(\widehat{W}^N) = \frac{1}{T}\int_0^T \|\nabla \overline{W} * \widehat{\mu}_t^N - \nabla \widehat{W}^N * \widehat{\mu}_t^N\|^2_{L^2(\widehat{\mu}_t^N)} \mathrm{d}t \leqslant \|\nabla \overline{W} - \nabla \widehat{W}^N\|^2_\infty \ .$$

Then, by our assumption on the initial conditions and [BFHM17, Theorem 1.1] we have that

$$\lim_{N\to\infty} d_2^2(\widehat{\mu}_t^N, \widehat{\mu}_t) = 0 \ ,$$

which gives the first part of our statement. Finally, by the coercitivity condition we have that

$$c_T \frac{1}{T}\int_0^T \|(\nabla \overline{W} - \nabla W) * \mu_t(x)\|^2_{L^2(\mu_t)} \mathrm{d}t \leqslant \tilde{\mathcal{E}}_\infty(\overline{W}) \ ,$$

and, since $\overline{W}$ is a minimizer of $\tilde{\mathcal{E}}_\infty$, we can conclude that $W = \overline{W}$ in $L^2(\mu_t)$. □

Remark 2.3.3. The result is not generalized to the cases with diffusion because the method of proof of [BFHM17, Theorem 1.1] requires the solution of (2.1.1) to have compact support. This is not guaranteed if diffusion is present. The extension for the cases of linear or nonlinear diffusion are left for future work.

2.3.2 McKean-Vlasov dynamics with solution dependent diffusion

Next, we consider an extension of the stability estimate of Proposition 2.3.1 when (2.1.1) is a McKean-Vlasov equation with solution dependent diffusion coefficient. Namely, let $\mu \in C([0,T], \mathcal{P}_2(\mathbb{R}^d))$ be the weak solution of the following equation

$$\partial_t \mu = \nabla \cdot (\mu(\nabla W * \mu)) + \sum_{i,j=1}^d \partial^2_{x_i,x_j}[\sigma(K*\mu)^\top \sigma(K*\mu)\mu] \ , \tag{2.3.7}$$

where, as before, W is the interaction potential, σ is the diffusion coefficient which is allowed to depend on the solution μ through its convolution with a kernel K, and the superscript $\top$ denotes the transpose of a matrix as before. We recall that (2.3.7) can be interpreted as the evolution of the law of the solution of the following stochastic differential equation (SDE) [Szn91, Car16]

$$\begin{aligned} dX_t &= \nabla W * \mu_t(X_t)dt + \sqrt{2}\sigma(K*\mu_t(X_t))dB_t \ , \\ X_0 &= X^0 \in L^2 \text{ independent of } (B_t)_{t\in[0,T]} \ , \end{aligned} \tag{2.3.8}$$

where $(B_t)_{t\in[0,T]} \in \mathbb{R}^d$ is a Brownian motion. We present now our stability estimate for this case.

Proposition 2.3.3. Let $\mu, \widehat{\mu} \in C([0,T], \mathcal{P}_2(\mathbb{R}^d))$ be weak solutions to (2.3.7) with coefficients

$\nabla W, \sigma(K)$ and $\nabla\widehat{W}, \hat{\sigma}(\widehat{K})$, where $W \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$, and $\sigma : \mathbb{R}^n \to \mathbb{R}^d \times \mathbb{R}^d$ and $K : \mathbb{R}^d \to \mathbb{R}^n$ are Lipschitz and bounded. Furthermore, assume that for any $\nu \in \mathcal{P}(\mathbb{R}^d)$, $(\nabla W * \nu(X_t))_{t\in[0,T]}$ and $(\sigma(K * \nu(X_t)))_{t\in[0,T]}$ are progressively measurable and square integrable. Let X_t denote the solution to (2.3.8) and $\widehat{X}_t$ denote the solution to an analogous SDE with coefficients, $\widehat{W}$, $\hat{\sigma}$ and $\widehat{K}$. Furthermore, assume that the initial data X_0 and $\widehat{X_0}$ are chosen such that $d_2^2(\mu(0), \hat{\mu}(0)) = \mathbb{E}|X_0 - \widehat{X}_0|^2$. Then we have the following stability estimate

$$d_2^2(\mu_t, \hat{\mu}_t) \leqslant C(T)\left(d_2^2(\mu_0, \hat{\mu}_0) + \tilde{\mathcal{E}}_\infty(\widehat{W}) + \|\sigma - \hat{\sigma}\|_\infty^2 + \int_0^t \|(K - \widehat{K}) * \mu_s\|^2_{L^2(\mu_s)}\mathrm{d}s\right) ,$$

where $C(T)$ is a non-negative constant depending on T and the Lipschitz constants of $K, \widehat{K}, \sigma, \hat{\sigma}$ and W.

Proof. We refer the reader to Appendix A.0.2 for a proof of this proposition. □

Remark 2.3.4. In a similar way to Proposition 2.3.1, we note that our estimate in Proposition 2.3.3 allows the interaction potential, the diffusion coefficient and the kernels K and $\widehat{K}$ to differ between the equations being compared. As before, this would be particularly relevant in a situation where not only the interaction potential, but also the diffusion coefficient as well as the kernel K have to be inferred. Note that the bound in the previous proposition depends on the uniform norm of the difference between σ and $\hat{\sigma}$ suggesting that deeper modifications of the error functional would be required to allow for inference of these functions in an L^2 framework. As before, the reader can think of $K = \widehat{K}$ and $\sigma = \hat{\sigma}$.

2.3.3 Nonlinear diffusion

In this section, we obtain a similar type of stability estimate for an aggregation-diffusion equation with nonlinear diffusion coefficient. This now corresponds to the full equation (2.1.1). Throughout this section, we will assume $\mu(t, x) = \rho(t, x)\mathrm{d}x$. Thus, we will consider the following Cauchy-problem for a curve of probability densities $\rho \in C([0, T], \mathcal{P}_2(\mathbb{R}^d))$

$$\begin{aligned} \partial_t\rho + \nabla \cdot (\rho v(\rho)) = 0 \ , \\ \rho(0) = \rho^0 \in \mathcal{P}_2(\mathbb{R}^d) \ , \end{aligned} \tag{2.3.9}$$

where $v : \mathbb{R}^d \to \mathbb{R}$ is the following velocity field

$$v(\rho) := -\nabla(H'(\rho) + W * \rho + V) \ . \tag{2.3.10}$$

Here, $H : [0, +\infty] \to \mathbb{R}$ is the internal energy density given by $H(z) = \kappa\frac{z^m}{m-1}$ where $m \neq 1$,

$m \geqslant 1 - \frac{1}{d}$ and $m > \frac{d}{d+2}$, $W \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$ is an interaction potential and $V \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$ is a confinement potential. Following [Ott01, Proposition 1], for some $\Omega \subset \mathbb{R}^d$ convex with $\partial\Omega$ smooth we will consider two smooth, non-negative solutions $\rho, \widehat{\rho} : [0,T] \times \Omega \to \mathbb{R}^d$, to the following problem

$$\begin{aligned} \frac{\partial \rho}{\partial \tau} + \nabla \cdot (\rho v(\rho)) = 0 \quad \text{in } [0,T] \times \Omega \\ \rho v(\rho) \cdot \nu = 0 \quad \text{on } [0,T] \times \partial\Omega \end{aligned} \tag{2.3.11}$$

where $v : \mathbb{R}^d \to \mathbb{R}$ is the velocity field in (2.3.10) and $\widehat{\rho}$ satisfies an analogous problem for the velocity field $\widehat{v}(\widehat{\rho}) = -\nabla(H'(\widehat{\rho}) + \widehat{W} * \widehat{\rho} + \widehat{V})$. As before, W is the ground truth interaction potential and $\widehat{W}$ the learned interaction potential. We are now ready to present our stability estimate for the nonlinear diffusion case

Proposition 2.3.4. Let ρ be a smooth solution to (2.3.11) with velocity field (2.3.10), and $\widehat{\rho}$ be another solution of the analogous equation driven by the velocity field $\widehat{v}$. Then, if the conditions of Lemma A.0.1 in Appendix A.0.2.1 are satisfied, we have the following stability estimate for any $t \in [0,T]$

$$\begin{aligned} d_2^2(\rho_t, \widehat{\rho}_t) \leqslant & \exp\{2(1 + L_{\widehat{V}}^2 + L_{\widehat{W}}^2)t\} \\ & \times \left(d_2^2(\widehat{\rho}_0, \rho_0) + 2T\tilde{\mathcal{E}}_\infty(\widehat{W}) + 2\int_0^t \|\nabla V - \nabla \widehat{V}\|_{L^2(\rho_s)}^2 \mathrm{d}s \right) . \end{aligned}$$

Proof. We refer the reader to Appendix A.0.2.1 for the proof of this proposition as well as the statement and proof of Lemma A.0.1. □

Remark 2.3.5. Note that in Proposition 2.3.4 we assumed that the solutions $\rho, \widehat{\rho}$ are smooth. By well known properties of the porous medium equation, this implies that the solutions are bounded away from 0. However, this restriction can be removed by following the approximation arguments of the proof of Theorem 1 in [Ott01]. This yields weak solutions in $L^1(\Omega)$ by approximating with smooth solutions as the one considered in Proposition 2.3.4. We note that here we consider a slightly more general energy functional than in [Ott01], since our case includes an interaction term. However, the approximation argument goes through with minor modifications and we omit it it here for the sake of brevity.

2.4 Proposed Numerical Schemes

In practical scenarios, our access is limited to discrete-time data. Consequently, this section outlines the numerical discretization of the error functional (2.2.11) and of the Basis Pursuit method introduced in Section 2.2. We approximate all the integrals by a numerical quadrature

rule and all the computations are carried out on a regular mesh. We present the fully discretized estimator for the 1D case. The generalisation to higher dimensions can be derived analogously. We consider a computational domain $[0,T]\times\Omega$, with $T>0$, $\Omega=[-R,R]$ for $R>0$ and R chosen large enough such that the essential support of ρ is contained in $[0,T]\times\Omega$. We will not estimate the cut-off error produced by this assumption as it is zero if the solution is compactly supported for all times with support in Ω. This is the case for all PDEs (2.1.1) if the diffusion is degenerate at zero.

2.4.1 Discrete error functional

Let us take a space-time mesh size of $(\Delta x,\Delta t)$ and denote $t_\ell=\ell\Delta t$ where ℓ ranges from 0 to $\lceil T/\Delta t\rceil$, and $x_m=m\Delta x$ where m spans from $-M$ to M, with M defined as $\lceil 2R/\Delta x\rceil$. For any function $v(t,x)$, denote $v_m^\ell\approx v(t_\ell,x_m)$, $v^\ell\approx v(t_\ell,x)$ and $v_m\approx v(t,x_m)$. The given discrete input data is $\{\rho_m^\ell:=\rho(t_\ell,x_m)\}_{m=-M,\ell=1}^{M,L+1}$. We define the standard (forward (+)) finite difference operators δ_t^+ and δ_x^+ to approximate $\partial_t v$ and $\partial_x v$, respectively,

$$(\delta_t^+ v)_m^\ell=\frac{v_m^{\ell+1}-v_m^\ell}{\Delta t}\ ,\qquad (\delta_x^+ v)_m^\ell=\begin{cases}-v_m^\ell/\Delta x, & \text{if } m=M\ ,\\ \frac{v_{m+1}^\ell-v_m^\ell}{\Delta x}, & \text{if } m<M\ .\end{cases}\tag{2.4.1}$$

For simplicity of the notation, we omit the parenthesis in the previous definitions and we write $\delta_x^+ v_m^\ell$ and $\delta_t^+ v_m^\ell$. We will employ numerical quadratures utilizing discrete-time data to approximate the continuous integrals necessary for calculating A and b, as detailed in (2.2.9) and (2.2.10). This approach leads to a discretized version of A and b, denoted by $A_{n,M,L}$ and $b_{n,M,L}$, respectively, satisfying the approximations:

$$A\approx A_{n,M,L}\ ,\qquad b\approx b_{n,M,L}\ .$$

Then we can write the discrete error functional, similar to (2.2.11), as

$$\mathcal{E}_\infty^{n,M,L}(c)=c^\top A_{n,M,L}c-2b_{n,M,L}^\top c\ .\tag{2.4.2}$$

In principle, choosing quadrature methods that correspond to the smoothness of the integrands is crucial for effective computation. Following the approach in [LL22], we use a straightforward first-order forward Euler scheme, which makes minimal assumptions about the smoothness of the integrands. We introduce a series of functionals that are useful in defining our numerical scheme. For $t\in[0,T]$, $x\in\Omega$ and $\nabla\Psi_i(x)=\psi_i(|x|)\frac{x}{|x|}$ the i-th basis function for the potential

and its derivative, respectively, with $\psi_i(r) = \Psi_i'(r)$ for $r \in \mathbb{R}^+$, we define:

$$R^i_{n,M,L}(t,x) := \sum_{m=-M}^{M} \Psi_i(x - x_m)\rho(t,x_m)\Delta x \approx \int_{\mathbb{R}} \Psi_i(x-y)\rho(t,y)dy = \Psi_i * \rho \ , \tag{2.4.3}$$

$$C^i_{n,M,L}(t,x) := \sum_{m=-M}^{M} \nabla\Psi_i(x - x_m)\rho(t,x_m)\Delta x \approx \int_{\mathbb{R}} \nabla\Psi_i(x-y)\rho(t,y)dy = \nabla\Psi_i * \rho \ , \tag{2.4.4}$$

$$\widehat{\partial}_t\rho(t,x) := \sum_{\ell=1}^{L} (\delta_t^+\rho)^\ell \ \mathbf{1}_{[t_\ell, t_{\ell+1})}(t) \ , \tag{2.4.5}$$

$$F_{M,L}(t,x) := \rho\Big(\sum_{m=-M}^{M} (\delta_x^+ H'(\rho))_m \mathbf{1}_{[x_m, x_{m+1})}(x) + \partial_x V \Big) \approx \rho \ \partial_x(H'(\rho) + V) \ . \tag{2.4.6}$$

In (2.4.6), the notation $\partial_x V(x)$ indicates that we compute this term analytically since the confinement potential $V(x)$ is a known function in our setting. Then we get a discretization of A and b in (2.2.9) and (2.2.10) such that for $i, j = 1, \ldots, n$ we write

$$A(i,j) \approx A_{n,M,L}(i,j) : = \frac{1}{T} \sum_{m=-M,\ell=1}^{M,L} (C^i_{n,M,L})^\ell_m (C^j_{n,M,L})^\ell_m \rho^\ell_m \Delta x \Delta t \ , \tag{2.4.7}$$

$$b(i) \approx b_{n,M,L}(i) : = -\frac{1}{T} \sum_{m=1,\ell=1}^{M,L} \Big((\widehat{\partial}_t \rho R^i_{n,M,L})^\ell_m + (C^i_{n,M,L} F_{M,L})^\ell_m \Big) \Delta x \Delta t \ . \tag{2.4.8}$$

Note that in our numerical examples, we have the solution ρ (approximately) compactly supported on $[-R, R]$ and therefore the integral kernel G defined in (2.2.7) is (approximately) supported on $[-2R, 2R] \times [-2R, 2R]$. In practice, we rewrite the previous approximations to compute $A(i,j)$ using the formulas in (2.2.9) as follows. We approximate G on the extension of the solution mesh: for $y, z \in \mathbb{R}$, $i, j = 1, \cdots, n$,

$$G(y,z) \approx G_{M,L}(y,z) = \sum_{\ell=1}^{L} \sum_{m=-M}^{M} \rho(t_\ell, x_m - y)\rho(t_\ell, x_m - z)\rho(t_\ell, x_m)\Delta x \Delta t \ . \tag{2.4.9}$$

Hence, we can rewrite equation (2.4.7) as

$$A_{n,M,L}(i,j) = \frac{1}{T} \sum_{\ell=1}^{L} \sum_{m,m'=-2M}^{2M} \nabla\Psi_i(x_m)\nabla\Psi_j(x_{m'}) G_{M,L}(x_m, x_{m'}) (\Delta x)^2 \ . \tag{2.4.10}$$

These empirical quantities give rise to a linear system $A_{n,M,L}c = b_{n,M,L} + e$ where

$$e = (A_{n,M,L} - A)c + (b - b_{n,M,L}) \ . \tag{2.4.11}$$

We then solve the following basis pursuit problem

$$\begin{aligned} &\text{minimize}_{\hat{c}\in\mathbb{R}^n} \|\hat{c}\|_1 \ , \\ &\text{such that } A_{n,M,L}\hat{c} = b_{n,M,L} \ , \end{aligned} \tag{2.4.12}$$

which is a perturbed version of (2.2.13). Finally, we write the discrete estimator as $\hat{\Psi}_{n,M,L} = \sum_{i=1}^{n} \hat{c}_i \Psi_i$.

2.4.2 Error bounds

It is expected that $A_{n,M,L}$ and $b_{n,M,L}$ will converge to A and b as $\Delta x,\ \Delta t \to 0$ and the convergence rate depends on the regularity of the solutions and the basis functions. Therefore we first introduce some preliminary assumptions on ρ and the basis functions of $\mathcal{H}$.

Assumption 2. Assume that $\rho \in W^{2,\infty}([0,T]\times\Omega)$ and $\mathcal{J} = H'(\rho) + V \in \mathcal{W}^{2,\infty}([0,T]\times\Omega)$.

Assumption 3. Assume $\mathcal{H} = \text{span}\{\Psi_i\}_{i=1}^n$ consists of radial functions, $\nabla\Psi_i(x) = \psi_i(|x|)\frac{x}{|x|}$, with $\psi_i \in \mathcal{W}^{2,\infty}(\bar{\Omega})$ for $\bar{\Omega} = [-2R, 2R]$.

The convergence analysis is addressed in [LL22] and our two assumptions above are based on [LL22, Assumption 3.1] and [LL22, Assumption 3.2], respectively. However, two differences are present in our approach. In [LL22], $\mathcal{J}$ is the linear diffusion term whose regularity is determined by ρ, whereas our work extends regularity assumptions to a more general form of $\mathcal{J}$ that includes nonlinear diffusion. Furthermore, [LL22] assumes that the basis functions are compactly supported, motivated by the use of a local spline basis. However, in our context, neither the basis functions nor the external potential functions need to be compactly supported. Instead, our approach involves considering their restrictions within a bounded domain. This is evident from (2.4.10) and the compact support property of ρ, where the numerical error analysis only needs to be applied to functions defined over $[-2R, 2R]$ or $[-R, R]$.

Proposition 2.4.1. Under the Assumptions 2 and 3, the discretization error of $A_{n,M,L}$ and $b_{n,M,L}$ in (2.4.7) and (2.4.8) are bounded by

$$|A(i,j) - A_{n,M,L}(i,j)| \leqslant \alpha(\Delta t + \Delta x) \ , \tag{2.4.13}$$

$$|b(i) - b_{n,M,L}(i)| \leqslant \beta(\Delta t + \Delta x) \ , \tag{2.4.14}$$

for any $i, j = 1, \ldots, n$, where α is a constant depending on R, $\|\rho\|_{1,\infty}$ and the bounds of the basis functions $\|\psi_i\|_{1,\infty}$, $i = 1, \ldots, n$, and β depends on R, $\|\rho\|_{2,\infty}$, $\|H'(\rho) + V\|_{2,\infty}$, and the bounds of the basis functions $\|\psi_i\|_{2,\infty}$, $i = 1, \ldots, n$.

Proof. The proof of (2.4.13) is identical to the one presented in [LL22]. For (2.4.14), the only difference lies in estimating $\|F - F_{M,L}\|_\infty$ where $F = \rho\partial_x J$ and its quadrature is defined in (2.4.6). Note that

$$\|F - F_{M,L}\|_\infty \leqslant \|\rho\|_\infty \Big|\partial_x(H'(\rho) + V) - \sum_{m=-M}^{M} (\delta_x^+(H'(\rho) + V))_m \mathbf{1}_{[x_m, x_{m+1})}\Big|_\infty \leqslant \|\rho\|_\infty C\Delta x\,,$$

with $C = \|H'(\rho) + V\|_{2,\infty}$. So the above estimate slightly generalizes [LL22] by considering a general form of diffusion that satisfies the same smoothness assumption as ρ. □

Remark 2.4.1. In Proposition 2.4.1 we assume the solution data is exact and there is no forward error from the numerical solver. The error committed in the approximation of $A_{n,M,L}$ is only due to the numerical integration. Since our quadrature rule is the middle point formula and there is no derivative involved in the expression of A, (2.4.13) can be improved to spatial accuracy $(\Delta x)^2$. Note that one can use centered finite difference to approximate the spatial derivative, and the result can be improved from Δx to $(\Delta x)^2$ in (2.4.14). If a higher order quadrature rule is used in time, we expect analogous improvements in the approximation with respect to time for (2.4.13). However, we cannot expect the improvement on (2.4.14) as we need to perform numerical quadrature on

$$\partial_t \rho \Psi_i * \rho \in W^{1,\infty}([0,T] \times \Omega)\,,$$

in approximating b where $\mathcal{O}(\Delta t)$ is already optimal.

Implications for Optimal Estimation Accuracy: Consider the true support of the coefficient vector c, denoted by $\mathcal{I} \subseteq \{1, \dots, n\}$. If PartInv accurately identifies $\mathcal{I}$, we define our estimator as $\hat{c}(\mathcal{I}) = \big((A_{n,M,L})_\mathcal{I}\big)^+ b_{n,M,L}$. Despite the potential ill-conditioning of A, it is plausible to assume that $A_\mathcal{I}$ possesses a bounded inverse, especially when $\mathcal{I}$ is a relatively small set. Assuming Δx and Δt are small enough for $(A_{n,M,L})_\mathcal{I}$ to maintain a bounded inverse, we derive the following equation:

$$\begin{aligned}
\|\hat{c} - c\| &= \|[(A_{n,M,L})_\mathcal{I}]^+ b_{n,M,L} - A_\mathcal{I}^+ b\| \\
&= \|([(A_{n,M,L})_\mathcal{I}]^+ - A_\mathcal{I}^+) b_{n,M,L} + A_\mathcal{I}^+ (b_{n,M,L} - b)\| \\
&= \mathcal{O}(\|A_\mathcal{I} - (A_{n,M,L})_\mathcal{I}\| + \|b_{n,M,L} - b\|) = \mathcal{O}(n(|\mathcal{I}| + 1)(\Delta x + \Delta t))\,.
\end{aligned}$$

This derivation employs the perturbation bounds on the pseudo-inverse as discussed in [Ste77]. The accuracy of the estimator $\hat{c}$ is thus influenced by both, the accuracy of identifying $\mathcal{I}$ and the perturbations in A and b, quantified in terms of Δx and Δt.

