



Delft University of Technology

Document Version

Final published version

Citation (APA)

Junné, J. (2026). *Interacting Particle Systems and Riemannian Geometry*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:c0b0715c-8845-4a01-bcf9-cfd69456c6bc>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Interacting Particle Systems *and* Riemannian Geometry

Jonathan Junné

Interacting Particle Systems and Riemannian Geometry

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
Prof.dr.ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Wednesday, the 28th of October at 17:30 o'clock.

by

Jonathan JUNNÉ

This dissertation has been approved by the (co)promoters.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof.dr. F. Redig,	Delft University of Technology, promotor
Dr. R. Versendaal,	Delft University of Technology, copromotor

Independent Members:

Prof.dr. M.C. Veraar,	Delft University of Technology
Prof.dr. F. Baudoin,	Aarhus University, Denmark
Prof.dr. P. Gonçalves,	Instituto Superior Técnico, University of Lisbon, Portugal
Prof.dr. N. Gozlan,	Université Paris Cité, France
Dr. R.C. Kraaij,	Delft University of Technology
Prof.dr.ir. G. Jongbloed,	Delft University of Technology, reserve member

This research was financially supported by the Dutch Research Council (NWO) through the project *Interacting particle systems and Riemannian geometry*, with project number OCENW.M20.251, as part of the research program Open Competitie ENW.



Key words: interacting particle systems; Riemannian geometry; hydrodynamic limits;
random walks; Wasserstein distance; empirical measures; Fisher information
Copyright: 2026 Jonathan Junné

An electronic version of this dissertation is available at: <http://repository.tudelft.nl>

Contents

1	Introduction	1
1.1	General introduction	1
1.2	Overview of the main results of the thesis	7
1.2.1	Chapter 2 – Invariance principle for lifted geodesic random walks	7
1.2.2	Chapter 3 – Hydrodynamic limit of the symmetric exclusion process on manifolds	9
1.2.3	Chapter 4 – Concentration and expectations for empirical measures on manifolds	12
1.2.4	Chapter 5 – Well-posedness of the multi-species Landau equation	15
2	Invariance Principle for Lifted Geodesic Random Walks	19
2.1	Introduction	19
2.2	Random walks and horizontal random walks	20
2.2.1	Geodesic random walk	20
2.2.2	Horizontal lift of geodesic random walks	22
2.3	Invariance principle for Riemannian submersions	23
2.3.1	Main result	23
2.3.2	Important examples	28
3	Hydrodynamic Limit of the Symmetric Exclusion Process on Manifolds	33
3.1	Introduction and main results	33
3.1.1	Introduction	33
3.1.2	Main results	35
3.2	Existence of the SEP	38
3.3	Abstract hydrodynamic limit via duality	45
3.3.1	Duality and hypotheses on the random neighborhood graph	45
3.3.2	Hydrodynamic limit of the SEP	47
3.4	Convergence of discrete random walks	50
3.5	Lifting the SEP to principal bundles	56
3.5.1	Geometric setting	56
3.5.2	Hydrodynamic limit of the lifted SEP	58
3.6	Appendix	62

4	Concentration and Expectation for Empirical Measures on Manifolds	71
4.1	Introduction	71
4.2	Definitions and main results	72
4.2.1	Wasserstein distance and relative entropy	72
4.2.2	Geometric setting	74
4.2.3	Upper bounds for the expected Wasserstein distance	77
4.2.4	Quantitative Sanov's theorem	79
4.2.5	Concentration for aggregation-diffusion	82
4.3	Proofs for independent random variables	83
4.3.1	Expected Wasserstein distance	83
4.3.2	Concentration inequalities	87
4.4	An interacting particle system	100
4.4.1	Description of the interacting particle system	100
4.4.2	Assumptions on the potentials	102
4.4.3	Preliminary results	103
4.4.4	Coupling the interacting and non interacting proxy	110
4.5	Appendix	115
5	Well-Posedness of the Multi-Species Landau Equation	123
5.1	Introduction	123
5.1.1	Main results	126
5.1.2	Summary of previous results	128
5.1.3	Preliminaries and structure of the chapter	130
5.2	The multi-species Landau system with finite masses	131
5.2.1	Lifting argument for the multi-species system	131
5.2.2	Dissipation of the weighted Fisher information	133
5.2.3	Local well-posedness	141
5.2.4	Global well-posedness	147
5.3	The two-species Landau system with large mass ratio	148
5.3.1	Non-monotonicity of the Fisher information: a counterexample	148
5.3.2	Dissipation identities for the relative entropy	154
5.3.3	A new Lyapunov functional	156
5.4	Appendix	162
5.4.1	Spherical calculus	162
	Bibliography	165
	Summary	173
	Acknowledgments	175
	Curriculum Vitae	177
	Publications	179

CHAPTER 1

Introduction

1.1 General introduction

In statistical physics, complex systems are made up of a large number of individual constituents, called particles. These particles may move, jump, collide, or interact, which makes their precise behavior complex to describe explicitly. Moreover, tracking the evolution of each particle individually is often not that informative. Instead, one is interested in the evolution of macroscopic conserved quantities under microscopic dynamics. This leads to one of the central questions of statistical mechanics:

How do macroscopic equations emerge from microscopic dynamics?

Deriving macroscopic equations from microscopic laws is in general a difficult task, even if the latter are simple.

A particularly successful framework for studying such a question is the theory of interacting particle systems (IPS). An IPS is a stochastic model consisting of many particles evolving according to prescribed rules. The configuration of an IPS records the state of the system and evolves as a continuous-time Markov process. The word interacting in IPS means that the evolution of one particle depends on the other particles.

The symmetric exclusion process (SEP) is a well-known IPS introduced by Spitzer [93] in the 70's as a stochastic counterpart of a gas of particles subject to hard-core exclusion. This IPS will be central in this thesis, and thus we describe it in detail. Let $G = (V, E)$ be a graph. The state space of the SEP is $\{0, 1\}^V$ and a configuration is a function

$$\eta : V \longrightarrow \{0, 1\}$$

whose value at a vertex $x \in V$, called a site, means that there is a particle at the site x if $\eta(x) = 1$ and that the site is empty if $\eta(x) = 0$. The SEP evolves as follows. Each particle occupying a site is equipped with an independent Poisson clock. When the clock attached to a particle rings, this particle attempts to jump to one of the neighboring sites. The exclusion rule says that at most one particle may occupy each site; jumps to already

occupied sites are suppressed. Denoting by $W(x, y