2.4.3 Noisy data

To test the robustness of the proposed method, we also consider the case where the solution data is corrupted by observational noise. In particular, we analyze the effects of adding i.i.d. random noise with zero mean to the discretized samples of ρ. Hence, in this case, the final data set used for the estimation of the interaction kernel in the numerical examples is given by

$$\{\tilde{\rho}(t_\ell, x_m)\}_{m=-M,\ell=1}^{M,L} \ , \tag{2.4.15}$$

where $\tilde{\rho}(t_\ell, x_m) = \rho(t_\ell, x_m) + \epsilon_m^\ell$. In our numerical examples, we used $\epsilon_m^\ell \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. To ensure that the perturbation due to the noise is on a similar scale to the solution ρ we will set

$$\sigma = \frac{p}{100}\left(\sum_{\ell=1}^{L}\sum_{m=-M}^{M}(\rho_m^\ell)^2\Delta x\Delta t\right)^{\frac{1}{2}} \ ,$$

for some constant $p \in [0, 100]$. We refer to this as the noise being p-percent. In what follows we denote by $\|\cdot\|_{L^2(\varepsilon)}$ the L^2 norm over the probability space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P})$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-algebra and $\mathbb{P}$ is a probability measure, i.e. for a random variable defined over this space we have

$$\|X\|_{L^2(\varepsilon)} := \sqrt{\mathbb{E}|X|^2} \ .$$

Let us define

$$\widetilde{A}_{n,M,L}(i,j) := \frac{1}{T}\sum_{m=-M,\ell=1}^{M,L}(\tilde{C}_{n,M,L}^i)_m^\ell(\tilde{C}_{n,M,L}^j)_m^\ell\tilde{\rho}_m^\ell\Delta x\Delta t \ , \tag{2.4.16}$$

$$\widetilde{b}_{n,M,L}(i) := -\frac{1}{T}\sum_{m=-M,l=1}^{M,L}\left((\widehat{\partial}_t\tilde{\rho}\tilde{R}_{n,M,L}^i)_m^\ell + (\tilde{C}_{n,M,L}^i\tilde{F}_{M,L})_m^\ell\right)\Delta x\Delta t \tag{2.4.17}$$

where $\tilde{C}_{n,M,L}, \tilde{R}_{n,M,L}$ and $\tilde{F}_{n,M,L}$ are defined analogously to (2.4.3)-(2.4.6), but depending on $\tilde{\rho}$. In this framework, we can obtain the following extension of the error bounds in Proposition 2.4.1.

Proposition 2.4.2. The numerical error of $A_{n,M,L}$ in (2.4.7) when we consider the perturbed solution $\tilde{\rho}$ as in (2.4.15) is

$$\|A - \widetilde{A}_{n,M,L}\|_{L^2(\epsilon)} \leqslant \alpha n(\Delta t + \Delta x) + nC(\sigma\sqrt{\Delta t\Delta x} + \sigma^2\Delta x) \ , \tag{2.4.18}$$

where α is as in Proposition 2.4.1 and $C > 0$ is a constant depending on R, T and $\|\psi_i\|_\infty$, $i = 1, \ldots, n$.

Proof. The error induced by random perturbations is additive, and as a consequence

$$\|A - \widetilde{A}_{n,M,L}\|_{L^2(\epsilon)} \leqslant \|A - A_{n,M,L}\|_{L^2(\epsilon)} + \|A_{n,M,L} - \widetilde{A}_{n,M,L}\|_{L^2(\epsilon)}.$$

The first part is estimated as in Proposition 2.4.1 while the second term is discussed in Appendix B. □

Remark 2.4.2. If a centered finite differences method is used and we have $H(\rho) = \frac{\rho^2}{2}$, then we can obtain the following error estimate for the numerical error of $b_{n,M,L}$ with added noise,

$$\|b - \widetilde{b}_{n,M,L}\|_{L^2(\epsilon)} \leqslant \beta\sqrt{n}(\Delta t + \Delta x) + \sqrt{n}C\sigma^2(\Delta x^{-1} + \Delta x \Delta t^{-1}) \ , \tag{2.4.19}$$

where we note that the inverse dependence on the mesh size is due to the discrete derivatives in $b_{n,M,L}$. We remark that if an upwind scheme is used for the computation of derivatives or we have an arbitrary free energy kernel, H, the nonlinearities impede any explicit numerical error estimate.

Proof. As for (2.4.18), the error stemming from the random noise is additive so we have

$$\|b - \widetilde{b}_{n,M,L}\|_{L^2(\epsilon)} \leqslant \|b - b_{n,M,L}\|_{L^2(\epsilon)} + \|b_{n,M,L} - \widetilde{b}_{n,M,L}\|_{L^2(\epsilon)} \ .$$

Again, the first term is controlled as in Proposition 2.4.1 and the second term is treated in Appendix B. □

2.4.4 Support pruning algorithm

In this section, we discuss strategies for finding the right support of the coefficient vector when the data are not accurate. Once the true support is identified, we can perform restricted least squares on the support set and therefore improve the robustness of the algorithm.

Given the discrete data, we apply Algorithm 1 on the BP problem (2.4.12). When the discretization error terms in (2.4.18) and (2.4.19), are small, it is effective to select the sparsity level of PartInv algorithm $K = s$, i.e., the exact sparsity of the true coefficient vector c. However, in cases where this error becomes significant such as when the dimension of the basis (i.e. n) is large, or errors coming from the discretization and noise, $(\Delta x, \Delta t, \sigma)$ increase, this choice often results in inaccurate support identification, adversely affecting the recovery of the interaction potential (see Figure 2.4 (a)). In such situations, it is advantageous to choose $K \geqslant s + 1$ in our PartInv method as PartInv consistently produces a support set $\mathcal{I}^{(k)}$ that contains the true support $\mathcal{I}$ as a subset. However, when we perform restricted least squares regression—a method

where the regression coefficients are estimated under certain linear constraints—on $\mathcal{I}^{(k)}$, we may still encounter large estimation errors. This is often due to the ill-conditioning of the regression matrix, which can adversely affect the accuracy of the estimates (see Figure 2.4 (b)). So it is necessary to prune $\mathcal{I}^{(k)}$ to identify the true support $\mathcal{I}$.

We note that it is possible to skip the basis pursuit step, and perform restricted least squares on all possible combinations of indices from the beginning. However, the computational cost in this case is very high. The PartInv helps to reduce the number of combinatorial trials, and increases the computational efficiency of the estimation procedure. We propose the following algorithm, which combines residual error and time evolution error:

Step 1: For each subset J from $\mathcal{I}^{(k)}$, the PartInv output, we compute the coefficient vector $c_J \in \mathbb{R}^n$ using

$$c_J(J) = (A_{n,M,L})_J^+ b_{n,M,L} \ ,$$

and $c_J(J^c) = 0$. We then calculate its associated residual error (RE)

$$c_J^\top A_{n,M,L} c_J - 2\langle c_J, b_{n,M,L}\rangle \ .$$

Step 2: We sort the REs in descending order and identify a cluster of subsets whose residual errors are close to the smallest one falling within a predefined precision threshold τ. This threshold τ is defined as a proportion of the norm of the error vector $|e|$, previously defined in (2.4.11).

Step 3: For each subset within the identified cluster, we use the interaction potential associated with c_J to incorporate it into (2.1.1). Subsequently, we perform a forward solver on a much smaller space-time mesh size $(\widehat{\Delta x}, \widehat{\Delta t})$ than $(\Delta x, \Delta t)$ and calculate the time evolution error (TEE) using the formula

$$\mathrm{TEE}^2 = \sum_{m=-M',\ell=1}^{M',L'} |\widehat{\rho}_m^\ell - \tilde{\rho}_m^\ell|^2 \Delta x \Delta t \ ,$$

where we may use a subset of training data on a smaller time interval $[0, \widehat{T}]$ for validation.

Step 2 draws its motivation from Proposition 2.2.1, guiding the pursuit of estimators capable of accurately reproducing the training data. Particularly when e is in a reasonable range, this strategy effectively narrows down candidate estimators for Step 3, providing computational efficiency given the potentially high computational cost of this subsequent step. In our numerical experiments, it is often easy for us to identify such a cluster of values that are close to a minimum.

It is noteworthy that while the literature on sparse signal processing does present support pruning algorithms, our learning problem distinctively diverges due to the nonlinear relationship between the coefficient vector and the solution data. TEE, initially proposed in [KLL21], is employed for support pruning in the sparse identification of nonlinear PDEs, using a LASSO-based algorithm. At its core, the fundamental notion is that if the true PDE identifies the underlying dynamics, any further refinement in the discretization of the time domain should adhere to the given data. This adherence is ensured by the consistency, stability, and convergence of a numerical scheme.

To conclude, it is important to note that when e is large, estimators may be yielded with approximately equivalent TEEs. In such instances, it is prudent to select the estimator yielding the sparser solution, aligning with the Akaike information criteria.

2.5 Numerical examples

In this section we systematically apply the algorithm outlined in Section 2.4 for the estimation of the interaction potential, to several instances of

$$\partial_t \rho = \nabla \cdot [\rho \nabla (H'(\rho) + V(x) + W * \rho)] \,, \tag{2.5.1}$$

showcasing a wide range of dynamics. In particular, we consider examples with different initial data and potentials, as well as dynamics modulated by an external potential V, in one and two dimensions. In the examples below, we either consider nonlinear diffusion, where $H(\rho) = \kappa \frac{\rho^m}{m-1}$, or linear diffusion, where $H(\rho) = \kappa \rho(\log \rho - 1)$.

The evaluation of the algorithm's performance hinges on the computation of the relative reconstruction error defined as

$$E_{\text{reconst}} = \frac{\|c - \widehat{c}\|_2}{\|c\|_2}. \tag{2.5.2}$$

2.5.1 Data generation

To evaluate the estimation approach, the data is produced by solving (2.5.1) employing a finite volume method on a grid of high resolution, using a space-time mesh size of $(\delta x, \delta t)$, and the solution is obtained over the time interval $[0, T]$. One could choose very fine δx, and δt so that we minimize the numerical error from the solver to a negligible level. More precisely, we use a semi-discrete (discrete in space only) second-order finite volume scheme as presented in [CCH15]. This scheme uses a third-order strong preserving Runge-Kutta ODE solver [GST01]. It preserves positivity of the average solution in each cell provided a CFL

condition, $\delta t \leqslant \frac{\delta x}{2\max_m \left\{u^{\ell\,+}_{m+\frac{1}{2}}, -u^{\ell\,-}_{m-\frac{1}{2}}\right\}}$, is satisfied, where $u^{\ell\,+}_{m+\frac{1}{2}}$ and $u^{\ell\,-}_{m-\frac{1}{2}}$ are the right and left discrete velocity fields in each cell, respectively. Since this finite volume scheme is obtained by integrating equation (2.5.1) over each cell, it is easily generalized to higher dimensions where, in the 2D case, the velocity field is computed over squared cells. We consider no-flux boundary conditions in all cases.

Subsequently, this simulated data is constrained to a coarser grid characterized by a mesh of size $(\Delta x, \Delta t)$, where $\Delta x = C_x \delta x$ and $\Delta t = C_t \delta t$. Here C_x and C_t are referred to as the downsampling factors. These factors represent the level of resolution present in the observational data.

Notation	Description
$(\delta x, \delta t)$	Space-time step size used in finite volume solver
$\Delta x = C_x \delta x$	Space size in observational data
$\Delta t = C_t \delta t$	Time step size in observational data
$(\widehat{\Delta} x, \widehat{\Delta} t)$	Space-time step size used in finite volume solver in support pruning step

Table 2.1: Notations of space-time step size.

Overview of numerical experiments. In the following section, we test the effectiveness of algorithms over 1D and 2D numerical examples that display various collective behaviours.

- For each example, we assess the effectiveness of PartInv across different data scenarios by using the error metric defined in (2.5.2). Firstly, we examine the case of noise-free data, we first generate our data by using an approximation of the PDE obtained in a very fine mesh. Then, the major source of errors arises from the evaluation error of the functional (2.4.2) introduced in the observational data by the downsampling procedure above. Secondly, we explore scenarios with noise contamination, keeping the space-time resolution constant. It is important to note that introducing Gaussian noise might lead to negative values in the solution data. This scenario is at odds with the reality that the actual solution data should be positive. However, in this study, we intentionally avoid using any denoising techniques. Our aim is to evaluate the resilience of PartInv even when the solution data deviates from physical constraints. Finally, we also test the robustness of the method when the solution data is obtained at coarse scale in a 2D example (See Example 5).

- We test the effects of different choices of the sparsity parameter K in the PartInv algorithm on the reconstruction accuracy and show how the support pruning algorithm can help

stabilize the results. We thereby provide a comprehensive check of robustness for PartInv.

- We show regularization is necessary in our estimation problem and sparsity-promoting is effective. Indeed, the least squares estimator yields inaccurate estimates while promoting sparsity can yield very accurate estimations. See Figure 2.2.

- We perform comparative tests between PartInv and standard solvers in the field of PDE sparse identification: LASSO-type estimators, Greedy methods such as Subspace pursuit[1], and Sequential Thresholded least squares (SINDy), all within the framework of basis pursuit, see Figure 2.7 and 2.9 in Example 2 and 3. Additional examination is performed to contrast the proposed data-fidelity term with that invoked by the strong form of PDEs, commonly utilized in PDE literature; for a relevant example we refer to Figure 2.9.

2.5.2 One dimensional examples

Consider the one-dimensional aggregation-diffusion equation given by

$$\partial_t \rho = (\rho(\kappa \rho^{m-1} + W * \rho + V)_x)_x \ ,$$

where $W(x) = \Phi(|x|)$ and $\Phi'(|x|) = \phi(|x|)\text{sign}(x)$.

Example 1 (Nonlinear diffusion and compactly supported attraction potential) We consider the nonlinear diffusion case where $m = 2$, $\kappa = 0.2$ and $V = 0$. The initial condition is $\rho_0(x) = \mathbf{1}_{[-2,2]}(x)$ and we have a compactly supported interaction potential given by

$$W(x) = -5(1 - |x|)_+ \ .$$

δt	δx	Time domain	Spatial domain	Initial condition	$\phi(\|x\|)$
$0.5 * 10^{-4}$	10^{-2}	$[0, 0.5]$	$[-6, 6]$	$\mathbf{1}_{[-2,2]}(x)$	$5\mathbf{1}_{[0,1](\|x\|)}$

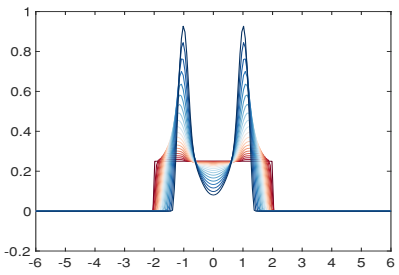
Table 2.2: (CP) Parameters to produce the solution data using a finite volume scheme.

The solution data is produced with the parameters in Table 2.2. These dynamics have the capability to simulate formation of clustered solutions which, after some time, merge together as a result of the attraction potential and the very weak diffusion, see the profile of trajectory data used in our training in Figure 2.1 (a). Note that the solution profile obtained is a transient state and we expect these two bumps to merge together at longer times, given the attraction range of

[1]the algorithm only differs from CoSaMP in choosing sparsity.

the potential. Considering further away initial conditions, or weaker interaction potentials, leads to a steady state of disconnected support [CCH15].

Applications of this particular dynamics can be found for instance in interacting populations of cells. Cells from different colonies can start moving towards each other if they are at a certain sensing distance, forming bigger aggregates as a survival mechanism. In time, this will be observed in Figure 2.1 (a).

(a) 

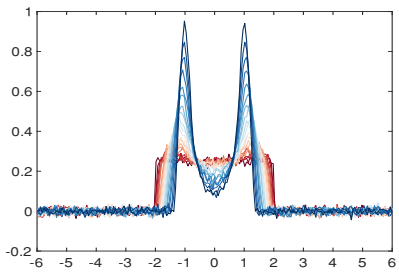
(b) 

Figure 2.1: Profile of a subset of trajectory data used in our training where we choose $\Delta x = 6\delta x$ and $\Delta t = 50\delta t$. A spectrum of colors transitioning from red to blue to symbolize the passage of time. (a) data generated from the numerical solver (b) data with 3% noise added.

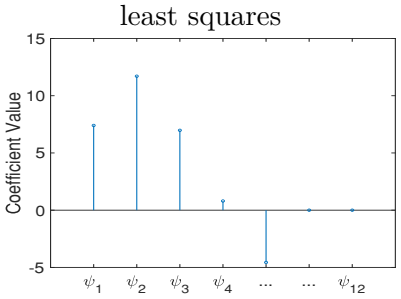
We consider the estimation of the interaction kernel ϕ on the positive axis, and the results on the negative axis will follow automatically by employing the radial symmetry. We use a local piecewise linear ($p = 0, 1$) or constant basis ($p = 0$) of the form $\{x^p \cdot \mathbf{1}_{[\frac{6j}{n}, \frac{6(j+1)}{n}]}(x) : j = 0, \cdots, n-1\}$. We choose $n = 12$. In the context of the piecewise linear basis (dimension = 24), the true interaction kernel is 2-sparse with respect to this particular basis representation. Similarly, when using the piecewise constant basis (dimension = 12), the true interaction kernel also exhibits a 2-sparse characteristic in relation to its basis representation.

In Figure 2.2, we show the efficacy of sparsity-promoting in the proposed algorithm by comparing the least squares estimator using the pseudoinverse depicted in (a) with our estimator in (b) using piecewise constant basis for the noise-free data. We see the least squares estimator failed in this case, while our estimator obtained from the sparsity-promoting algorithm produced an accurate estimate since it identified a correct 2-sparse representation.

In addition, we test the robustness of PartInv with respect to data perturbations coming from observation noise and discretization errors, and summarize the results in Figure 2.3. In this set of experiments, we observed in all challenging data regimes (large discretization error and/or large noise) PartInv accurately identified the correct support and the reconstruction error got amplified because of the corruption of the data. In addition, for a fixed space-time mesh size, we observe in Figure 2.3(a) that the reconstruction error depends linearly on the noise variance. Furthermore, we display the relative reconstruction errors with different choices of space-time mesh size $(\Delta x, \Delta t)$ in Figure 2.3(b). We see the error depends roughly linearly with respect to

Δx, but did not vary much with respect to (Δt) in our selected range. This is possible as in our error analysis, the coefficient in front of $(\Delta t)^2$ may be relatively small, and at the current scale it is dominated by the errors in the Δx term.

(a)



(b)

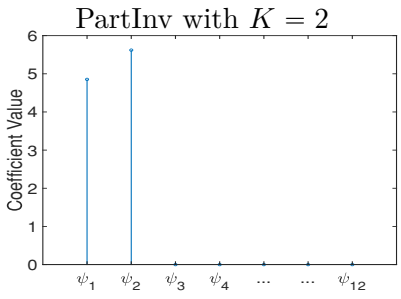
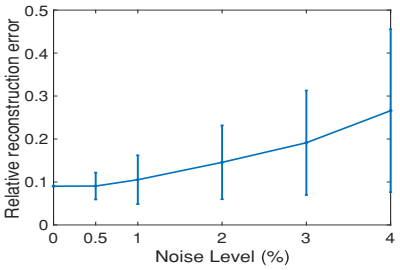


Figure 2.2: Results with piecewise constant basis where we choose $\Delta x = 6\delta x$ and $\Delta t = 50\delta t$. From (b), we clearly see that sparsity effectively regularizes the inverse problem and finds a solution that closely aligns with the ground truth coefficient vector $[5, 5]$ with respect to the basis $[\psi_1, \psi_2]$.

(a)



(b)

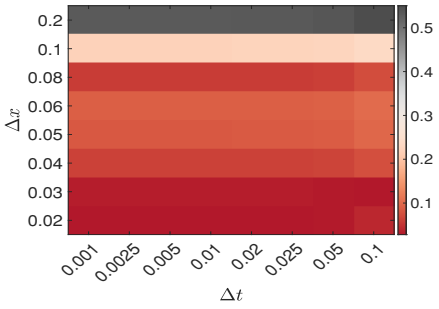


Figure 2.3: Results of PartInv with sparsity $K = 2$ using piecewise constant basis. (a) Accuracy for different levels of noise where we display the mean and standard deviation of relative errors over 100 trials. (b) Relation between the relative reconstruction error, given by the different color intensities, and the mesh size $(\Delta x, \Delta t)$.

The effectiveness of PartInv depends, in part, on the choice of the basis. To illustrate this, we investigate the performance of the algorithm over a piecewise linear basis. As evidenced in Figure 2.4 (a), using the piecewise linear basis leads to inaccurate recovery using the same noise-free training data and parameters as in Figure 2.2. The reason is that using a larger dictionary increases the probability of obtaining a matrix A whose columns present high coherence with those corresponding to the true support, making the identification of the true support particularly difficult in such instances.

To circumvent this challenge, it is advantageous to choose a larger K and implement our support pruning algorithm, introduced in Section 2.4.4. We see from Figure 2.4 (b) that PartInv outputs a support set $\{1, 2, 3\}$. Then we apply the support pruning algorithm, as depicted in

Figure 2.4 (c), where we identify the right support set $\{1, 3\}$ [2] using the numerical values in Table 2.3, yielding accurate coefficient estimation.

Active terms	Coefs	RE	TEE	Active Terms	Coefs	RE	TEE
ψ_1	11.56	-0.24		$[\psi_1, \psi_2]$	[-52.97,208.05]	0.59	
ψ_2	37.88	-0.29		$[\psi_1, \psi_3]$	**[4.69,5.63]**	**-0.40**	**0.04**
ψ_3	8.40	-0.31		$[\psi_2, \psi_3]$	[16.49,5.28]	-0.39	0.16
$[\psi_1, \psi_2, \psi_3]$	[9.55, -17.18, 6.01]	-0.38	0.40				

Table 2.3: Numerical results for the pruning algorithms where we refine the finite volume solution using a mesh size $\widehat{\Delta}x = \frac{\delta x}{2}$ and $\widehat{\Delta}t = \frac{\delta t}{4}$.

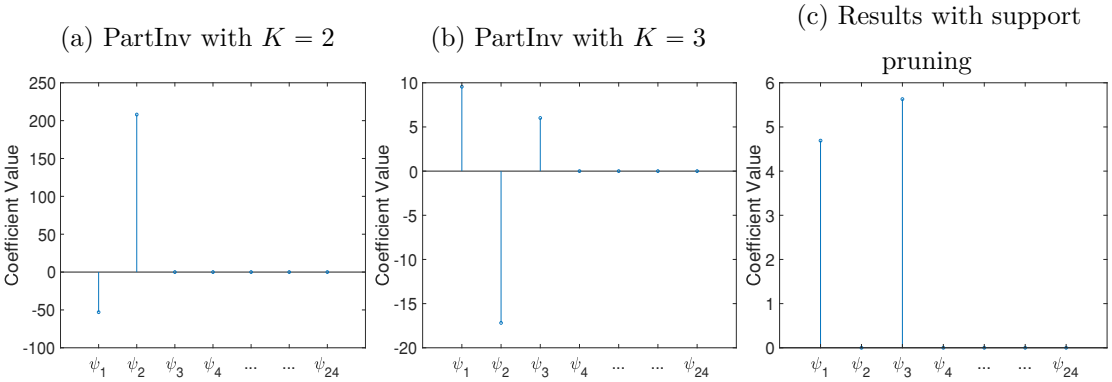


Figure 2.4: Results for PartInv with piecewise linear basis where we use the same training data as in Figure 2.2. (a)-(b) are the cases without support pruning. The case with support pruning with $K = 3$ is presented in (c). We see it produced the most accurate estimation of the true coefficient $[5, 5]$ with respect to the basis (ψ_1, ψ_3).

Example 2 (Nonlinear diffusion and nonlocal attraction potential) We consider the nonlinear diffusion case where this time $m = 3$, $\kappa = 0.48$ and $V = 0$. We have a nonlocal attraction interaction potential given by

$$W(x) = -2\frac{\exp(-|x|^2)}{\sqrt{\pi}} - 2\frac{\exp(-|x|^2/2)}{\sqrt{2\pi}}.$$

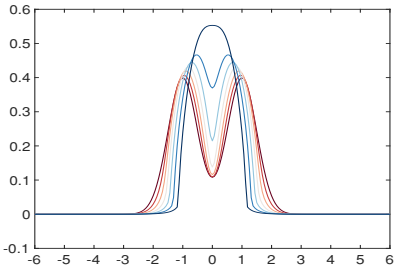
This equation describes spontaneous biological aggregation of e.g. bacteria colonies [TBL06]. An extensive study of the steady states for an analogous example was carried out in [BFH14] where it was observed that, when $m > 2$ the attraction dominates the dynamics leading to compactly supported steady states as observed in Figure 2.5. The dynamics in this case is governed by a competition between the nonlocal attraction, characterized by the term $W * \rho$, and the nonlinear diffusion with exponent m, which represents a local repulsion. To generate

[2] We choose the natural ordering in our piecewise linear basis $\{x^0\mathbf{1}_{[0,1/2]}, x^1\mathbf{1}_{[0,1/2]}, x^0\mathbf{1}_{[1/2,1]}, x^1\mathbf{1}_{[1/2,1]}, ...\}$ so the true interaction kernel is spanned by ψ_1 and ψ_3.

the solution data we used as initial condition $\frac{\mathcal{N}(1,0.5^2)+\mathcal{N}(-1,0.5^2)}{2}$. The solution data profile is plotted in Figure 2.5 (a) and its noisy version in (b).

δt	δx	Time domain	Spatial domain	$\phi(\lvert x\rvert)$
10^{-4}	$1.25 \cdot 10^{-2}$	$[0, 1.5]$	$[-6, 6]$	$\frac{4}{\sqrt{\pi}}\lvert x\rvert \exp(-\lvert x\rvert^2) + \frac{2\lvert x\rvert}{\sqrt{2\pi}} \exp(-\frac{\lvert x\rvert^2}{2})$

Table 2.4: (1D Metastable) Parameters to produce the solution data using a finite volume scheme.

(a) 

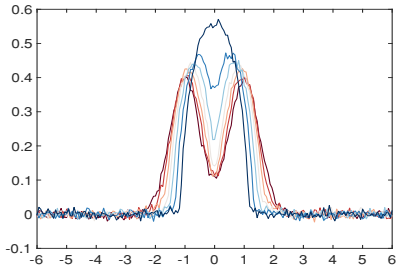
(b) 

Figure 2.5: Profile of the solution for $\Delta x = 5\delta x, \Delta t = 2500\delta t$. (a) a subset of solution data generated from the numerical solver (b) the solution data with 1% noise added.

To estimate the interaction kernel, we use a set of exponential basis of the form $\{\frac{|x|}{6}\exp(-w|x|^2) : w = 0.5 : 0.5 : 5\}$, i.e. the values of w range from 0.5 to 5 in incrementes of 0.5. Then, the true interaction kernel is 2-sparse with respect to this particular basis representation. Figure 2.6 (a) shows that it yields a very coherent basis in our sparse learning problem, as the coherence parameter ranges from 0.982 to 1 (see discussion in Section 2.2.3). In the algorithm, we set the $K = 2$ and PartInv can yield a very accurate estimation as observed in Figure 2.7 (a), for the solution data in Figure 2.5 (a) and (b), where the time observations are very sparse. We also explore its robustness with respect to different Δx and different noise levels for the choice $\Delta x = 5\delta x, \Delta t = 2500\delta t$ and summarize the result in Figure 2.7 (b)-(d), where we also compare with the subspace pursuit and LASSO approach. We see that the reconstruction error is significantly smaller using our approach.

Example 3 (Linear diffusion with external potential V) In this one-dimensional example, we consider an external confinement potential given by a double-well function and linear diffusion. Therefore we have

$$\partial_t \rho = \kappa \partial_{xx}^2 \rho + (\rho(W * \rho + V)_x)_x \ , \text{ where}$$

$$W(x) = \frac{|x|^2}{2} \ , \qquad V(x) = \frac{|x|^4}{4} - \frac{|x|^2}{2} \ .$$

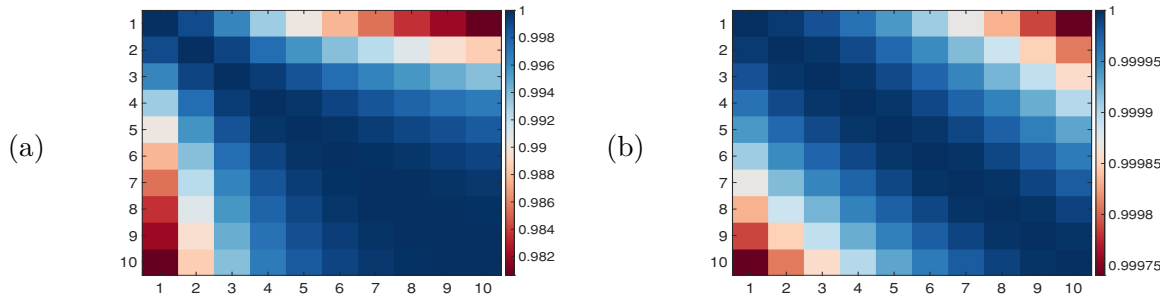


Figure 2.6: Patterns of incoherence in the regression matrix. (a) corresponds to Example 2, where we use Gaussian type basis of size 10. (b) corresponds to Example 3, where we use polynomial basis of size 10.

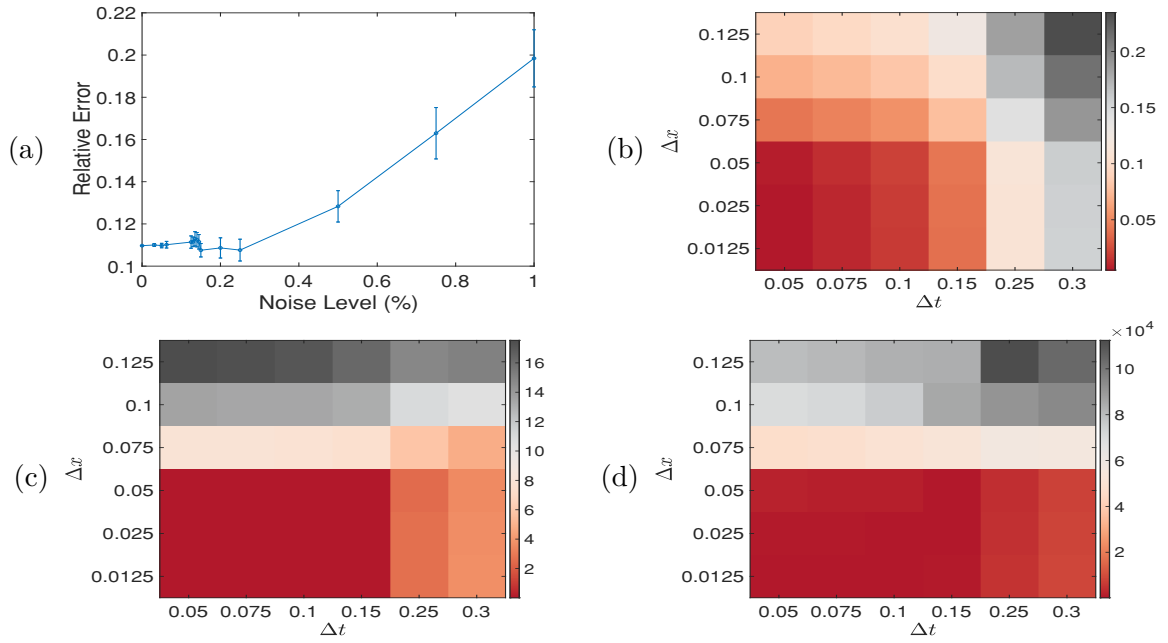


Figure 2.7: Reconstruction errors for solution data in Figure 2.5. (a) We have different levels of noise perturbation where we display the mean and standard deviation of relative errors over 100 trials. (b) Accuracy of PartInv with $K = 2$ using different $(\Delta x, \Delta t)$. (c) Accuracy of subspace pursuit with $K = 2$. (d) Accuracy of LASSO. For the LASSO algorithm, the Matlab-built-in LASSO solver was employed with the 'IndexMinMSE' option.

This equation describes a model for self-propelled agents [BCnCD16] with a noisy term given by the linear diffusion. The confinement potential describes the tendency of individuals to move in a preferred direction while the interaction potential W models the alignment component of the movement. We consider $\kappa = 0.1$ and simulation parameters are provided in Table 2.5. As expected, low values of the diffusion coefficient result in flocking for certain initial conditions [CGPS20].

δt	δx	Time domain	Spatial domain	Initial condition	$\phi(\lvert x\rvert)$
10^{-2}	$1.2 \cdot 10^{-2}$	$[0, 5]$	$[-6, 6]$	$\mathcal{N}(0, 0.3^2)$	$\lvert x\rvert$

Table 2.5: (KF) Parameters to produce the solution data using the finite volume scheme.

Figure 2.8 illustrates the solution profile used as training data. Given that we started with symmetric initial data, a symmetric steady state is anticipated [BCH20]. For kernel estimation, we use a polynomial basis of the form $\left\{\frac{1}{6}\left(|x|\right)^n : n = 0, \ldots, 9\right\}$, such that ϕ is 1-sparse relative to this dictionary. More explicitly, the coefficient for ψ_2 is 6. The normalization factor of 6 on the basis ensures bounded entries in the matrix $A_{n,M,L}$.

Provided continuous-time trajectory data, the uniqueness of a 1-sparse solution to the normal equation (2.2.12) lies in the prerequisite that any two distinct columns of matrix A are linearly independent. This 1-sparse solution is the coefficient of the true interaction kernel.

Nonetheless, our numerical result reveals that every pair of columns in A tends to exhibit near-linear dependence, a phenomenon evidenced by the coherence patterns manifested in its empirical regression matrix $A_{n,M,L}$, as shown in Figure 2.6 (b). This suggests that sparse identification of kernels from discrete noisy data is expected to be difficult, even though the ground truth is 1-sparse. This anticipation is, in part, foreseeable, considering that the potential of type $|x|^n$ is capable of promoting analogous collective dynamics.

The unfavorable coherent patterns lead to failures in LASSO and SINDy estimators, as depicted in Figure 2.9 (c)-(d). In contrast, the greedy type methods yield much more accurate estimations see Figure 2.9 (a) for PartInv and (b) for CoSamP.

When the solution data is tainted by noise, kernel identification becomes considerably more challenging. In Figure 2.9 (e), we employ PartInv on the data, introducing a noise level of 0.5% and imposing a sparsity constraint $K = 2$. Notably, even when PartInv outputs a support set $\{1, 2\}$ encompassing the true support $\{2\}$, the estimated coefficients diverge substantially from the ground truth (see 3rd row of Table 2.6). We first used RE to narrow down two candidates and then compute their TEEs utilizing a space-time mesh size $(\delta x, 0.1\delta t)$ and evolve the PDE over the time interval $[0, 0.1]$. In this instance, the TEE emerges as a robust quantitative metric, aiding in the identification of the correct support set $\{2\}$. As a result, we obtain a significantly accurate estimation.

For this example, we also compare our loss function with the PDE residual. In Figure 2.9 (f), we show that, even when provided with the true support, the restricted least squares (note that this is the optimal outcome attainable through a sparsity-promoting algorithm) yields an estimated coefficient that is notably divergent from the ground truth of 6. This demonstrates

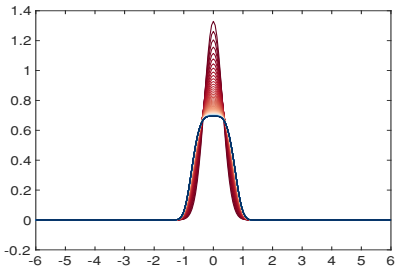
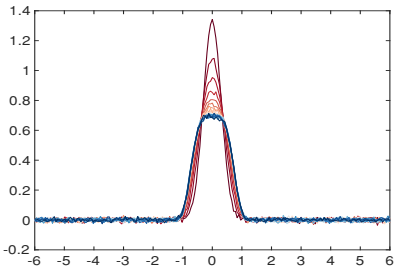
(a)  (b) 

Figure 2.8: Profile of the solution for $\Delta x = 5\delta x, \Delta t = 5\delta t$. (a) a subset of the solution data generated from the numerical solver (b) the solution data with 0.5% noise added.

the advantages of our loss functional over the PDE residual.

Active terms	Coefs	RE	TEE
ψ_1	0.87	-1.22	0.39
ψ_2	**6.91**	-3.34	**0.13**
$[\psi_1, \psi_2]$	[40.20, -313.11]	121.19	

Table 2.6: Numerical values for support pruning. When $K = 2$, PartInv produced $I^{(k)} = \{1, 2\}$ and we perform support pruning by calculating the TEE on time steps $0 : 5\delta t : 10\delta t$.

2.5.3 Two dimensional examples

In this part, we delve into the performance of two dimensional examples. Compared to one dimensional scenarios, 2D examples require considerably more computational resources since the computation of the integration kernel G, given by (2.2.8), becomes increasingly complex with higher dimensionality, reaching computational limits if the mesh size exceeds 100. We present two specific examples: the first involves closed-form analytic solutions, a scenario free from forward errors. The second example features observations on a very coarse scale, designed to evaluate the effectiveness and robustness of the PartInv method with the implementation of support pruning.

Example 4 (2D Fokker Planck Equation with Nonlinear diffusion) In this 2D scenario, we examine a nonlinear diffusion case characterized by $m = 2$, with κ set to 1. The functions $W(x)$ and $V(x)$ are defined as follows:

$$W(x) = \frac{|x|^2}{2}, \qquad V(x) = 0.$$

In this case we have

$$\nabla W(x) = \nabla \Phi(|x|) = \phi(|x|)\frac{x}{|x|} \ .$$

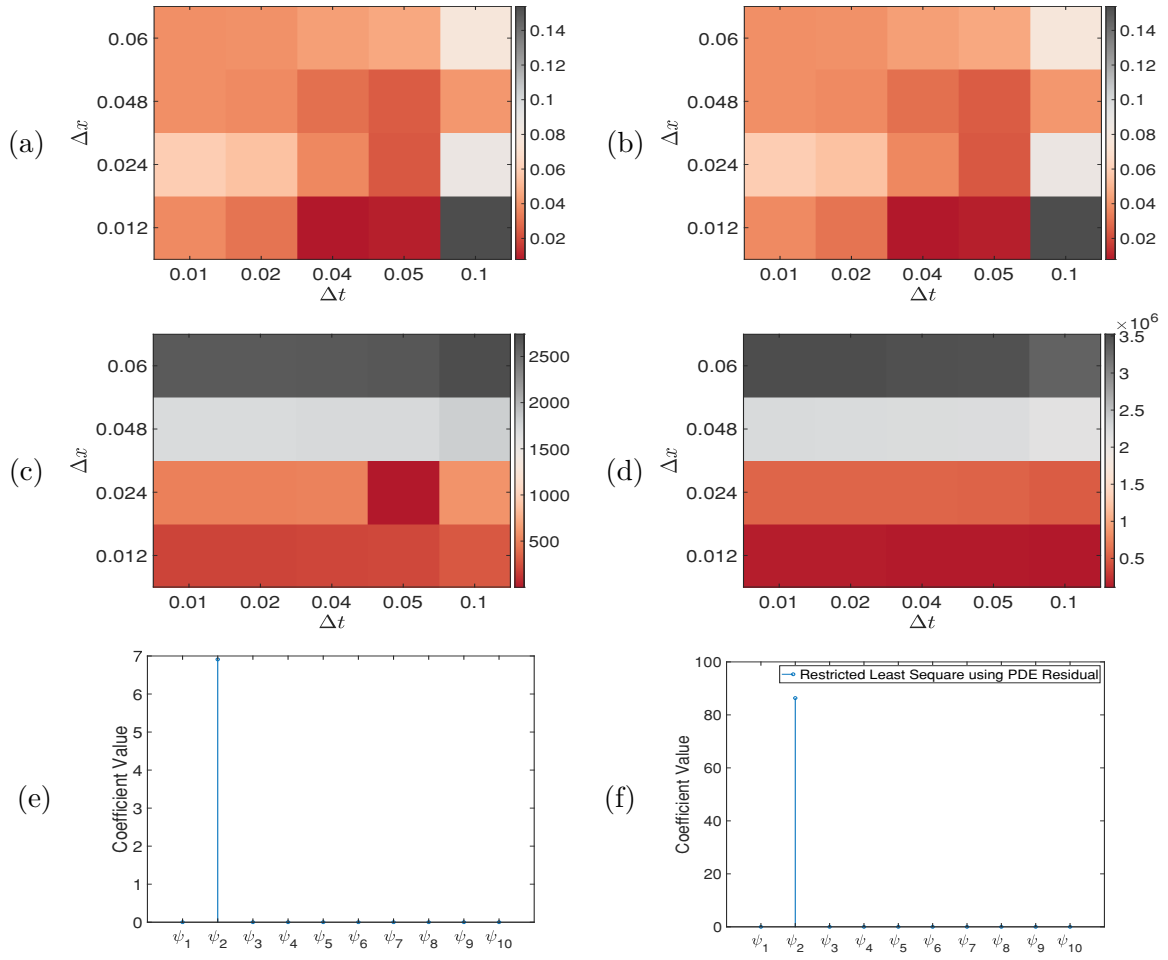


Figure 2.9: (a) the learning outcomes of PartInv with a parameter $K = 3$ with respect to space-time mesh size. (b)-(d): a comparative analysis with subspace pursuit algorithms (b), LASSO (c), and SINDy (d) using identical training data, where the subspace pursuit algorithm was configured with a sparsity level set to 3. We found results in (a) and (b) are very close. In (e) we used the same training data as in Figure 2.8 (b) by choosing $K = 2$ and performing support tuning where the numerical values are summarized in Table 2.6. In (f), the training data is the same as in (e) and we display the restricted least squares estimator using the PDE residual provided the true support set $\{2\}$.

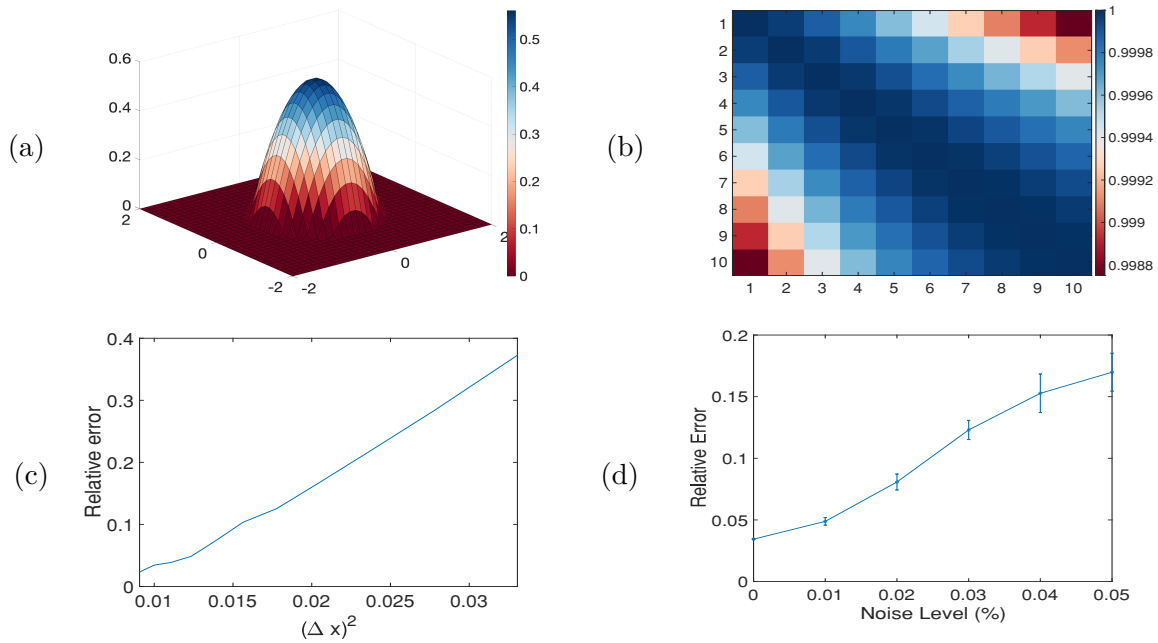


Figure 2.10: **Top Panel:** (a) The profile of the stationary solution at a single time instance (b) The coherence pattern of the regression matrix $A_{n,M,L}$. **Bottom Panel:** (c) We run PartInv with $K = 1$ and display the relative error versus the square of space mesh size and find an approximately linear relationship. (d) We test the robustness of PartInv with $K = 1$ for a variety of noise levels.

Note that now $x \in \mathbb{R}^2$, and to avoid instability issues when $\boldsymbol{x}$ is close to the origin, we learn the kernel $\frac{\phi(|\boldsymbol{x}|)}{|\boldsymbol{x}|}$ instead. Analogous to the Example 3, we employ the polynomial basis $\{1, |\boldsymbol{x}|, \ldots, |\boldsymbol{x}|^{n-1}\}$ with $n = 10$, so the true kernel is 1-sparse to this dictionary of size 10. For the training data, we use the closed form of its stationary solution given by $\rho_t(x) = \max\left(\sqrt{\frac{1}{\pi}} - |x|^2, 0\right)$. The computational parameters are summarized in Table 2.7.

Time domain	Spatial domain	$\frac{\phi(\lvert x\rvert)}{\lvert x\rvert}$
$[0, 0.1]$	$[-2, 2] \times [-2, 2]$	1

Table 2.7: (2DFP) Parameters to produce the solution data using the finite volume scheme.

In the context of this identification problem, we observe a phenomenon within the regression matrix $A_{n,M,L}$ that mirrors Example 2 and 3: it has highly coherent columns (see Figure 2.10 (b)). Consequently, the estimation of the coefficient is acutely sensitive to the choice of nonzero locations, requiring the identification of the accurate support of the ground truth.

In Figure 2.10 (c), across various mesh sizes defined as $\Delta x = \Delta y = \frac{4}{[22:2:40]}$, the PartInv algorithm with $K = 1$ produced accurate estimators by identifying the correct support set $\{1\}$, and therefore effectively tackles the data corruptions coming from discrete-time observations in

this challenging basis pursuit problem. For competitor methods, even with the training data where we have smallest mesh size $\Delta x = \Delta y = 0.1$, the CoSaMP (also subspace pursuit) and LASSO estimators are not accurate and even failed to find the right support set.

As we use an analytic solution for the underlying PDE, the only error source is coming from the discrete time observations. In Figure 2.10 (c), we found that the convergence rate of the relative error with respect to space-time mesh size aligns with our theoretical error analysis in Proposition 2.4.1. Given the absence of numerical error from the solver, the second-order convergence is achieved as we do not need to compute the time derivatives here. The robustness relative to the measurement noise was also tested, as depicted in Figure 2.10 (d).

Example 5 (2D nonlinear diffusion with nonlocal interaction) In this example, we explore the performance of our algorithm using very coarse scale data that are subjected to both discretization and numerical solver forward errors. We consider an initial condition and interaction potential given by

$$\rho_0(x) = 5\Big(\frac{e^{-((x+0.5)^2+(y+0.5)^2)}}{0.2} + \frac{e^{-((x-0.5)^2+(y-0.5)^2)}}{0.2}\Big)\,, \quad W(x) = -3e^{-2|x|^2}\,,$$

respectively. We set $m = 2$, $\kappa = 1$ and the rest of the computational parameters are summarized in Table 2.8.

δt	δx	Time domain	Spatial domain	$\frac{\phi(\lvert x\rvert)}{\lvert x\rvert}$
10^{-3}	$2 * 10^{-1}$	$[0, 0.05]$	$[-2.1, 2.1] \times [-2.1, 2.1]$	$12e^{-2\lvert x\rvert^2}$

Table 2.8: Parameters to produce the solution data using the finite volume scheme.

(a) (b)

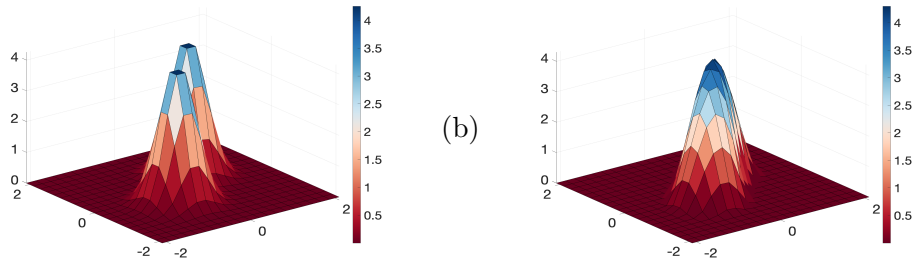
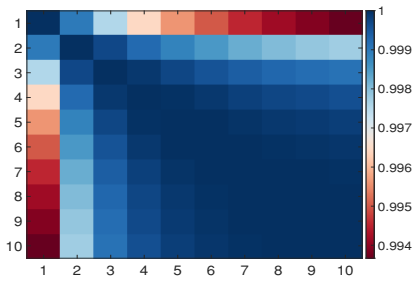


Figure 2.11: (a)-(b): the profile of the solution at $t = 0$ and $t = 0.05$.

To estimate the interaction kernel as in Table 2.8, we use an exponential basis of the form $\{-2w\exp(-w|x|^2) : w = 1 : 1 : 10\}$. Then the true interaction kernel is 1-sparse with respect to this particular basis representation. Figure 2.12 (c) shows that it yields a very coherent basis in our sparse learning problem. In this example, we use solution data defined on a coarse mesh

(c) 

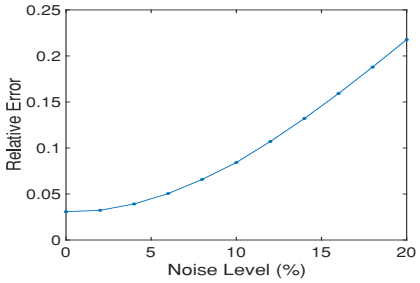
(d) 

Figure 2.12: (c) The coherence pattern of the regression matrix $A_{n,M,L}$. (d) We run PartInv with $K = 2$ and used the support pruning algorithm for kernel estimation. For each noise level, we run 100 trials and display the relative reconstruction error bar versus noise levels but variances are very small in this set of experiments and therefore are not visible in the current scale.

as shown in Figure 2.11 (a) and (b) corresponding to different times. When we set the $K = 1$, PartInv failed to find the right support. We then set $K = 2$, and run our support pruning algorithm setting $\widehat{\Delta}t = 10^{-4}$ and $\widehat{\Delta}x = 10^{-1}$. We observed that under a variety of noise levels, our algorithm enables accurate estimation thanks to the support pruning step and consistently found the right support. See the results in Figure 2.12 (d).

2.6 Conclusion and future work

In this study we present a new sparse identification algorithm designed to estimate the nonlocal interaction kernel within a broad spectrum of nonlocal gradient flow equations using noisy and discrete data. We establish new stability estimates that demonstrate the ability of our learned estimator to accurately reflect the training data. Additionally, we conduct an error analysis of our estimators and elucidate the dependency of their accuracy on factors such as the noise level and the mesh discretization. When compared to alternative sparse regression algorithms, our PartInv algorithm stands out for its simplicity in implementation and hyperparameter tuning while effectively addressing coherent regression matrices. It surpasses other methods like LASSO, subspace pursuit, and SINDy in performance. The main constraint is the prerequisite of selecting a suitable basis that ensures exact sparsity in the interaction kernel.

Future work will delve into the development of robust techniques tailored to solution data for aggregation diffusion equations. This will include the exploration of advanced denoising techniques and regularization algorithms, such as sparse Bayesian methods, aimed at reducing the dependence of the prior knowledge on the estimated kernels. Another direction of future work is to extend the current algorithm to cover systems with multiple interaction kernels, which models the heterogeneous interactions in multi-species systems.

Appendix A

Proofs of Dobrushin-type stability estimates

A.0.1 Proof of Proposition 2.3.1

Proof. Recalling standard results [AGS05], it is known that, given our assumptions on W, $\widehat{W}, V$ and $\widehat{V}$, the solutions of (2.3.2) are of the form $\mu_t = \Phi_t \# \mu_0$, $\widehat{\mu}_t = \widehat{\Phi}_t \# \widehat{\mu}_0$, where Φ_t, $\widehat{\Phi}_t$ are the flow maps induced by the velocity fields $\nabla W * \mu_t + \nabla V$ and $\nabla \widehat{W} * \widehat{\mu}_t + \nabla \widehat{V}$, respectively. Then we have the following estimate

$$
\begin{aligned}
d_2^2(\mu_t, \widehat{\mu}_t) &= d_2^2(\Phi_t \# \mu_0, \widehat{\Phi}_t \# \widehat{\mu}_0) && \text{(A.0.1)} \\
&\leqslant d_2^2(\Phi_t \# \mu_0, \widehat{\Phi}_t \# \mu_0) + d_2^2(\widehat{\Phi}_t \# \mu_0, \widehat{\Phi}_t \# \widehat{\mu}_0) \\
&\leqslant \int_{\mathbb{R}^d} |\Phi_t(x) - \widehat{\Phi}_t(x)|^2 \mathrm{d}\mu_0(x) + d_2^2(\widehat{\Phi}_t \# \mu_0, \widehat{\Phi}_t \# \widehat{\mu}_0) \ . && \text{(A.0.2)}
\end{aligned}
$$

We can bound the final term above in the following way. Denote the product measure $\Pi_t := (\widehat{\Phi}_t \times \widehat{\Phi}_t) \# \Pi_0$, where Π_0 is the optimal transport plan between μ_0 and $\widehat{\mu}_0$. Then, by definition of the 2-Wasserstein metric we have that

$$
\begin{aligned}
d_2^2(\widehat{\Phi}_t \# \mu_0, \widehat{\Phi}_t \# \widehat{\mu}_0) &\leqslant \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \mathrm{d}\Pi_t = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\widehat{\Phi}_t(x) - \widehat{\Phi}_t(y)|^2 \mathrm{d}\Pi_0 \\
&\leqslant e^{2(L_W + L_V)t} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \mathrm{d}\Pi_0 \leqslant e^{2(L_W + L_V)t} d_2^2(\mu_0, \widehat{\mu}_0) \ .
\end{aligned}
$$

Where in the second line we used the Lipschitzness of the flow map $\widehat{\Phi}_t$. We then have the following estimate for the integrand in the first term of (A.0.1)

$$
\begin{aligned}
|\Phi_t(x)-\widehat{\Phi}_t(x)|^2 \leqslant & t\int_0^t |(\nabla W*\mu_s)(\Phi_s(x))+\nabla V(\Phi_s(x)) \\
& -(\nabla\widehat{W}*\widehat{\mu}_s)(\widehat{\Phi}_s(x))-\nabla\widehat{V}(\widehat{\Phi}_s(x))|^2\mathrm{d}s \\
\leqslant & 4t\int_0^t |(\nabla W*\mu_s)(\Phi_s(x))-(\nabla\widehat{W}*\mu_s)(\Phi_s(x))|^2\mathrm{d}s \\
& +4t\int_0^t |(\nabla\widehat{W}*\mu_s)(\Phi_s(x))-(\nabla\widehat{W}*\widehat{\mu}_s)(\widehat{\Phi}_s(x))|^2\mathrm{d}s \\
& +4t\int_0^t |\nabla V(\Phi_s(x))-\nabla\widehat{V}(\Phi_s(x))|^2\mathrm{d}s \\
& +4t\int_0^t |\nabla\widehat{V}(\Phi_s(x))-\nabla\widehat{V}(\widehat{\Phi}_s(x))|^2\mathrm{d}s \\
\leqslant & 4t\int_0^t \left|\int_{\mathbb{R}^d}\left[\nabla W(\Phi_s(x)-y)-\nabla\widehat{W}(\Phi_s(x)-y)\right]\mathrm{d}\mu_s(y)\right|^2\mathrm{d}s \\
& +8t\int_0^t \left|\int_{\mathbb{R}^d}\nabla\widehat{W}(\Phi_s(x)-y)\mathrm{d}\mu_s(y)-\nabla\widehat{W}(\Phi_s(x)-y)\mathrm{d}\widehat{\mu}_s(y)\right|^2\mathrm{d}s \\
& +8t\int_0^t \left|\int_{\mathbb{R}^d}\nabla\widehat{W}(\Phi_s(x)-y)\mathrm{d}\widehat{\mu}_s(y)-\nabla\widehat{W}(\widehat{\Phi}_s(x)-y)\mathrm{d}\widehat{\mu}_s(y)\right|^2\mathrm{d}s \\
& +4t\int_0^t |\nabla V(\Phi_s(x))-\nabla\widehat{V}(\Phi_s(x))|^2\mathrm{d}s \\
& +4tL_{\widehat{V}}^2\int_0^t |\Phi_s(x)-\widehat{\Phi}_s(x)|^2\mathrm{d}s \\
\leqslant & 4t\int_0^t \left|\int_{\mathbb{R}^d}\left[\nabla W(\Phi_s(x)-y)-\nabla\widehat{W}(\Phi_s(x)-y)\right]\mathrm{d}\mu_s(y)\right|^2\mathrm{d}s \\
& +8t\int_0^t \int_{\mathbb{R}^d}\left|\nabla\widehat{W}(\Phi_s(x)-y)(\mathrm{d}\mu_s(y)-\mathrm{d}\widehat{\mu}_s(y))\right|^2\mathrm{d}s \\
& +8t\int_0^t \left|\int_{\mathbb{R}^d}\nabla\widehat{W}(\Phi_s(x)-y)-\nabla\widehat{W}(\widehat{\Phi}_s(x)-y)\mathrm{d}\widehat{\mu}_s(y)\right|^2\mathrm{d}s \\
& +4t\int_0^t |\nabla V(\Phi_s(x))-\nabla\widehat{V}(\Phi_s(x))|^2\mathrm{d}s \\
& +4tL_{\widehat{V}}^2\int_0^t |\Phi_s(x)-\widehat{\Phi}_s(x)|^2\mathrm{d}s \\
\leqslant & 4t\int_0^t \left|\int_{\mathbb{R}^d}\left[\nabla W(\Phi_s(x)-y)-\nabla\widehat{W}(\Phi_s(x)-y)\right]\mathrm{d}\mu_s(y)\right|^2\mathrm{d}s \\
& +8tL_{\widehat{W}}^2\int_0^t e^{2(L_W+L_V)s}d_2^2(\mu_s,\widehat{\mu}_s)\mathrm{d}s \\
& +8tL_{\widehat{W}}^2\int_0^t \left|\Phi_s(x)-\widehat{\Phi}_s(x)\right|^2\mathrm{d}s \\
& +4t\int_0^t |\nabla V(\Phi_s(x))-\nabla\widehat{V}(\Phi_s(x))|^2\mathrm{d}s \\
& +4tL_{\widehat{V}}^2\int_0^t |\Phi_s(x)-\widehat{\Phi}_s(x)|^2\mathrm{d}s\ .
\end{aligned}
$$

Then, integrating with respect to μ_0 yields

$$\begin{aligned}
\int_{\mathbb{R}^d} |\Phi_t(x) - \widehat{\Phi}_t(x)|^2 \mathrm{d}\mu_0(x) \leqslant 4t \int_0^t \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} \Big[\nabla W(\Phi_s(x) - y) \\
- \nabla \widehat{W}(\Phi_s(x) - y) \Big] \mathrm{d}\mu_s(y) \Big|^2 \mathrm{d}\mu_0(x) \mathrm{d}s \\
+ 8tL_{\widehat{W}}^2 \int_0^t e^{2(L_W + L_V)s} d_2^2(\mu_s, \widehat{\mu}_s) \mathrm{d}s \\
+ 8tL_{\widehat{W}}^2 \int_0^t \int_{\mathbb{R}^d} \left| \Phi_s(x) - \widehat{\Phi}_s(x) \right|^2 \mathrm{d}\mu_0(x) \mathrm{d}s \\
+ 4t \int_0^t \int_{\mathbb{R}^d} |\nabla V(\Phi_s(x)) - \nabla \widehat{V}(\Phi_s(x))|^2 \mathrm{d}\mu_0(x) \mathrm{d}s \\
+ 4tL_{\widehat{V}}^2 \int_0^t \int_{\mathbb{R}^d} |\Phi_s(x) - \widehat{\Phi}_s(x)|^2 \mathrm{d}\mu_0(x) \mathrm{d}s \\
= 4t \int_0^t \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \left[\nabla W(x - y) - \nabla \widehat{W}(x - y) \right] \mathrm{d}\mu_s(y) \right|^2 \mathrm{d}\mu_s(x) \mathrm{d}s \\
+ 8tL_{\widehat{W}}^2 \int_0^t e^{2(L_W + L_V)s} d_2^2(\mu_s, \widehat{\mu}_s) \mathrm{d}s \\
+ 4t(2L_{\widehat{W}}^2 + L_{\widehat{V}}^2) \int_0^t \int_{\mathbb{R}^d} |\Phi_s(x) - \widehat{\Phi}_s(x)|^2 \mathrm{d}\mu_0(x) \mathrm{d}s \\
+ 4t \int_0^t \int_{\mathbb{R}^d} |\nabla V(\Phi_s(x)) - \nabla \widehat{V}(\Phi_s(x))|^2 \mathrm{d}\mu_0(x) \mathrm{d}s.
\end{aligned}$$

Then, an application of Grönwall's inequality yields

$$\begin{aligned}
\int_{\mathbb{R}^d} |\Phi_t(x) - \widehat{\Phi}_t(x)|^2 \mathrm{d}\mu_0(x) \leqslant \Big(8tL_{\widehat{W}}^2 \int_0^t e^{2(L_W + L_V)s} d_2^2(\mu_s, \widehat{\mu}_s) \mathrm{d}s \\
+ 8t \int_0^t \| \nabla W * \mu_s - \nabla \widehat{W} * \mu_s \|_{L^2(\mu_s)}^2 \mathrm{d}s \\
+ 4t \int_0^t \| \nabla V - \nabla \widehat{V} \|_{L^2(\mu_s)}^2 \mathrm{d}s \Big) e^{2(2L_{\widehat{W}}^2 + L_{\widehat{V}}^2)t^2} .
\end{aligned}$$

Then going back to our original estimate in (A.0.1) we have after an additional application of Grönwall's lemma that

$$d_2^2(\mu_t, \widehat{\mu}_t) \leqslant C_1 \int_0^t \| \nabla W * \mu_s - \nabla \widehat{W} * \mu_s \|_{L^2(\mu_s)}^2 \mathrm{d}s + C_2 \int_0^t \| \nabla V - \nabla \widehat{V} \|_{L^2(\mu_s)}^2 \mathrm{d}s + C_3 d_2^2(\mu_0, \widehat{\mu}_0)$$

where C_1, C_2 and C_3 are non-negative constants that depend on $T, L_W, L_V, L_{\widehat{W}}$ and $L_{\widehat{V}}$. Recalling the definition of the error functional $\tilde{\mathcal{E}}_\infty$ concludes the proof. □

A.0.2 Assumptions and proof of Proposition 2.3.3

Equation (A.0.3) can be interpreted as the evolution of the law of the solution of a stochastic differential equation (SDE) [Szn91, Car16]. Namely, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space and let $(B_t)_{t\in[0,T]}$ be an adapted Brownian motion in $\mathbb{R}^d$. Let us denote by

$$\mathbb{H}^{2,k} := \left\{ Z : [0,T] \times \mathbb{R}^d \to \mathbb{R}^k \mid Z \text{ is progressively measurable, } \mathbb{E}\int_0^T |Z_s|^2 \, ds < \infty \right\} .$$

Then, (2.3.7) can be interpreted as the evolution of the law of the solution of the following SDE [Car16]

$$\begin{aligned} dX_t &= \nabla W * \mu_t(X_t)dt + \sqrt{2}\sigma(K * \mu_t(X_t))dB_t \ , \\ X_0 &= X^0 \in L^2 \text{ independent of } (B_t)_{t\in[0,T]} \ , \end{aligned} \tag{A.0.3}$$

where $\mu_t = \mathcal{L}(X_t)$ denotes the law of solution of (A.0.3) at time $t \in [0,T]$. We will need the following assumptions

Assumption 4.

1. $W \in \mathcal{W}^{2,\infty}(\mathbb{R}^d)$, $\sigma : \mathbb{R}^n \to \mathbb{R}^d \times \mathbb{R}^d$ and $K : \mathbb{R}^d \to \mathbb{R}^n$ are Lipschitz and bounded.

2. For any $\nu \in \mathcal{P}(\mathbb{R}^d)$, $(\nabla W * \nu(X_t))_{t\in[0,T]} \in \mathbb{H}^{2,d}$ and $(\sigma(K * \nu(X_t)))_{t\in[0,T]} \in \mathbb{H}^{2,d\times d}$.

With these assumptions we are now ready to prove Proposition 2.3.3. We note that for ease of notation, throughout the proof we use subscripts to denote the time argument of functions that depend on time.

Proof. We begin by noting that by the definition of the 2-Wasserstein distance we have

$$d_2^2(\mu_t, \widehat{\mu}_t) \leqslant \mathbb{E} \sup_{s\in[0,t]} \left| X_s - \widehat{X}_s \right|^2 \ . \tag{A.0.4}$$

Next, we consider the following estimate

$$\begin{aligned} \mathbb{E} \sup_{s\in[0,t]} \left| X_s - \widehat{X}_s \right|^2 = 3\mathbb{E} \sup_{s\in[0,t]} \Bigg(& \left| X_0 - \widehat{X_0} \right|^2 \\ & + \left| \int_0^s \left(\nabla W * \mu_r(X_r) - \nabla \widehat{W} * \widehat{\mu}_r(\widehat{X_r}) \right) dr \right|^2 \\ & + \left| \int_0^s \sigma(K * \mu_r(X_r)) - \widehat{\sigma}(\widehat{K} * \widehat{\mu}_r(\widehat{X}_r)) dB_r \right|^2 \Bigg) \\ =: & \ I + II + III \ . \end{aligned}$$

We then have

$$
\begin{aligned}
II \leqslant & 3\mathbb{E} \sup_{s\in[0,t]} s\int_0^s \left|\nabla W * \mu_r(X_r) - \nabla \widehat{W} * \widehat{\mu}_r(\widehat{X_r})\right|^2 \mathrm{d}r \\
\leqslant & 9T\mathbb{E} \sup_{s\in[0,t]} \Bigg(\int_0^s \left|\nabla W * \mu_r(X_r) - \nabla \widehat{W} * \mu_r(X_r)\right|^2 \mathrm{d}r \\
& + \int_0^s \left|\nabla \widehat{W} * \mu_r(X_r) - \nabla \widehat{W} * \widehat{\mu}_r(X_r)\right|^2 \mathrm{d}r \\
& + \int_0^s \left|\nabla \widehat{W} * \widehat{\mu}_r(X_r) - \nabla \widehat{W} * \widehat{\mu}_r(\widehat{X_r})\right|^2 \mathrm{d}r \Bigg) \\
\leqslant & 9T \int_0^t \|\nabla W * \mu_s - \nabla \widehat{W} * \mu_s\|^2_{L^2(\mu_s)} \mathrm{d}s \\
& + 9TL^2_{\widehat{W}} \int_0^t d_2^2(\mu_s, \widehat{\mu}_s)\mathrm{d}s + 9TL^2_{\widehat{W}} \int_0^t \mathbb{E} \sup_{r\in[0,s]} |X_r - \widehat{X_r}|^2 \mathrm{d}s \ .
\end{aligned}
$$

For the third term, using the Burkholder-Davis-Gundy inequality we have

$$
\begin{aligned}
III = & 6\mathbb{E} \sup_{s\in[0,t]} \left|\int_0^s \sigma(K * \mu_r(X_r)) - \widehat{\sigma}(\widehat{K} * \widehat{\mu}_r(\widehat{X}_r))\mathrm{d}B_r\right|^2 \\
\leqslant & 6CT\mathbb{E} \int_0^t \left|\sigma(K * \mu_s(X_s)) - \widehat{\sigma}(\widehat{K} * \widehat{\mu}_s(\widehat{X}_s))\right|^2 \mathrm{d}s \\
\leqslant & 24CT\mathbb{E} \int_0^t \left|\sigma(K * \mu_s(X_s)) - \sigma(\widehat{K} * \mu_s(X_s))\right|^2 \mathrm{d}s \\
& + 24CT\mathbb{E} \int_0^t \left|\sigma(\widehat{K} * \mu_s(X_s)) - \sigma(\widehat{K} * \widehat{\mu}_s(X_s))\right|^2 \mathrm{d}s \\
& + 24CT\mathbb{E} \int_0^t \left|\sigma(\widehat{K} * \widehat{\mu}_s(X_s)) - \sigma(\widehat{K} * \widehat{\mu}_s(\widehat{X}_s))\right|^2 \mathrm{d}s \\
& + 24CT\mathbb{E} \int_0^t \left|\sigma(\widehat{K} * \widehat{\mu}_s(\widehat{X}_s)) - \widehat{\sigma}(\widehat{K} * \widehat{\mu}_s(\widehat{X_s}))\right|^2 \mathrm{d}s \\
=: & (i) + (ii) + (iii) + (iv) \ .
\end{aligned}
$$

Next, we have the following estimates for $(i), (ii), (iii)$ and (iv).

$$
(i) \leqslant 24CTL^2_\sigma \int_0^t \|K * \mu_s - \widehat{K} * \mu_s\|^2_{L^2(\mu_s)} \mathrm{d}s \ .
$$

For (ii) we have

$$
\begin{aligned}
(ii) \leqslant & 24CTL^2_\sigma \mathbb{E} \int_0^t \left|\int_{\mathbb{R}^d} \widehat{K}(X_s - y)\mathrm{d}\mu(y) - \int_{\mathbb{R}^d} \widehat{K}(X_s - z)\mathrm{d}\widehat{\mu}(z)\right|^2 \mathrm{d}s \\
\leqslant & 24CTL^2_\sigma L^2_{\widehat{K}} \int_0^t d_2^2(\mu_s, \widehat{\mu}_s)\mathrm{d}s \ .
\end{aligned}
$$

Term (iii) can be bounded by

$$(iii) \leqslant 24CTL_\sigma^2 L_{\widehat{K}}^2 \int_0^t \sup_{r\in[0,s]} \mathbb{E}|X_r - \widehat{X}_r|^2 \mathrm{d}s \ ,$$

and finally, we have the following bound for term (iv)

$$(iv) \leqslant 24CT^2 \|\sigma - \widehat{\sigma}\|_\infty^2.$$

Putting all the previous estimates together we get

$$\begin{aligned}
\mathbb{E} \sup_{s\in[0,t]} \left|X_s - \widehat{X}_s\right|^2 \leqslant & 3\mathbb{E}\left|X_0 - \widehat{X}_0\right|^2 + 9T \int_0^t \|\nabla W * \mu_s - \nabla \widehat{W} * \mu_s\|_{L^2(\mu_s)}^2 \mathrm{d}s \\
& + 9TL_{\widehat{W}}^2 \int_0^t d_2^2(\mu_s, \widehat{\mu}_s)\mathrm{d}s + 9TL_{\widehat{W}}^2 \int_0^t \mathbb{E} \sup_{r\in[0,s]} |X_r - \widehat{X_r}|^2 \mathrm{d}s \\
& + 24CTL_\sigma^2 \int_0^t \|K * \mu_s - \widehat{K} * \mu_s\|_{L^2(\mu_s)}^2 \mathrm{d}s \\
& + 24CTL_\sigma^2 L_{\widehat{K}}^2 \int_0^t d_2^2(\mu_s, \widehat{\mu}_s)\mathrm{d}s + 24CT^2\|\sigma - \widehat{\sigma}\|_\infty^2 \\
& + 24CTL_\sigma^2 L_{\widehat{K}}^2 \int_0^t \sup_{r\in[0,s]} \mathbb{E}|X_r - \widehat{X}_r|^2 \mathrm{d}s \, .
\end{aligned}$$

A first application of Grönwall's inequality then yields

$$\begin{aligned}
\mathbb{E} \sup_{s\in[0,t]} \left|X_s - \widehat{X}_s\right|^2 \leqslant & \left(3\mathbb{E}\left|X_0 - \widehat{X}_0\right|^2 + \left(9L_{\widehat{W}}^2 T + T24CL_\sigma^2 L_{\widehat{K}}^2\right) \int_0^t d_2^2\left(\widehat{\mu}_s, \mu_s\right) \mathrm{d}s \right. \\
& + 9T \int_0^t \|\nabla W * \mu_s - \nabla \widehat{W} * \mu_s\|_{L^2(\mu_s)}^2 \mathrm{d}s + 24CT^2\|\sigma - \widehat{\sigma}\|_\infty^2 \\
& \left. + 24CTL_\sigma^2 \int_0^t \|K * \mu_s - \widehat{K} * \mu_s\|_{L^2(\mu_s)}^2 \mathrm{d}s\right) \exp\left(12CL_\sigma^2 L_{\widehat{K}}^2 T^2 + \frac{9}{2}T^2 L_{\widehat{W}}^2\right).
\end{aligned}$$

Then, using (A.0.4) and a final application Grönwall's inequality yields

$$\begin{aligned}
d_2^2(\mu_t, \widehat{\mu}_t) \leqslant & C(T, L_\sigma, L_{\widehat{K}}, L_{\widehat{W}})\left(3d_2^2(\mu_0, \widehat{\mu}_0) + 9T \int_0^t \|\nabla W * \mu_s - \nabla \widehat{W} * \mu_s\|_{L^2(\mu_s)}^2 \mathrm{d}s \right. \\
& \left. + 24CT^2\|\sigma - \widehat{\sigma}\|_\infty^2 + 24CTL_\sigma^2 \int_0^t \|K * \mu_s - \widehat{K} * \mu_s\|_{L^2(\mu_s)}^2 \mathrm{d}s\right),
\end{aligned}$$

where $C(T, L_\sigma, L_{\widehat{K}}, L_{\widehat{W}}) > 0$ is a constant depending on the Lipschitz coefficients of $\sigma, \widehat{K}$ and $\widehat{W}$, and T. Using the definition of the error functional $\tilde{\mathcal{E}}_\infty$ concludes the proof. □

A.0.2.1 Assumptions and proof of Proposition 2.3.4

Let us denote by $\Pi(\sigma,\nu) \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ the set of transport plans between $\sigma \in \mathcal{P}(\mathbb{R}^d)$ and $\nu \in \mathcal{P}(\mathbb{R}^d)$ for the quadratic cost function. As it will become clear later on, we will need the following estimate on the energy $H(\rho) := \int_{\mathbb{R}^d} h(\rho(x))\mathrm{d}x$.

Lemma A.0.1. Let $h : [0,+\infty] \to \mathbb{R}$ be the internal energy density given by $h(z) = \kappa\frac{z^m}{m-1}$ where $m \neq 1$, $m \geqslant 1 - \frac{1}{d}$ and $m > \frac{d}{d+2}$. Let $\rho, \widetilde{\rho}$ be two smooth solutions to (2.3.9) and let $\gamma_0 \in \Pi(\rho_0, \widetilde{\rho}_0)$. Then, we have the following estimate

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} (\nabla h'(\widetilde{\rho}(0,y)) - \nabla h'(\rho(0,x))) \cdot (y-x)\gamma_0(\mathrm{d}x,\mathrm{d}y) \geqslant 0 \ . \tag{A.0.5}$$

Proof. Following [Ott01, Section 5.2] (see also [ABH+13, Proposition 3.38]), the functional $H(\rho)$ is displacement convex. As a consequence, we deduce the following estimate for the difference of the internal energy for two smooth solutions $\rho, \widetilde{\rho}$ of (2.3.9)

$$H(\widetilde{\rho}_0) - H(\rho_0) \geqslant \int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla h'(\rho(0,x)) \cdot (y-x)\gamma_0(\mathrm{d}x,\mathrm{d}y) \ , \tag{A.0.6}$$

where γ_0 is a transference plan between ρ and $\widetilde{\rho}$. Notice this is nothing else than the characterization of convexity by supporting hyperplanes. By symmetry we also have

$$H(\rho_0) - H(\widetilde{\rho}_0) \geqslant -\int_{\mathbb{R}^d\times\mathbb{R}^d} \nabla h'(\widetilde{\rho}(0,y)) \cdot (y-x)\gamma_0(\mathrm{d}x,\mathrm{d}y) \ . \tag{A.0.7}$$

Adding (A.0.6) and (A.0.7) yields

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} (\nabla h'(\widetilde{\rho}(0,y)) - \nabla h'(\rho(0,x)) \cdot (y-x)\gamma_0(\mathrm{d}x,\mathrm{d}y) \geqslant 0 \ .$$

□

Furthermore, we note that by our assumption, $\rho, \widehat{\rho}$ are smooth solutions to continuity equations and hence, by [AGS05, Proposition 8.1.8] they admit the representations $\rho_\tau = \Phi_{\tau\#}\rho_0$ and $\widehat{\rho}_\tau = \widehat{\Phi}_{\tau\#}\widehat{\rho}_0$, where Φ is the flow map associated to the problem

$$\begin{aligned} \frac{d}{d\tau} r(\tau,x) &= v(\tau,r), \\ r(0) &= x \in \mathbb{R}^d, \end{aligned}$$

and $\widehat{\Phi}$ is the flow map for the analogous problem with the velocity field $\widehat{v}(\widehat{\rho}) = -\nabla(H'(\widehat{\rho}) + \widehat{W} * \widehat{\rho} + \widehat{V})$. Let $\gamma_0 \in \Pi(\rho_0, \widehat{\rho}_0)$ be an optimal transport plan between ρ_0 and $\widehat{\rho}_0$ and note that, for

$\tau \in (0,T]$ by the representations of ρ_τ and $\widehat{\rho}_\tau$, we have that $\gamma_\tau = (\Phi_\tau \times \widehat{\Phi}_\tau)_{\#}\gamma_0$ is an admissible transport plan between $\rho_\tau, \widehat{\rho}_\tau$. Now we are ready to present the proof of Proposition 2.3.4.

Proof. Closely following arguments similar to [Ott01, Section 5.2] (see also [San15, Corollary 5.2.5]), we have that by the representation of the solution $\rho_t, \widehat{\rho}_t$ in terms of the associated flows $\Phi_t, \widehat{\Phi}_t$ we can obtain the following estimate

$$\begin{aligned}\frac{1}{t}\left(d_2^2(\rho_t,\widehat{\rho}_t) - d_2^2(\rho_0,\widehat{\rho}_0)\right) &\leqslant \frac{1}{t}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |y-x|^2 \,\mathrm{d}\gamma_t(x,y) - \int_{\mathbb{R}^d\times\mathbb{R}^d} |y-x|^2 \,\mathrm{d}\gamma_0(x,y)\right)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{t}\left(\left|\widehat{\Phi}(t,y) - \Phi(t,x)\right|^2 - |y-x|^2\right)\mathrm{d}\gamma_0(x,y)\ .\end{aligned}$$

Then, letting $t \to 0^+$ and by the definition of the flow maps $\Phi, \widehat{\Phi}$ we obtain

$$\left.\frac{d^+}{d\tau}\right|_{\tau=0} d_2^2(\rho_t,\widehat{\rho}_t) \leqslant 2\int_{\mathbb{R}^d\times\mathbb{R}^d} (\widehat{v}(0,y) - v(0,x))\cdot(y-x)\,\mathrm{d}\gamma_0(x,y)\,,$$

Using our estimate from Lemma A.0.1 and integrating in time yields

$$\begin{aligned}d_2^2(\rho_t,\widehat{\rho}_t) &\leqslant d_2^2(\rho_0,\widehat{\rho}_0) + 2\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} (\widehat{v}(\widehat{\rho})(s,y) - v(\rho)(s,x))\cdot(y-x)\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\leqslant -2\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} (\nabla\widehat{V}(y) - \nabla V(x))\cdot(y-x)\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\quad - 2\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} (\nabla\widehat{W} * \widehat{\rho}_s(y) - \nabla W * \rho_s(x))\cdot(y-x)\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\quad + d_2^2(\rho_0,\widehat{\rho}_0)\ .\end{aligned}$$

Taking absolute value on both sides of the previous estimate and an application of Young's inequality yields

$$\begin{aligned}d_2^2(\rho_t,\widehat{\rho}_t) &\leqslant \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} |\nabla\widehat{V}(y) - \nabla V(x)|^2\mathrm{d}\gamma_s(x,y) + \int_{\mathbb{R}^d\times\mathbb{R}^d} |y-x|^2\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\quad + \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} |(\nabla\widehat{W} * \widehat{\rho}_s(y) - \nabla W * \rho_s(x))|^2\mathrm{d}\gamma_s(x,y) + \int_{\mathbb{R}^d\times\mathbb{R}^d} |y-x|^2\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\quad + d_2^2(\rho_0,\widehat{\rho}_0)\\ &\leqslant \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} |\nabla\widehat{V}(y) - \nabla V(x)|^2\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\quad + \int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} |(\nabla\widehat{W} * \widehat{\rho}_s(y) - \nabla W * \rho_s(x))|^2\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ &\quad + 2\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d} |y-x|^2\mathrm{d}\gamma_s(x,y)\mathrm{d}s + d_2^2(\rho_0,\widehat{\rho}_0)\ .\end{aligned}$$

With similar calculations to the ones in Proposition 2.3.1 we have that

$$\begin{aligned}|\nabla\widehat{V}(y)-\nabla V(x)|^2 &\leqslant 2|\nabla\widehat{V}(y)-\nabla\widehat{V}(x)|^2+2|\nabla\widehat{V}(x)-\nabla V(x)|^2\\ &\leqslant 2L_{\widehat{V}}^2|y-x|^2+2|\nabla\widehat{V}(x)-\nabla V(x)|^2\ ,\end{aligned}$$

and integrating with respect to γ_s yields

$$\int_{\mathbb{R}^d\times\mathbb{R}^d}|\nabla\widehat{V}(y)-\nabla V(x)|^2\mathrm{d}\gamma_s(x,y)\leqslant 2L_{\widehat{V}}^2\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-x|^2\mathrm{d}\gamma_s(x,y)+2\|\nabla V-\nabla\widehat{V}\|_{L^2(\rho_s)}^2\ .$$

Similarly, for the term involving the interaction potential we have

$$\begin{aligned}|\nabla\widehat{W}*\widehat{\rho}_s(y)-\nabla W*\rho_s(x)|^2 \leqslant&|\nabla\widehat{W}*\widehat{\rho}_s(y)-\nabla\widehat{W}*\rho_s(x)|^2\\ &+2|\nabla\widehat{W}*\rho_s(x)-\nabla W*\rho_s(x)|^2\\ \leqslant&\ 2L_{\widehat{W}}^2|y-x|^2+2|\nabla\widehat{W}*\rho_s(x)-\nabla W*\rho_s(x)|^2\ .\end{aligned}$$

Integrating with respect to γ_s then yields

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d}|\nabla\widehat{W}*\widehat{\rho}_s(y)-\nabla W*\rho_s(x)|^2\mathrm{d}\gamma_s(x,y)\leqslant 2L_{\widehat{W}}^2\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-x|^2\mathrm{d}\gamma_s(x,y)\\ +2\|\nabla\widehat{W}*\rho_s-\nabla W*\rho_s\|_{L^2(\rho_s)}^2\ .\end{aligned}$$

Hence, all in all we have

$$\begin{aligned}d_2^2(\widehat{\rho}_t,\rho_t)\leqslant d_2^2(\widehat{\rho}_0,\rho_0)+2\int_0^t(1+L_{\widehat{V}}^2+L_{\widehat{W}}^2)\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-x|^2\mathrm{d}\gamma_s(x,y)\mathrm{d}s\\ +2T\tilde{\mathcal{E}}_\infty(\widehat{W})+2\int_0^t\|\nabla V-\nabla\widehat{V}\|_{L^2(\rho_s)}^2\mathrm{d}s\ .\end{aligned}$$

Taking the infimum with respect to $\gamma_s\in\Pi(\widehat{\rho}_s,\rho_s)$ in the previous equation gives

$$\begin{aligned}d_2^2(\widehat{\rho}_t,\rho_t)\leqslant d_2^2(\widehat{\rho}_0,\rho_0)+2\int_0^t(1+L_{\widehat{V}}^2+L_{\widehat{W}}^2)d_2^2(\widehat{\rho}_s,\rho_s)\mathrm{d}s\\ +2T\tilde{\mathcal{E}}_\infty(\widehat{W})+2\int_0^t\|\nabla V-\nabla\widehat{V}\|_{L^2(\rho_s)}^2\mathrm{d}s\ .\end{aligned}$$

An application of Grönwall's inequality then yields

$$\begin{aligned}d_2^2(\rho_t,\widehat{\rho}_t)\leqslant&\exp\{2(1+L_{\widehat{V}}^2+L_{\widehat{W}}^2)t\}\\ &\times\left(d_2^2(\widehat{\rho}_0,\rho_0)+2T\tilde{\mathcal{E}}_\infty(\widehat{W})+2\int_0^t\|\nabla V-\nabla\widehat{V}\|_{L^2(\rho_s)}^2\mathrm{d}s\right)\ ,\end{aligned}$$

which concludes the proof. □

Appendix B

Error estimate with noisy data

Proof of Proposition 2.4.2

Proof. We need to estimate $\mathcal{B} := A_{n,M,L}(i,j) - \widetilde{A}_{n,M,L}(i,j)$. Note that $\mathcal{B}$ does not depend on the indices $(i,j) \in \{1,\dots,n\}^2$ as our final estimate is independent of the indices considered. Then we have

$$\mathcal{B} = \frac{1}{T} \sum_{\ell=1,m=-M}^{L,M} [(C^i_{n,M,L} \cdot C^j_{n,M,L})^\ell_m \rho^\ell_m - (\tilde{C}^i_{n,M,L} \cdot \tilde{C}^j_{n,M,L})^\ell_m \tilde{\rho}^\ell_m] \Delta x \Delta t \,.$$

Define $\mathcal{C}^i_m(g^\ell) = \sum_{k=-M}^{M} (\nabla \Psi_i)_{m-k} g(t_\ell, x_k) \Delta x$ for any function g defined on the mesh. Then we can write $\mathcal{B}$ using this notation to emphasize the dependence of $\tilde{C}^i_{n,M,L}$ on the noise as follows

$$\mathcal{B} = \frac{1}{T} \sum_{\ell=1,m=-M}^{L,M} [(C^i_{n,M,L} \cdot C^j_{n,M,L})^\ell_m \rho^\ell_m - \mathcal{C}^i_m(\rho^\ell + \varepsilon^\ell)\mathcal{C}^j_m(\rho^\ell + \varepsilon^\ell)(\rho^\ell_m + \varepsilon^\ell_m)] \Delta x \Delta t \,.$$

By linearity of $\mathcal{C}^i_m$, expanding the second term yields

$$|\mathcal{B}| = |\Lambda^{i,j}| \,,$$

where $\Lambda^{i,j}$ is composed of the terms in $\mathcal{B}$ that depend on the noise and is given by

$$\begin{aligned}
\Lambda^{i,j} &= \frac{1}{T} \sum_{v=1}^{7} \Lambda^{i,j}_v \\
&= \frac{1}{T} \sum_{\ell=1,m=-M}^{L,M} \Big(\mathcal{C}^i_m(\varepsilon^\ell)\mathcal{C}^j_m(\rho^\ell)\rho^\ell_m + \mathcal{C}^i_m(\rho^\ell)\mathcal{C}^j_m(\rho^\ell)\varepsilon^\ell_m \\
&\quad + \mathcal{C}^i_m(\varepsilon^\ell)\mathcal{C}^j_m(\varepsilon^\ell)\rho^\ell_m + \mathcal{C}^i_m(\rho^\ell)\mathcal{C}^j_m(\varepsilon^\ell)\rho^\ell_m \\
&\quad + \mathcal{C}^i_m(\varepsilon^\ell)\mathcal{C}^j_m(\rho^\ell)\varepsilon^\ell_m + \mathcal{C}^i_m(\rho^\ell)\mathcal{C}^j_m(\varepsilon^\ell)\varepsilon^\ell_m
\end{aligned}$$

$$+ \mathcal{C}_m^i(\varepsilon^\ell)\mathcal{C}_m^j(\varepsilon^\ell)\varepsilon_m^\ell\Big)\Delta x\Delta t \ .$$

For brevity, we only display the estimates for the terms yielding the error order reported in Proposition 2.4.2 as the rest of the terms are of higher order. Namely, we will consider the terms

$$\Lambda_1^{i,j} = \sum_{\ell=1,m=-M}^{L,M} \mathcal{C}_m^i(\varepsilon^\ell)\mathcal{C}_m^j(\rho^\ell)\rho_m^\ell\Delta x\Delta t \ ,$$
$$\Lambda_2^{i,j} = \sum_{\ell=1,m=-M}^{L,M} \mathcal{C}_m^i(\rho^\ell)\mathcal{C}_m^j(\rho^\ell)\varepsilon_m^\ell\Delta x\Delta t \ ,$$

and

$$\Lambda_3^{i,j} = \sum_{\ell=1,m=-M}^{L,M} \mathcal{C}_m^i(\varepsilon^\ell)\mathcal{C}_m^j(\varepsilon^\ell)\rho_m^\ell\Delta x\Delta t \ .$$

Note that $\Lambda_4^{i,j}$ is analogous to $\Lambda_1^{i,j}$, so it will have the same error. We begin with the estimate for $\Lambda_1^{i,j}$ given by

$$\|\Lambda_1^{i,j}\|_{L^2(\varepsilon)} = \left(\mathbb{E}\left|\sum_{\ell=1,m=-M}^{L,M} \mathcal{C}_m^i(\varepsilon^\ell)\mathcal{C}_m^j(\rho^\ell)\rho_m^\ell\Delta x\Delta t\right|^2\right)^{\frac{1}{2}} .$$

By expanding the square, using the triangle inequality and noting that the expectation of terms for different points in the time mesh $\ell_1 \neq \ell_2$ for $\ell_1, \ell_2 \in \{1, \dots, L\}$ vanishes, we obtain

$$\begin{aligned}\|\Lambda_1^{i,j}\|_{L^2(\varepsilon)} \leqslant & \left(\sum_{\ell=1,m=-M}^{L,M} \mathbb{E}\left|(\mathcal{C}_m^i(\varepsilon^\ell))^2(\mathcal{C}_m^j(\rho^\ell)\rho_m^\ell)^2\Delta x^2\Delta t^2\right|\right)^{\frac{1}{2}} \\ & + \left(\sum_{\ell=1,m_1\neq m_2}^{L,M} \mathbb{E}\left|(\mathcal{C}_{m_1}^i(\varepsilon^\ell)\mathcal{C}_{m_1}^j(\rho^\ell)\rho_{m_1}^\ell)(\mathcal{C}_{m_2}^i(\varepsilon^\ell)\mathcal{C}_{m_2}^j(\rho^\ell)\rho_{m_2}^\ell)\Delta x^2\Delta t^2\right|\right)^{\frac{1}{2}} . \\ = & \ (a) + (b) \ .\end{aligned}$$

Note that for any $m \in \{-M, \dots, M\}$ and any $\ell \in \{1, \dots, L\}$ we have the following estimate

$$|(\mathcal{C}_m^i(\rho^\ell))^2| = \left|\sum_{k=-M}^{M} (\nabla\Psi_i)_{m-k}^2(\rho_k^\ell)^2\Delta x^2 + \sum_{k_1\neq k_2}^{M} (\nabla\Psi_i)_{m-k_1}(\nabla\Psi_i)_{m-k_2}\rho_{k_1}^\ell\rho_{k_2}^\ell\Delta x^2\right| \leqslant C \ , \tag{B.0.1}$$

for some positive constant $C = C(\|\rho\|_\infty, \|\nabla\Psi\|_\infty)$. In the estimates that follow we will write C to denote a generic positive constant which can depend on $R, T, \|\rho\|_\infty, \|\Psi\|_\infty, \|\nabla\Psi\|_\infty$ and can

change from line to line. Recalling that we defined $M = 2R/\Delta x$ and $L = T/\Delta t$, we have

$$\begin{aligned}(a) &\leqslant C\left(\frac{2R}{\Delta x}\frac{T}{\Delta t}\left(\frac{2R}{\Delta x}\|\nabla\Psi_i\|_\infty^2\sigma^2\Delta x^2\right)\Delta x^2\Delta t^2\right)^{1/2}\\ &\leqslant C\sigma\Delta x\Delta t^{1/2}\ .\end{aligned}$$

For (b) we begin by noting that, similarly to the estimate (B.0.1) we have $|(\mathcal{C}^j_{m_1}(\rho^\ell))(\mathcal{C}^j_{m_2}(\rho^\ell))| \leqslant C = C(\|\rho\|_\infty, \|\nabla\Psi\|_\infty)$ for any $m_1, m_2 \in \{-M, \dots, M\}$ and $\ell \in \{1, \dots, L\}$. Then we have that

$$\begin{aligned}(b) &\leqslant C\left(\frac{4R^2}{\Delta x^2}\frac{T}{\Delta t}\left(\frac{2R}{\Delta x}\|\nabla\Psi_i\|^2\sigma^2\Delta x^2\right)\Delta x^2\Delta t^2\right)^{1/2}\\ &\leqslant C\sigma\sqrt{\Delta x\Delta t}\ ,\end{aligned}$$

from which $\|\Lambda_1^{i,j}\|_{L^2(\varepsilon)} \leqslant C\sigma\sqrt{\Delta x\Delta t}$ follows. The calculations for $\Lambda_2^{i,j}$ are analogous. The only term that survives after expanding the square, using the triangle inequality and taking expectation is

$$\|\Lambda_2^{i,j}\|_{L^2(\varepsilon)} = \left(\sum_{\ell=1,m=-M}^{L,M}\mathbb{E}|\mathcal{C}^i_m(\rho^\ell)\mathcal{C}^j_m(\rho^\ell)\varepsilon^\ell_m|^2\Delta x^2\Delta t^2\right)^{1/2}$$

Then we have the following estimate

$$\begin{aligned}\|\Lambda_2^{i,j}\|_{L^2(\varepsilon)} &\leqslant \left(C\frac{2RT}{\Delta x\Delta t}\sigma^2\Delta x^2\Delta t^2\right)^{1/2}\\ &\leqslant\ C\sigma\sqrt{\Delta x\Delta t}\ .\end{aligned}$$

Finally, for term $\Lambda_3^{i,j}$ we have

$$\begin{aligned}\|\Lambda^{i,j}\|_{L^2(\varepsilon)} \leqslant &\left(\sum_{\ell=1,m=-M}^{L,M}\mathbb{E}|\mathcal{C}^i_m(\varepsilon^\ell)\mathcal{C}^j_m(\varepsilon^\ell)\rho^\ell_m|^2\Delta x^2\Delta t^2\right)^{1/2}\\ &+\left(\sum_{\ell_1\neq\ell_2=1,m=-M}^{L,M}\mathbb{E}\left|(\mathcal{C}^i_m(\varepsilon^{\ell_1})\mathcal{C}^j_m(\varepsilon^{\ell_1})\rho^{\ell_1}_m)(\mathcal{C}^i_m(\varepsilon^{\ell_2})\mathcal{C}^j_m(\varepsilon^{\ell_2})\rho^{\ell_2}_m)\Delta x^2\Delta t^2\right|\right)^{\frac{1}{2}}\\ &+\left(\sum_{\ell=1,m_1\neq m_2=-M}^{L,M}\mathbb{E}\left|(\mathcal{C}^i_{m_1}(\varepsilon^{\ell})\mathcal{C}^j_{m_1}(\varepsilon^{\ell})\rho^{\ell}_{m_1})(\mathcal{C}^i_{m_2}(\varepsilon^{\ell})\mathcal{C}^j_{m_2}(\varepsilon^{\ell})\rho^{\ell}_{m_2})\Delta x^2\Delta t^2\right|\right)^{\frac{1}{2}}\\ &+\left(\sum_{\ell_1\neq\ell_2=1,m_1\neq m_2=-M}^{L,M}\mathbb{E}\left|(\mathcal{C}^i_{m_1}(\varepsilon^{\ell_1})\mathcal{C}^j_{m_1}(\varepsilon^{\ell_1})\rho^{\ell_1}_{m_1})(\mathcal{C}^i_{m_2}(\varepsilon^{\ell_2})\mathcal{C}^j_{m_2}(\varepsilon^{\ell_2})\rho^{\ell_2}_{m_2})\Delta x^2\Delta t^2\right|\right)^{\frac{1}{2}}\\ &= (i)+(ii)+(iii)+(iv)\ .\end{aligned}$$

The expectation in all terms can be bounded by

$$C\sigma^4\Delta x^4\Delta t^2\ .$$

Then, the worst estimate comes from (iv) which has a larger number of terms giving

$$\begin{aligned}\|\Lambda_3^{i,j}\|_{L^2(\varepsilon)} &\leqslant \left(C\frac{T^2}{\Delta t^2}\frac{4R^2}{\Delta x^2}\sigma^4\Delta x^4\Delta t^2\right)^{1/2}\\ &\leqslant C\sigma^2\Delta x\ .\end{aligned}$$

The estimates for the rest of the terms in $\Lambda^{i,j}$ are obtained in an analogous way and one can check that they are of higher order than the terms presented above. Hence we can conclude that for any $(i,j)\in\{1,\dots,n\}^2$

$$\|A_{n,M,L}(i,j)-\widetilde{A}_{n,M,L}(i,j)\|_{L^2(\varepsilon)}\leqslant C(\sigma\sqrt{\Delta x\Delta t}+\sigma^2\Delta x)\ .$$

□

Proof of (2.4.19)

Proof. The structure of the proof is analogous to the one of the proof of Proposition 2.4.2. We need to estimate $\mathcal{D}:=b_{n,M,L}(i)-\widetilde{b}_{n,M,L}(i)$. Note that $\mathcal{D}$ does not depend on the indices $i\in\{1,\dots,n\}$ as the final estimate is again independent of these. Let us put $\mathcal{R}^i_m(g^\ell)=\sum_{k=-M}^{M}(\Psi_i)_{m-k}g(t_\ell,x_k)\Delta x$ for any function g defined on the mesh. Then $\mathcal{D}$ reads

$$\begin{aligned}\mathcal{D}=\ &-\frac{1}{T}\sum_{\ell=1,m=-M}^{L,M}\Bigg[\Big((\widehat{\partial}_t\rho R^i_{n,M,L})^\ell_m+(C^i_{n,M,L}F_{M,L})^\ell_m\Big)\\ &-\Big(\delta_t^+(\rho^\ell_m+\varepsilon^\ell_m)\mathcal{R}^i_m(\rho^\ell+\varepsilon^\ell)+(\rho^\ell_m+\varepsilon^\ell_m)\delta_x^+(\rho^\ell_m+\varepsilon^\ell_m)\mathcal{C}^i_m(\rho^\ell+\varepsilon^\ell)\Big)\Delta x\Delta t\ .\end{aligned}$$

By linearity of $\mathcal{C}$ and $\mathcal{R}$, expanding the second term yields

$$|\mathcal{D}|\leqslant|\zeta^i_t|+|\zeta^i_x|\ ,$$

where ζ^i_t is the sum of all the terms in the expansion of the second term that depend on the noise and its discrete time derivative and ζ^i_x is the sum of all the terms that depend on the noise and its discrete space derivative, i.e.

$$\begin{aligned}\zeta^i_t=&\frac{1}{T}\sum_{v=1}^{3}\zeta^i_{t,v}\\ =&\frac{1}{T}\sum_{\ell=1,m=-M}^{L,M}\Big[\delta_t^+\varepsilon^\ell_m\mathcal{R}^i_m(\rho^\ell)+\delta_t^+\varepsilon^\ell_m\mathcal{R}^i_m(\varepsilon^\ell)+\delta_t^+\rho^\ell_m\mathcal{R}^i_m(\varepsilon^\ell)\Big]\Delta x\Delta t\ ,\end{aligned}$$

and

$$\begin{aligned}\zeta_x^i =& \frac{1}{T}\sum_{v=1}^{7}\zeta_{x,v}^i\\ =&\frac{1}{T}\sum_{\ell=1,m=-M}^{L,M}\Big[\rho_m^\ell\delta_x^+\rho_m^\ell\mathcal{C}_m^i(\varepsilon^\ell)+\rho_m^\ell\delta_x^+\varepsilon_m^\ell\mathcal{C}_m^i(\rho^\ell)+\rho_m^\ell\delta_x^+\varepsilon_m^\ell\mathcal{C}_m^i(\varepsilon^\ell)\\ &+\varepsilon_m^\ell\delta_x^+\rho_m^\ell\mathcal{C}_m^i(\rho^\ell)+\varepsilon_m^\ell\delta_x^+\rho_m^\ell\mathcal{C}_m^i(\varepsilon^\ell)+\varepsilon_m^\ell\delta_x^+\varepsilon_m^\ell\mathcal{C}_m^i(\rho^\ell)+\varepsilon_m^\ell\delta_x^+\varepsilon_m^\ell\mathcal{C}_m^i(\varepsilon^\ell)\Big]\Delta x\Delta t\ .\end{aligned}$$

As before, for brevity, we only display the calculations for the terms yielding the error reported in Proposition 2.4.19 as the rest of the terms are of higher order. Namely, we will consider the terms

$$\zeta_{t,2}^i=\sum_{\ell=1,m=-M}^{L,M}\delta_t^+\varepsilon_m^\ell\mathcal{R}_m^i(\varepsilon^\ell)\Delta x\Delta t\ ,$$

and

$$\zeta_{x,6}^i=\sum_{\ell=1,m=-M}^{L,M}\varepsilon_m^\ell\delta_x^+\varepsilon_m^\ell\mathcal{C}_m^i(\rho^\ell)\Delta x\Delta t\ .$$

Let us begin estimating $\|\zeta_{t,2}^i\|_{L^2(\varepsilon)}$

$$\|\zeta_{t,2}^i\|_{L^2(\varepsilon)}=\left(\mathbb{E}\left|\sum_{\ell=1,m=-M}^{L,M}\delta_t^+\varepsilon_m^\ell\mathcal{R}_m^i(\varepsilon^\ell)\Delta x\Delta t\right|^2\right)^{1/2}\ .$$

Expanding the square and an application of the triangle inequality yields

$$\begin{aligned}\|\zeta_{t,2}^i\|_{L^2(\varepsilon)}\leqslant&\left(\sum_{\ell=1,m=-M}^{L,M}\mathbb{E}\left|\delta_t^+\varepsilon_m^\ell\mathcal{R}_m^i(\varepsilon^\ell)\Delta x\Delta t\right|^2\right)^{1/2}\\ &+\left(\sum_{\ell_1\neq\ell_2=1,m=-M}^{L,M}\mathbb{E}\delta_t^+\varepsilon_m^{\ell_1}\mathcal{R}^i(\varepsilon^{\ell_1})\delta_t^+\varepsilon_m^{\ell_2}\mathcal{R}^i(\varepsilon^{\ell_2})\Delta x^2\Delta t^2\right)^{1/2}\\ &+\left(\sum_{\ell=1,m_1\neq m_2=-M}^{L,M}\mathbb{E}\delta_t^+\varepsilon_{m_1}^\ell\mathcal{R}^i(\varepsilon^\ell)\delta_t^+\varepsilon_{m_2}^\ell\mathcal{R}^i(\varepsilon^\ell)\Delta x^2\Delta t^2\right)^{1/2}\\ &+\left(\sum_{\ell_1\neq\ell_2=1,m_1\neq m_2=-M}^{L,M}\mathbb{E}\delta_t^+\varepsilon_{m_1}^{\ell_1}\mathcal{R}^i(\varepsilon^{\ell_1})\delta_t^+\varepsilon_{m_2}^{\ell_2}\mathcal{R}^i(\varepsilon^{\ell_2})\Delta x^2\Delta t^2\right)^{1/2}\\ =&(i)+(ii)+(iii)+(iv)\ .\end{aligned}$$

We begin by bounding (i)

$$(i)\leqslant\left(\frac{T}{\Delta t}\frac{2R}{\Delta x}\frac{C}{\Delta t^2}\sigma^4\Delta x^2\Delta t^2\right)^{1/2}=C\sigma^2\Delta x^{1/2}\Delta t^{-1/2}\ .$$

For terms $(ii), (iii)$ and (iv) the expectation can be bounded by $C\frac{\sigma^4}{\Delta t^2}\Delta x^4\Delta t^2$. Since the only difference among these terms is how many of them we need to consider, it follows that (iv) produces the worst case yielding

$$\begin{aligned}(iv) &\leqslant \left(\frac{4R^2}{\Delta x^2}\frac{T^2}{\Delta t^2}C\frac{\sigma^4}{\Delta t^2}\Delta x^4\Delta t^2\right)^{1/2}\\ &= C\sigma^2\Delta x\Delta t^{-1}\ .\end{aligned}$$

This is the worst error associated to the noise with respect to time. The estimate for $\|\zeta^i_{x,6}\|_{L^2(\varepsilon)}$ can be obtained in an analogous way. Indeed, we have by expanding the square and an application of the triangle inequality that

$$\begin{aligned}\|\zeta_{x,6}\|_{L^2(\varepsilon)} \leqslant & \sum_{\ell=1,m=-M}^{L,M}\mathbb{E}|\varepsilon^\ell_m\delta^+_x\varepsilon^\ell_m\mathcal{C}^i_m(\rho^\ell)|^2\Delta x^2\Delta t^2\\ &+\sum_{\ell_1\neq\ell_2=1,m=-M}^{L,M}\mathbb{E}|\varepsilon^{\ell^1}_m\varepsilon^{\ell_2}_m\delta^+_x\varepsilon^{\ell_1}_m\delta^+_x\varepsilon^{\ell_2}_m\mathcal{C}^i_m(\rho^{\ell_2})\mathcal{C}^i_m(\rho^{\ell_1})|^2\Delta x^2\Delta t^2\\ &+\sum_{\ell=1,m_1\neq m_2=-M}^{L,M}\mathbb{E}|\varepsilon^\ell_{m_1}\varepsilon^\ell_{m_2}\delta^+_x\varepsilon^\ell_{m_1}\delta^+_x\varepsilon^\ell_{m_2}\mathcal{C}^i_{m_1}(\rho^\ell)\mathcal{C}^i_{m_2}(\rho^\ell)|^2\Delta x^2\Delta t^2\\ &+\sum_{\ell_1\neq\ell_2=1,m_1\neq m_2=-M}^{L,M}\mathbb{E}|\varepsilon^{\ell_1}_{m_1}\varepsilon^{\ell_2}_{m_2}\delta^+_x\varepsilon^{\ell_1}_{m_1}\delta^+_x\varepsilon^{\ell_2}_{m_2}\mathcal{C}^i_{m_1}(\rho^{\ell_1})\mathcal{C}^i_{m_2}(\rho^{\ell_2})|^2\Delta x^2\Delta t^2\\ &= (i)+(ii)+(iii)+(iv)\end{aligned}$$

Term (i) can be bounded as

$$\|(i)\|_{L^2(\varepsilon)} \leqslant C\sigma^2\Delta t^{1/2}\Delta x^{-1/2}\ .$$

Furthermore, the expectations in terms $(ii), (iii)$ and (iv) can be bounded by $C\sigma^4\Delta t^2$. As before, since the only difference among these terms is how many of them we need to consider, it follows that (iv) produces the worst error yielding

$$(iv) \leqslant \left(\frac{4R^2}{\Delta x^2}\frac{T^2}{\Delta t^2}C\sigma^4\Delta t^2\right)^{1/2} = C\sigma^2\Delta x^{-1}$$

Thus, we can conclude that

$$\|b_{n,M,L}(i) - \widetilde{b}_{n,M,L}(i)\|_{L^2(\varepsilon)} \leqslant C\sigma^2(\Delta x^{-1} + \Delta x\Delta t^{-1})$$

□

Bibliography

[ABCvB14] Giacomo Albi, Daniel Balagué, José A. Carrillo, and James von Brecht. Stability analysis of flock and mill rings for second order models in swarming. *SIAM Journal on Applied Mathematics*, 74(3):794–818, 2014.

[ABH+13] Luigi Ambrosio, Alberto Bressan, Dirk Helbing, Axel Klar, Enrique Zuazua, Luigi Ambrosio, and Nicola Gigli. A user's guide to optimal transport. *Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel Rascle*, pages 1–155, 2013.

[ABH+17] Giacomo Albi, Martin Burger, Jan Haskovec, Peter Markowich, and Matthias Schlottbom. Continuum modeling of biological network formation. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 1–48. Birkhäuser/Springer, Cham, 2017.

[AD23] Nathalie Ayi and Nastassia Pouradier Duteil. Graph limit for interacting particle systems on weighted random graphs. *Preprint arXiv:2307.12801*, 2023.

[AD24] Nathalie Ayi and Nastassia Pouradier Duteil. Large-population limits of non-exchangeable particle systems. *Preprint arXiv:2401.07748*, 2024.

[ADP24] Nathalie Ayi, Nastassia Pouradier Duteil, and David Poyato. Mean-field limit of non-exchangeable multi-agent systems over hypergraphs with unbounded rank. *Preprint arXiv:2406.04691*, 2024.

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[AGS11] Pierre Alquier, Eric Gautier, and Gilles Stoltz. *Inverse problems and high-dimensional estimation: stats in the Château Summer School, August 31-September 4, 2009*, volume 203. Springer Science & Business Media, 2011.

[AJM+24] Md Sayeed Anwar, S Nirmala Jenifer, Paulsamy Muruganandam, Dibakar Ghosh, and Timoteo Carletti. Synchronization in adaptive higher-order networks. *Preprint arXiv:2408.12235*, 2024.

[APD21] Nathalie Ayi and Nastassia Pouradier Duteil. Mean-field and graph limits for collective dynamics models with time-varying weights. *J. Differential Equations*, 299:65–110, 2021.

[AS14] Charu Aggarwal and Karthik Subbian. Evolutionary network analysis: A survey. *ACM Computing Surveys (CSUR)*, 47(1):1–36, 2014.

[ASM22] Arash Amini, Qiyu Sun, and Nader Motee. Learning nonlinear couplings in swarm of agents from a single sample trajectory. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 168–173. IEEE, 2022.

[BCC+08] Michele Ballerini, Nicola Cabibbo, Raphael Candelier, Andrea Cavagna, Evaristo Cisbani, Irene Giardina, Vivien Lecomte, Alberto Orlandi, Giorgio Parisi, Andrea Procaccini, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the national academy of sciences*, 105(4):1232–1237, 2008.

[BCE+22] Krishna Balasubramanian, Sinho Chewi, Murat A Erdogdu, Adil Salim, and Shunshi Zhang. Towards a theory of non-log-concave sampling:first-order stationarity guarantees for langevin monte carlo. In Po-Ling Loh and Maxim Raginsky, editors, *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 2896–2923. PMLR, 02–05 Jul 2022.

[BCGC24] Rafael Bailo, José A. Carrillo, and David Gómez-Castro. Aggregation-diffusion equations for collective behaviour in the sciences. *Preprint arXiv:2405.16679*, 2024.

[BCH20] Rafael Bailo, José A. Carrillo, and Jingwei Hu. Fully discrete positivity-preserving and energy-dissipating schemes for aggregation-diffusion equations with a gradient-flow structure. *Communications in Mathematical Sciences*, 18(5):1259–1303, 2020.

[BCL09] Adrien Blanchet, José A. Carrillo, and Philippe Laurençot. Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions. *Calc. Var. Partial Differential Equations*, 35(2):133–168, 2009.

[BCLR13] Daniel Balagué, José A. Carrillo, Thomas Laurent, and Gaël Raoul. Dimensionality of local minimizers of the interaction energy. *Arch. Ration. Mech. Anal.*, 209(3):1055–1088, 2013.

[BCM07] Martin Burger, Vincenzo Capasso, and Daniela Morale. On an aggregation model with long and short range interactions. *Nonlinear Analysis: Real World Applications*, 8(3):939–958, 2007.

[BCMS20] Rafael Bailo, José A. Carrillo, Hideki Murakawa, and Markus Schmidtchen. Convergence of a fully discrete and energy-dissipating finite-volume scheme for aggregation-diffusion equations. *Math. Models Methods Appl. Sci.*, 30(13):2487–2522, 2020.

[BCnC11] François Bolley, José A. Cañizo, and José A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–2210, 2011.

[BCnCD16] Alethea Barbaro, José A. Cañizo, José A. Carrillo, and Pierre Degond. Phase transitions in a kinetic flocking model of Cucker–Smale type. *Multiscale Modeling & Simulation*, 14(3):1063–1088, 2016.

[BCP97] Dario Benedetto, Emanuele Caglioti, and Mario Pulvirenti. A kinetic equation for granular media. *ESAIM: Mathematical Modelling and Numerical Analysis*, 31(5):615–641, 1997.

[BdAS+20] Alexandre Benatti, Henrique F. de Arruda, Filipi Nascimento Silva, César Henrique Comin, and Luciano da Fontoura Costa. Opinion diversity and social bubbles in adaptive sznajd networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(2):023407, feb 2020.

[BDP06] Adrien Blanchet, Jean Dolbeault, and Benoît Perthame. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electron. J. Differential Equations*, pages No. 44, 32, 2006.

[BDS22] Benoît Bonnet, Nastassia Pouradier Duteil, and Mario Sigalotti. Consensus formation in first-order graphon models with time-varying topologies. *Math. Models Methods Appl. Sci.*, 32(11):2121–2188, 2022.

[BDT17] Nicola Bellomo, Pierre Degond, and Eitan Tadmor. *Active Particles, Volume 1: Advances in Theory, Models, and Applications*. Birkhäuser, 2017.

[BEB+09] Alethea Barbaro, Baldvin Einarsson, Björn Birnir, Sven Sigurðsson, Héðinn Valdimarsson, Ólafur Karvel Pálsson, Sveinn Sveinbjörnsson, and Þorsteinn Sigurðsson. Modelling and simulations of the migration of pelagic fish. *ICES Journal of Marine Science*, 66(5):826–838, 2009.

[BEW19] Jérémie Bigot, Paul Escande, and Pierre Weiss. Estimation of linear operators from scattered impulse responses. *Applied and Computational Harmonic Analysis*, 47(3):730–758, 2019.

[BFH14] Martin Burger, Razvan Fetecau, and Yanghong Huang. Stationary states and asymptotic behavior of aggregation models with nonlinear local repulsion. *SIAM Journal on Applied Dynamical Systems*, 13(1):397–424, 2014.

[BFHM17] Mattia Bongini, Massimo Fornasier, Markus Hansen, and Mauro Maggioni. Inferring interaction rules from observations of evolutive systems i: The variational approach. *Mathematical Models and Methods in Applied Sciences*, 27(05):909–951, 2017.

[BGK+23] Rico Berner, Thilo Gross, Christian Kuehn, Jürgen Kurths, and Serhiy Yanchuk. Adaptive dynamical networks. *Preprint arXiv:2304.05652*, 2023.

[BHOT05] Vincent D. Blondel, Julien M Hendrickx, Alex Olshevsky, and John N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 2996–3000. IEEE, 2005.

[BHP23] Martin Burger, Ina Humpert, and Jan-Frederik Pietschmann. Dynamic optimal transport on networks. *ESAIM Control Optim. Calc. Var.*, 29:Paper No. 54, 23, 2023.

[BL73] Henri Berliocchi and Jean-Michel Lasry. Intégrandes normales et mesures paramétrées en calcul des variations. *Bulletin de la Société Mathématique de France*, 101:129–184, 1973.

[BL14] Jacek Banasiak and Mirosław Lachowicz. *Methods of small parameter in mathematical biology*. Springer, 2014.

[BLSSS20] Fabian Baumann, Philipp Lorenz-Spreen, Igor M. Sokolov, and Michele Starnini. Modeling echo chambers and polarization dynamics in social networks. *Phys. Rev. Lett.*, 124(4):048301, 6, 2020.

[BP98] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.

[BPCJ24] Immanuel Ben-Porat, José A. Carrillo, and Pierre-Emmanuel Jabin. The graph limit for a pairwise competition model. *J. Differential Equations*, 413:329–369, 2024.

[Bre11] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*, volume 2. Springer, 2011.

[BS22] Ágnes Backhausz and Balázs Szegedy. Action convergence of operators and graphs. *Canad. J. Math.*, 74(1):72–121, 2022.

[BST22] Laurent Boudin, Francesco Salvarani, and Emmanuel Trélat. Exponential convergence towards consensus for non-symmetric linear first-order systems in finite and infinite dimensions. *SIAM Journal on Mathematical Analysis*, 54(3):2727–2752, 2022.

[BSW22] Davide Bianchi, Alberto G. Setti, and Radosław K. Wojciechowski. The generalized porous medium equation on graphs: existence and uniqueness of solutions with ℓ^1 data. *Calculus of Variations and Partial Differential Equations*, 61(5):171, 2022.

[BT11] Andrew J. Bernoff and Chad M. Topaz. A primer of swarm equilibria. *SIAM Journal on Applied Dynamical Systems*, 10(1):212–250, 2011.

[BTL20] Ghazaleh Beigi, Jiliang Tang, and Huan Liu. Social science–guided feature engineering: A novel approach to signed link analysis. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(1):1–27, 2020.

[Bur22] Martin Burger. Kinetic equations for processes on co-evolving networks. *Kinet. Relat. Models*, 15(2):187–212, 2022.

[BV06] Marek Bodnar and Juan José Lopéz Velázquez. An integro-differential equation arising as a limit of individual cell-based models. *Journal of Differential Equations*, 222(2):341–380, 2006.

[Can08] Emmanuel J. Candès. The restricted isometry property and its implications for compressed sensing. *C. R. Math. Acad. Sci. Paris*, 346(9-10):589–592, 2008.

[Car16] René Carmona. *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*. SIAM, 2016.

[CC06] Vincent Calvez and José A. Carrillo. Volume effects in the Keller-Segel model: energy estimates preventing blow-up. *J. Math. Pures Appl. (9)*, 86(2):155–175, 2006.

[CCH+14a] José A. Carrillo, Young-Pil Choi, Seung-Yeal Ha, Moon-Jin Kang, and Yongduck Kim. Contractivity of transport distances for the kinetic Kuramoto equation. *J. Stat. Phys.*, 156(2):395–415, 2014.

[CCH14b] José Antonio Carrillo, Young-Pil Choi, and Maxime Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lect.*, pages 1–46. Springer, Vienna, 2014.

[CCH15] José A. Carrillo, Alina Chertock, and Yanghong Huang. A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Communications in Computational Physics*, 17(1):233–258, 2015.

[CCY19] José A. Carrillo, Katy Craig, and Yao Yao. Aggregation-diffusion equations: dynamics, asymptotics, and singular limits. In *Active Particles, Volume 2*, pages 65–108. Springer, 2019.

[CDF+11] José A. Carrillo, Marco DiFrancesco, Alessio Figalli, Thomas Laurent, and Dejan Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[CDF+12] José A. Carrillo, Marco DiFrancesco, Alessio Figalli, Thomas Laurent, and Dejan Slepčev. Confinement in nonlocal interaction equations. *Nonlinear Anal.*, 75(2):550–558, 2012.

[CDN14] Guangliang Chen, Atul Divekar, and Deanna Needell. Guaranteed sparse signal recovery with highly coherent sensing matrices. *Sampling Theory in Signal and Image Processing*, 13:91–109, 2014.

[CEM24] José A. Carrillo, Antonio Esposito, and László Mikolás. Long-time behaviour for co-evolving graphs and the graph continuity equation. *In preparation*, 2024.

[CENR11] Emmanuel J. Candes, Yonina C. Eldar, Deanna Needell, and Paige Randall. Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, 31(1):59–73, 2011.

[CERMT24] José A. Carrillo, Gissell Estrada-Rodriguez, László Mikolás, and Sui Tang. Sparse identification of nonlocal interaction kernels in nonlinear gradient flow equations via partial inversion. *Preprint arXiv:2402.06355*, 2024.

[CEW24] José A. Carrillo, Antonio Esposito, and Jeremy Sheung-Him Wu. Nonlocal approximation of nonlinear diffusion equations. *Calculus of Variations and Partial Differential Equations*, 63(4):100, 2024.

[CFPT13] Marco Caponigro, Massimo Fornasier, Benedetto Piccoli, and Emmanuel Trélat. Sparse stabilization and optimal control of the Cucker-Smale model. *Math. Control Relat. Fields*, 3(4):447–466, 2013.

[CFPT15] Marco Caponigro, Massimo Fornasier, Benedetto Piccoli, and Emmanuel Trélat. Sparse stabilization and control of alignment models. *Math. Models Methods Appl. Sci.*, 25(3):521–564, 2015.

[CFTV10] José A. Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil. Particle, kinetic, and hydrodynamic models of swarming. *Mathematical modeling of collective behavior in socio-economic and life sciences*, pages 297–336, 2010.

[CGCYZ23] José A. Carrillo, David Gómez-Castro, Yao Yao, and Chongchun Zeng. Asymptotic simplification of aggregation-diffusion equations towards the heat kernel. *Arch. Ration. Mech. Anal.*, 247(1):Paper No. 11, 45, 2023.

[CGPS20] José A. Carrillo, Rishabh S. Gvalani, Grigorios A. Pavliotis, and André Schlichting. Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus. *Arch. Ration. Mech. Anal.*, 235(1):635–690, 2020.

[CGT18] Yongxin Chen, Tryphon T Georgiou, and Allen Tannenbaum. Vector-valued optimal mass transport. *SIAM Journal on Applied Mathematics*, 78(3):1682–1696, 2018.

[CGT20] Clément Cancès, Thomas O. Gallouët, and Gabriele Todeschi. A variational finite volume scheme for wasserstein gradient flows. *Numer. Math.*, 146(3):437–480, October 2020.

[CGTS21] Katy Craig, Nicolás García Trillos, and Dejan Slepčev. Clustering dynamics on graphs: From spectral clustering to mean shift through fokker–planck interpolation. In *Active Particles, Volume 3: Advances in Theory, Models, and Applications*, pages 105–151. Springer, 2021.

[CHL17] Young-Pil Choi, Seung-Yeal Ha, and Zhuchun Li. Emergent dynamics of the cucker–smale flocking model and its variants. *Active Particles, Volume 1: Advances in Theory, Models, and Applications*, pages 299–331, 2017.

[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.

[CHM14] José A. Carrillo, Yanghong Huang, and Stephan Martin. Explicit flock solutions for Quasi-Morse potentials. *European Journal of Applied Mathematics*, 25(5):553–578, 2014.

[CHMV18] José A. Carrillo, Franca Hoffmann, Edoardo Mainini, and Bruno Volzone. Ground states in the diffusion-dominated regime. *Calc. Var. Partial Differential Equations*, 57(5):Paper No. 127, 28, 2018.

[CHS18] José A. Carrillo, Yanghong Huang, and Markus Schmidtchen. Zoology of a nonlocal cross-diffusion model for two species. *SIAM J. Appl. Math.*, 78(2):1078–1104, 2018.

[CHVY19] José A. Carrillo, Sabine Hittmeir, Bruno Volzone, and Yao Yao. Nonlinear aggregation-diffusion equations: radial symmetry and long time asymptotics. *Invent. Math.*, 218(3):889–977, 2019.

[CJLV16] José A. Carrillo, François James, Frédéric Lagoutière, and Nicolas Vauchelet. The filippov characteristic flow for the aggregation equation with mildly singular potentials. *Journal of Differential Equations*, 260(1):304–338, 2016.

[CL06] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

[CLM14] José Antonio Carrillo, Stefano Lisini, and Edoardo Mainini. Uniqueness for Keller-Segel-type chemotaxis models. *Discrete Contin. Dyn. Syst.*, 34(4):1319–1338, 2014.

[CLP24] Weiqi Chu, Qin Li, and Mason A. Porter. Inference of interaction kernels in mean-field models of opinion dynamics. *SIAM J. Appl. Math.*, 84(3):1096–1115, 2024.

[CM19a] Hayato Chiba and Georgi S. Medvedev. The mean field analysis of the Kuramoto model on graphs I. The mean field equation and transition point formulas. *Discrete Contin. Dyn. Syst.*, 39(1):131–155, 2019.

[CM19b] Hayato Chiba and Georgi S. Medvedev. The mean field analysis of the Kuramoto model on graphs II. Asymptotic stability of the incoherent state, center manifold reduction, and bifurcations. *Discrete Contin. Dyn. Syst.*, 39(7):3897–3921, 2019.

[CM19c] Uthsav Chitra and Christopher Musco. Understanding filter bubbles and polarization in social networks. *Preprint arXiv:1906.08772*, 2019.

[CMS+19] José A. Carrillo, Hideki Murakawa, Makoto Sato, Hideru Togashi, and Olena Trush. A population dynamics model of cell-cell adhesion incorporating population pressure and density saturation. *Journal of theoretical biology*, 474:14–24, 2019.

[CMV03] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matemática Iberoamericana*, 19(3):971–1018, 2003.

[CPn16] Miguel Á. Carreira-Perpiñán. Clustering methods based on kernel density estimators: mean-shift algorithms. In *Handbook of cluster analysis*, Chapman & Hall/CRC Handb. Mod. Stat. Methods, pages 383–417. CRC Press, Boca Raton, FL, 2016.

[CPvG21] Mihai Cucuringu, Andrea Pizzoferrato, and Yves van Gennip. An MBO scheme for clustering and semi-supervised clustering of signed networks. *Commun. Math. Sci.*, 19(1):73–109, 2021.

[CRBD18] Ricky Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[CRT06] Emmanuel J. Candes, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

[CS07a] Felipe Cucker and Steve Smale. Emergent behavior in flocks. *IEEE Trans. Automat. Control*, 52(5):852–862, 2007.

[CS07b] Felipe Cucker and Steve Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2(1):197–227, 2007.

[CT00] José A. Carrillo and Giuseppe Toscani. Asymptotic L^1-decay of solutions of the porous medium equation to self-similarity. *Indiana University Mathematics Journal*, pages 113–142, 2000.

[CT04] José A. Carrillo and Giuseppe Toscani. Wasserstein metric and large–time asymptotics of nonlinear diffusion equations. In *New Trends in Mathematical Physics: In Honour of the Salvatore Rionero 70th Birthday*, pages 234–244. World Scientific, 2004.

[CT06] Emmanuel J. Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

[Dan67] John M. Danskin. *The theory of max-min and its application to weapons allocation problems*, volume V of *Econometrics and Operations Research*. Springer-Verlag New York, Inc., New York, 1967.

[DGRC+22] Francesco Di Giovanni, James Rowbottom, Benjamin Paul Chamberlain, Thomas Markovich, and Michael M. Bronstein. Graph neural networks as gradient flows: understanding graph convolutions via energy. 2022.

[DL15] Karoline Disser and Matthias Liero. On gradient structures for Markov chains and the passage to Wasserstein gradient flows. *Netw. Heterog. Media*, 10(2):233–253, 2015.

[DL21a] Zhiyan Ding and Qin Li. Ensemble kalman inversion: mean-field limit and convergence analysis. *Statistics and Computing*, 31:1–21, 2021.

[DL21b] Zhiyan Ding and Qin Li. Ensemble kalman sampler: Mean-field limit and convergence analysis. *SIAM Journal on Mathematical Analysis*, 53(2):1546–1578, 2021.

[DM09] Wei Dai and Olgica Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE transactions on Information Theory*, 55(5):2230–2249, 2009.

[DM11] Federico Dalmao and Ernesto Mordecki. Cucker-Smale flocking under hierarchical leadership and random interactions. *SIAM J. Appl. Math.*, 71(4):1307–1316, 2011.

[DM17] Alain Durmus and Éric Moulines. Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *The Annals of Applied Probability*, 27(3):1551 – 1587, 2017.

[DMH22] Laetitia Della Maestra and Marc Hoffmann. Nonparametric estimation for interacting particle systems: McKean-Vlasov models. *Probab. Theory Related Fields*, 182(1-2):551–613, 2022.

[DMT08] Bertram Düring, Daniel Matthes, and Giuseppe Toscani. Kinetic equations modelling wealth redistribution: a comparison of approaches. *Phys. Rev. E (3)*, 78(5):056103, 12, 2008.

[Dob79] Roland L'vovich Dobrushin. Vlasov equations. *Functional Analysis and Its Applications*, 13(2):115–123, 1979.

[DRD20] Arnak S. Dalalyan and Lionel Riou-Durand. On sampling from a log-concave density using kinetic Langevin diffusions. *Bernoulli*, 26(3):1956 – 1988, 2020.

[DRR22] Sara Daneri, Emanuela Radici, and Eris Runa. Deterministic particle approximation of aggregation-diffusion equations on unbounded domains. *Journal of Differential Equations*, 312:474–517, 2022.

[Ebe11] Andreas Eberle. Reflection coupling and wasserstein contractivity without convexity. *Comptes Rendus Mathematique*, 349(19):1101–1104, 2011.

[EBFE23] Imad El Bouchairi, Jalal M. Fadili, and Abderrahim Elmoataz. Continuum limit of p-laplacian evolution problems on graphs: L^q graphons and sparse graphs. *ESAIM: Mathematical Modelling and Numerical Analysis*, 57(3):1795–1838, 2023.

[EdVL07] Raluca Eftimie, Gerda de Vries, and Mark A Lewis. Complex spatial group patterns result from different animal communication mechanisms. *Proceedings of the National Academy of Sciences*, 104(17):6974–6979, 2007.

[EDVLL07] Raluca Eftimie, Gerda De Vries, Mark A Lewis, and Frithjof Lutscher. Modeling group formation and activity patterns in self-organizing collectives of individuals. *Bulletin of mathematical biology*, 69(5):1537–1565, 2007.

[EFMM22] Matthias Erbar, Dominik Forkert, Jan Maas, and Delio Mugnolo. Gradient flow formulation of diffusion equations in the Wasserstein space over a metric graph. *Netw. Heterog. Media*, 17(5):687–717, 2022.

[EGH00] Robert Eymard, Thierry Gallouët, and Raphaèle Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, volume VII of *Handb. Numer. Anal.*, pages 713–1020. North-Holland, Amsterdam, 2000.

[EGSS24] Antonio Esposito, Rishabh. S. Gvalani, André Schlichting, and Markus Schmidtchen. On a novel gradient flow structure for the aggregation equation. *Calc. Var. Partial Differential Equations*, 63(5):Paper No. 126, 33, 2024.

[EGZ19] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.*, 47(4):1982–2010, 2019.

[EHPS23] Antonio Esposito, Georg Heinze, Jan-Frederik Pietschmann, and André Schlichting. Graph-to-local limit for a multi-species nonlocal cross-interaction system. *PAMM*, 23(4):e202300094, 2023.

[EHS23] Antonio Esposito, Georg Heinze, and André Schlichting. Graph-to-local limit for the nonlocal interaction equation. *Preprint arXiv:2306.03475*, 2023.

[EM12] Matthias Erbar and Jan Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.*, 206(3):997–1038, 2012.

[EM24] Antonio Esposito and László Mikolás. On evolution PDEs on co-evolving graphs. *Discrete Contin. Dyn. Syst.*, 44(9):2660–2683, 2024.

[EPS23] Antonio Esposito, Francesco Saverio Patacchini, and André Schlichting. On a class of nonlocal continuity equations on graphs. *European J. Appl. Math.*, 35(1):109–126, 2023.

[EPSS21] Antonio Esposito, Francesco S. Patacchini, André Schlichting, and Dejan Slepčev. Nonlocal-interaction equation on graphs: gradient flow structure and continuum limit. *Arch. Ration. Mech. Anal.*, 240(2):699–760, 2021.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[Fal22] Carles Falcó. From random walks on networks to nonlinear diffusion. *Phys. Rev. E*, 106:054103, Nov 2022.

[FF23] Simone Fagioli and Gianluca Favre. Opinion formation on evolving network. the dpa method applied to a nonlocal cross-diffusion pde-ode system. *Preprint arXiv:2306.10844*, 2023.

[FJMAT24] Gianluca Favre, Gaspard Jankowiak, Sara Merino-Aceituno, and Lara Trussardi. Continuum modeling for opinion formation on a network. *Preprint arXiv:2406.18076*, 2024.

[FKRT24] Jinchao Feng, Charles Kulick, Yunxiang Ren, and Sui Tang. Learning particle swarming models from data with gaussian processes. *Mathematics of Computation*, 93(349):2391–2437, 2024.

[FL12] Albert Fannjiang and Wenjing Liao. Coherence pattern–guided compressive sensing with unresolved grids. *SIAM Journal on Imaging Sciences*, 5(1):179–202, 2012.

[FMP22] Dominik Forkert, Jan Maas, and Lorenzo Portinale. Evolutionary Γ-convergence of entropic gradient flow structures for Fokker–Planck equations in multiple dimensions. *SIAM J. Math. Anal.*, 54(4):4297–4333, 2022.

[Fol99] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[FPTT17] Giulia Furioli, Ada Pulvirenti, Elide Terraneo, and Giuseppe Toscani. Fokker-Planck equations in the modeling of socio-economic phenomena. *Math. Models Methods Appl. Sci.*, 27(1):115–158, 2017.

[FS21] Zachary Feinstein and Andreas Søjmark. Short communication: dynamic default contagion in heterogeneous interbank systems. *SIAM J. Financial Math.*, 12(4):SC83–SC97, 2021.

[GC24] David Gómez-Castro. Beginner's guide to aggregation-diffusion equations. *SeMA Journal*, pages 1–57, 2024.

[GCL24] Valentine Genon-Catalot and Catherine Larédo. Parametric inference for ergodic McKean-Vlasov stochastic differential equations. *Bernoulli*, 30(3):1971–1997, 2024.

[Ger20] Viktor Gerasimenko. Kinetic equations of granular media. In *Progress in Fine Particle Plasmas*, pages 163–177. IntechOpen, 2020.

[GJKM22] Marios Antonios Gkogkas, Benjamin Jüttner, Christian Kuehn, and Erik Andreas Martens. Graphop mean-field limits and synchronization for the stochastic Kuramoto model. *Chaos*, 32(11):Paper No. 113120, 13, 2022.

[GK22] Marios Antonios Gkogkas and Christian Kuehn. Graphop mean-field limits for Kuramoto-type models. *SIAM J. Appl. Dyn. Syst.*, 21(1):248–283, 2022.

[GKMP20] Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Homogenisation of one-dimensional discrete optimal transport. *J. Math. Pures Appl. (9)*, 139:204–234, 2020.

[GKMP23] Peter Gladbach, Eva Kopfer, Jan Maas, and Lorenzo Portinale. Homogenisation of dynamical optimal transport on periodic graphs. *Calc Var Partial Differ Equ*, 62(5), April 2023.

[GKX22] Marios Antonios Gkogkas, Christian Kuehn, and Chuang Xu. Mean field limits of co-evolutionary heterogeneous networks. *Preprint arXiv:2202.01742*, 2022.

[GKX23] Marios A. Gkogkas, Christian Kuehn, and Chuang Xu. Continuum limits for adaptive network dynamics. *Commun. Math. Sci.*, 21(1):83–106, 2023.

[GL95] Shay Gueron and Simon A Levin. The dynamics of group formation. *Mathematical biosciences*, 128(1-2):243–264, 1995.

[Gol16] François Golse. On the dynamics of large particle systems in the mean field limit. *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, pages 1–144, 2016.

[GPY17] Josselin Garnier, George Papanicolaou, and Tzu-Wei Yang. Consensus convergence with stochastic effects. *Vietnam J. Math.*, 45(1-2):51–75, 2017.

[GST01] Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM review*, 43(1):89–112, 2001.

[GTS16] Nicolás García Trillos and Dejan Slepčev. Continuum limit of total variation on point clouds. *Arch. Ration. Mech. Anal.*, 220(1):193–241, 2016.

[GTS18] Nicolás García Trillos and Dejan Slepčev. A variational approach to the consistency of spectral clustering. *Appl. Comput. Harmon. Anal.*, 45(2):239–281, 2018.

[GvGO22] Yoshikazu Giga, Yves van Gennip, and Jun Okamoto. Graph gradient flows: from discrete to continuum. *preprint arXiv:2211.03384*, 2022.

[GVR05] Sanjeev Goyal and Fernando Vega-Redondo. Network formation and social coordination. *Games and Economic Behavior*, 50(2):178–207, 2005.

[Has21] Jan Haskovec. A simple proof of asymptotic consensus in the Hegselmann–Krause and Cucker–Smale models with normalization and delay. *SIAM Journal on Applied Dynamical Systems*, 20(1):130–148, 2021.

[Has24] Jan Haskovec. Graph limit of the consensus model with self-delay. *Journal of Physics A: Mathematical and Theoretical*, 57(34):345203, 2024.

[HC13] Dan Hu and David Cai. Adaptation and optimization of biological transport networks. *Phys. Rev. Lett.*, 111:138701, Sep 2013.

[HCH10] Hanno Hildenbrandt, Cladio Carere, and Charlotte K Hemelrijk. Self-organized aerial displays of thousands of starlings: a model. *Behavioral Ecology*, 21(6):1349–1359, 2010.

[Hei24] Georg Heinze. *Graph-based nonlocal gradient systems and their local limits.* Ph.D. thesis, University of Augsburg, 2024.

[HJ23] Ben Hambly and Philipp Jettkant. Control of Mckean–Vlasov sdes with contagion through killing at a state-dependent intensity. *Preprint arXiv:2310.15854*, 2023.

[HKL+20] Yuchen He, Sung Ha Kang, Wenjing Liao, Hao Liu, and Yingjie Liu. Robust PDE identification from noisy data. *Preprint arXiv:2006.06557*, 2020.

[HKL+22] Yuchen He, Sung Ha Kang, Wenjing Liao, Hao Liu, and Yingjie Liu. Numerical identification of nonlocal potential in aggregation. *Preprint arXiv:2207.03358*, 2022.

[HKL+23] Yuchen He, Sung Ha Kang, Wenjing Liao, Hao Liu, and Yingjie Liu. Group projected subspace pursuit for identification of variable coefficient differential equations (GP-IDENT). *Journal of Computational Physics*, 494:112526, 2023.

[HKT22] Leonhard Horstmeyer, Christian Kuehn, and Stefan Thurner. Balancing quarantine and self-distancing measures in adaptive epidemic networks. *Bull. Math. Biol.*, 84(8):Paper No. 79, 22, 2022.

[Hol15] Petter Holme. Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88:1–30, 2015.

[HPS23a] Georg Heinze, Jan-Frederik Pietschmann, and Markus Schmidtchen. Nonlocal cross-interaction systems on graphs: energy landscape and dynamics. *Kinet. Relat. Models*, 16(6):795–827, 2023.

[HPS23b] Georg Heinze, Jan-Frederik Pietschmann, and Markus Schmidtchen. Nonlocal cross-interaction systems on graphs: nonquadratic Finslerian structure and nonlinear mobilities. *SIAM J. Math. Anal.*, 55(6):7039–7076, 2023.

[HS12] Petter Holme and Jari Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.

[HST23] Anastasiia Hraivoronska, André Schlichting, and Oliver Tse. Variational convergence of the Scharfetter–Gummel scheme to the aggregation-diffusion equation and vanishing diffusion limit. *Preprint arXiv:2306.02226*, 2023.

[HT23] Anastasiia Hraivoronska and Oliver Tse. Diffusive limit of random walks on tessellations via generalized gradient flows. *SIAM Journal on Mathematical Analysis*, 55(4):2948–2995, 2023.

[HTF09] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning*. Springer Series in Statistics. Springer, New York, second edition, 2009. Data mining, inference, and prediction.

[IMP+24] Reza Irankhah, Mahtab Mehrabbeik, Fatemeh Parastesh, Karthikeyan Rajagopal, Sajad Jafari, and Jürgen Kurths. Synchronization enhancement subjected to adaptive blinking coupling. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34(2), 2024.

[IRB+11] Lorenzo Isella, Mariateresa Romano, Alain Barrat, Ciro Cattuto, Vittoria Colizza, Wouter Van den Broeck, Francesco Gesualdo, Elisabetta Pandolfi, Lucilla Ravà, Caterina Rizzo, et al. Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. *PloS one*, 6(2):e17144, 2011.

[IY23] Ryosuke Ihara and Kazuyuki Yagasaki. Continuum limits of coupled oscillator networks depending on multiple sparse graphs. *J. Nonlinear Sci.*, 33(4):Paper No. 62, 44, 2023.

[Jab14] Pierre-Emmanuel Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models*, 7(4):661–711, 2014.

[JM14] Pierre-Emmanuel Jabin and Sebastien Motsch. Clustering and asymptotic behavior in opinion formation. *Journal of Differential Equations*, 257(11):4165–4187, 2014.

[JW17] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. *Active Particles, Volume 1: Advances in Theory, Models, and Applications*, pages 379–402, 2017.

[JWH+23] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor. *An introduction to statistical learning—with applications in Python*. Springer Texts in Statistics. Springer, Cham, 2023.

[JZ23] Pierre-Emmanuel Jabin and Datong Zhou. The mean-field limit of sparse networks of integrate and fire neurons. *Preprint arXiv:2309.04046*, 2023.

[Kas90] Raphael A Kasonga. Maximum likelihood theory for large interacting systems. *SIAM Journal on Applied Mathematics*, 50(3):865–875, 1990.

[KCB+13] Theodore Kolokolnikov, José A. Carrillo, Andrea Bertozzi, Razvan Fetecau, and Mark Lewis. Emergent behaviour in multi-particle systems with non-local interactions. *Physica D: Nonlinear Phenomena*, (260):1–4, 2013.

[KGK+20] Julian Kohne, Natalie Gallagher, Zeynep Melis Kirgil, Rocco Paolillo, Lars Padmos, and Fariba Karimi. *The Role of Network Structure and Initial Group Norm Distributions in Norm Conflict*, pages 113–140. Springer International Publishing, Cham, 2020.

[KLL21] Sung Ha Kang, Wenjing Liao, and Yingjie Liu. Ident: Identifying differential equations with numerical time evolution. *Journal of Scientific Computing*, 87:1–27, 2021.

[KS70] Evelyn F. Keller and Lee A. Segel. Initiation of slime mold aggregation viewed as an instability. *J. Theoret. Biol.*, 26(3):399–415, 1970.

[KS71] Evelyn F. Keller and Lee A. Segel. Model for chemotaxis. *Journal of theoretical biology*, 30(2):225–234, 1971.

[KSSF16] Srijan Kumar, Francesca Spezzano, VS Subrahmanian, and Christos Faloutsos. Edge weight prediction in weighted signed networks. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 221–230. IEEE, 2016.

[KTT22] Hajime Koike, Hideki Takayasu, and Misako Takayasu. Diffusion-localization transition point of gravity type transport model on regular ring lattices and bethe lattices. *Journal of Statistical Physics*, 186(3):44, 2022.

[Kue20] Christian Kuehn. Network dynamics on graphops. *New J. Phys.*, 22:053030, 11, 2020.

[Kur75] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International Symposium on Mathematical Problems in Theoretical Physics (Kyoto Univ., Kyoto, 1975)*, volume 39 of *Lecture Notes in Phys.*, pages 420–422. Springer, Berlin-New York, 1975.

[KVM18] Dmitry Kaliuzhnyi-Verbovetskyi and Georgi S. Medvedev. The mean field equation for the Kuramoto model on graph sequences with non-Lipschitz limit. *SIAM J. Math. Anal.*, 50(3):2441–2465, 2018.

[KX22] Christian Kuehn and Chuang Xu. Vlasov equations on digraph measures. *J. Differential Equations*, 339:261–349, 2022.

[Lac16] Daniel Lacker. A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Related Fields*, 165(3-4):581–648, 2016.

[Lan05] Carlo Lancellotti. On the Vlasov limit for systems of nonlinearly coupled oscillators without noise. *Transport Theory Statist. Phys.*, 34(7):523–535, 2005.

[Lau07] Thomas Laurent. Local and global existence for an aggregation equation. *Communications in Partial Differential Equations*, 32(12):1941–1964, 2007.

[LL22] Quanjun Lang and Fei Lu. Learning interaction kernels in mean-field equations of first-order systems of interacting particles. *SIAM J. Sci. Comput.*, 44(1):A260–A285, 2022.

[LL23a] Quanjun Lang and Fei Lu. Identifiability of interaction kernels in mean-field equations of interacting particles. *Found. Data Sci.*, 5(4):480–502, 2023.

[LL23b] Tim Laux and Jona Lelmi. Large data limit of the MBO scheme for data clustering: convergence of the dynamics. *Journal of Machine Learning Research*, 24(344):1–49, 2023.

[LL23c] Zhongyang Li and Fei Lu. On the coercivity condition in the learning of interacting particle systems. *Stoch. Dyn.*, 23(8):Paper No. 2340003, 26, 2023.

[LLEK10] Ryan Lukeman, Yue-Xian Li, and Leah Edelstein-Keshet. Inferring individual rules from collective behavior. *Proceedings of the National Academy of Sciences*, 107(28):12576–12580, 2010.

[LLM+21] Zhongyang Li, Fei Lu, Mauro Maggioni, Sui Tang, and Cheng Zhang. On the identifiability of interaction functions in systems of interacting particles. *Stochastic Process. Appl.*, 132:135–163, 2021.

[LLN19] Yulong Lu, Jianfeng Lu, and James Nolen. Accelerating langevin sampling with birth-death, 2019.

[LMT21] Fei Lu, Mauro Maggioni, and Sui Tang. Learning interaction kernels in heterogeneous systems of agents from multiple trajectories. *J. Mach. Learn. Res.*, 22:Paper No. 32, 67, 2021.

[LMT22] Fei Lu, Mauro Maggioni, and Sui Tang. Learning interaction kernels in stochastic systems of interacting particles from multiple trajectories. *Found. Comput. Math.*, 22(4):1013–1067, 2022.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[LPS+24] Matteo Lodi, Shirin Panahi, Francesco Sorrentino, Alessandro Torcini, and Marco Storace. Patterns of synchronized clusters in adaptive networks. *Communications Physics*, 7(1):198, 2024.

[LTB09] Andrew J. Leverentz, Chad M. Topaz, and Andrew J. Bernoff. Asymptotic dynamics of attractive-repulsive swarms. *SIAM Journal on Applied Dynamical Systems*, 8(3):880–908, 2009.

[LWLM24] Quanjun Lang, Xiong Wang, Fei Lu, and Mauro Maggioni. Interacting particle systems on networks: joint inference of the network and the interaction kernel. *Preprint arXiv:2402.08412*, 2024.

[LZTM19] Fei Lu, Ming Zhong, Sui Tang, and Mauro Maggioni. Nonparametric inference of interaction laws in systems of agents from trajectory data. *Proc. Natl. Acad. Sci. USA*, 116(29):14424–14433, 2019.

[Maa11] Jan Maas. Gradient flows of the entropy for finite markov chains. *Journal of Functional Analysis*, 261(8):2250–2292, 2011.

[MB21] Daniel A Messenger and David M Bortz. Weak SINDy for partial differential equations. *Journal of Computational Physics*, 443:110525, 2021.

[MB22] Daniel A Messenger and David M Bortz. Learning mean-field equations from particle data using WSINDy. *Physica D: Nonlinear Phenomena*, 439:133406, 2022.

[Med14a] Georgi S. Medvedev. The nonlinear heat equation on dense graphs and graph limits. *SIAM Journal on Mathematical Analysis*, 46(4):2743–2766, 2014.

[Med14b] Georgi S. Medvedev. The nonlinear heat equation on W-random graphs. *Arch. Ration. Mech. Anal.*, 212(3):781–803, 2014.

[Med19] Georgi S. Medvedev. Correction to: The nonlinear heat equation on W-random graphs. *Arch. Ration. Mech. Anal.*, 2ß31(2):1305–1308, 2019.

[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.

[MK19] Erik Andreas Martens and Konstantin Klemm. Cyclic structure induced by load fluctuations in adaptive transportation networks. In *Progress in Industrial Mathematics at ECMI 2018*, pages 147–155. Springer, 2019.

[MKB13] Ekaterina Merkurjev, Tijana Kostic, and Andrea L Bertozzi. An MBO scheme on graphs for classification and image processing. *SIAM Journal on Imaging Sciences*, 6(4):1903–1930, 2013.

[Mor05] Luc Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Control*, 50(2):169–182, 2005.

[MPKHT20] Tuan Minh Pham, Imre Kondor, Rudolf Hanel, and Stefan Thurner. The effect of social balance on social fragmentation. *Journal of the Royal Society Interface*, 17(172):20200752, 2020.

[MPPD19] Sean McQuade, Benedetto Piccoli, and Nastassia Pouradier Duteil. Social dynamics models with time-varying influence. *Math. Models Methods Appl. Sci.*, 29(4):681–716, 2019.

[MRT15] José M. Mazón, Julio D. Rossi, and Julián Toledo. Optimal mass transport on metric graphs. *SIAM J. Optim.*, 25(3):1609–1632, 2015.

[MTZM23] Jason Miller, Sui Tang, Ming Zhong, and Mauro Maggioni. Learning theory for inferring interaction kernels in second-order interacting agent systems. *Sampl. Theory Signal Process. Data Anal.*, 21(1):Paper No. 21, 58, 2023.

[NÁBV10] Máté Nagy, Zsuzsa Ákos, Dora Biro, and Tamás Vicsek. Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290):890–893, 2010.

[NDPW+18] Christian Nowke, Sandra Diaz-Pier, Benjamin Weyers, Bernd Hentschel, Abigail Morrison, Torsten W Kuhlen, and Alexander Peyser. Toward rigorous parameterization of underconstrained neural network models through interactive visualization and steering of connectivity generation. *Frontiers in neuroinformatics*, 12:32, 2018.

[Neu84] Helmut Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic theories and the Boltzmann equation (Montecatini, 1981)*, volume 1048 of *Lecture Notes in Math.*, pages 60–110. Springer, Berlin, 1984.

[NJW01] Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 14, 2001.

[NS11] Stamatios C. Nicolis and David Sumpter. A dynamical approach to stock market fluctuations. *International Journal of Bifurcation and Chaos*, 21(12):3557–3564, 2011.

[NT09] Deanna Needell and Joel A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and computational harmonic analysis*, 26(3):301–321, 2009.

[Oel90] Karl Oelschläger. Large systems of interacting particles and the porous medium equation. *J. Differential Equations*, 88(2):294–346, 1990.

[OSM04] Reza Olfati-Saber and Richard M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automat. Control*, 49(9):1520–1533, 2004.

[Ott01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26:101–174, 2001.

[Pat53] Clifford S. Patlak. Random walk with persistence and external bias. *Bull. Math. Biophys.*, 15:311–338, 1953.

[PCG23] Immanuel Ben Porat, José A. Carrillo, and Sondre T. Galtung. Mean field limit for one dimensional opinion dynamics with coulomb interaction and time dependent weights. *Preprint arXiv:2306.01099*, 2023.

[PD22] Nastassia Pouradier Duteil. Mean-field limit of collective dynamics with time-varying weights. *Netw. Heterog. Media*, 17(2):129–161, 2022.

[PS23] Mark A. Peletier and André Schlichting. Cosh gradient systems and tilting. *Nonlinear Anal.*, 231:Paper No. 113094, 113, 2023.

[PT24] Thierry Paul and Emmanuel Trélat. From microscopic to macroscopic scale dynamics: mean field, hydrodynamic and graph limits. *Preprint hal-04577170*, May 2024.

[RABK19] Samuel Rudy, Alessandro Alla, Steven L Brunton, and J Nathan Kutz. Data-driven identification of parametric partial differential equations. *SIAM Journal on Applied Dynamical Systems*, 18(2):643–660, 2019.

[Rao12] Gaël Raoul. Nonlocal interaction equations: stationary states and stability analysis. *Differential Integral Equations*, 25(5-6):417–440, 2012.

[RBPK17] Samuel H. Rudy, Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Data-driven discovery of partial differential equations. *Science advances*, 3(4):e1602614, 2017.

[Rey87] Craig W. Reynolds. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 25–34, 1987.

[RK02] Hegselmann Rainer and Ulrich Krause. Opinion dynamics and bounded confidence: models, analysis and simulation. 2002.

[Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[RVE22] G. M. Rotskoff and E. Vanden-Eijnden. Trainability and accuracy of artificial neural networks: an interacting particle system approach. *Comm. Pure Appl. Math.*, 75(9):1889–1935, 2022.

[SAB19] Guodong Shi, Claudio Altafini, and John S Baras. Dynamics over signed networks. *SIAM Review*, 61(2):229–257, 2019.

[San15] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

[Say14] Ali H. Sayed. Adaptive networks. *Proceedings of the IEEE*, 102(4):460–497, 2014.

[Sch17] Hayden Schaeffer. Learning partial differential equations via data discovery and sparse optimization. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2197):20160446, 2017.

[Sen79] Eugene Seneta. Coefficients of ergodicity: structure and applications. *Advances in applied probability*, 11(3):576–590, 1979.

[SKCS23] Valentina Shumovskaia, Mert Kayaalp, Mert Cemri, and Ali H. Sayed. Discovering influencers in opinion formation over social graphs. *IEEE Open Journal of Signal Processing*, 4:188–207, 2023.

[SM16] Pulipati Srilatha and Ramakrishnan Manjula. Similarity index based link prediction algorithms in social networks: A survey. *Journal of Telecommunications and Information Technology*, (2):87–94, 2016.

[Sor48] Thorvald Sorensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons. *Biologiske skrifter*, 5:1–34, 1948.

[Sor01] Didier Sornette. Fokker–Planck equation of distributions of financial returns and power laws. *Physica A: Statistical Mechanics and its Applications*, 290(1-2):211–217, 2001.

[SPBDM21] Kashin Sugishita, Mason A. Porter, Mariano Beguerisse-Díaz, and Naoki Masuda. Opinion dynamics on tie-decay networks. *Phys. Rev. Res.*, 3:023249, Jun 2021.

[Ste77] Gilbert W. Stewart. On the perturbation of pseudo-inverses, projections and linear least squares problems. *SIAM review*, 19(4):634–662, 1977.

[SW23] Dejan Slepčev and Andrew Warren. Nonlocal wasserstein distance: metric and asymptotic properties. *Calculus of Variations and Partial Differential Equations*, 62(9):238, 2023.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[TB04] Chad M. Topaz and Andrea L. Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM Journal on Applied Mathematics*, 65(1):152–174, 2004.

[TBL06] Chad M. Topaz, Andrea L. Bertozzi, and Mark A. Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of mathematical biology*, 68(7):1601–1623, 2006.

[THK18] Stefan Thurner, Rudolf Hanel, and Peter Klimek. *Introduction to the theory of complex systems*. Oxford University Press, Oxford, 2018.

[TKI+13] Kolbjørn Tunstrøm, Yael Katz, Christos C Ioannou, Cristián Huepe, Matthew J Lutz, and Iain D Couzin. Collective states, multistability and transitional behavior in schooling fish. *PLoS computational biology*, 9(2):e1002915, 2013.

[Tok18] Tomoki Tokuda. Statistical test for detecting community structure in real-valued edge-weighted graphs. *PloS one*, 13(3):e0194079, 2018.

[Tos04] Giuseppe Toscani. Kinetic and hydrodynamic models of nearly elastic granular flows. *Monatsh. Math.*, 142(1-2):179–192, 2004.

[Tos06] Giuseppe Toscani. Kinetic models of opinion formation. *Commun. Math. Sci.*, 4(3):481–496, 2006.

[TTT18] Koutarou Tamura, Hideki Takayasu, and Misako Takayasu. Diffusion-localization transition caused by nonlinear transport on complex networks. *Scientific Reports*, 8(1):5517, 2018.

[TTZ23] Sui Tang, Malik Tuerkoen, and Hanming Zhou. On the identifiablility of nonlocal interaction kernels in first-order systems of interacting particles on riemannian manifolds. *Preprint arXiv:2305.12340*, 2023.

[VÓ7] Juan Luis Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007. Mathematical theory.

[vBUKB12] James H. von Brecht, David Uminsky, Theodore Kolokolnikov, and Andrea L. Bertozzi. Predicting pattern formation in particle interactions. *Math. Models Methods Appl. Sci.*, 22:1140002, 31, 2012.

[WM22] John Wright and Yi Ma. *High-dimensional data analysis with low-dimensional models: Principles, computation, and applications*. Cambridge University Press, 2022.

[WT21] Bo Wang and Yu-Ping Tian. Consensus of discrete-time multi-agent systems with multiplicative uncertainties and delays. *Internat. J. Systems Sci.*, 52(11):2311–2323, 2021.

[YCY22] Rentian Yao, Xiaohui Chen, and Yun Yang. Mean-field nonparametric estimation of interacting particle systems. In *Conference on Learning Theory*, pages 2242–2275. PMLR, 2022.

[YDC14] Yang Yang, Yuxiao Dong, and Nitesh V. Chawla. Predicting node degree centrality with the node prominence profile. *Scientific reports*, 4(1):7236, 2014.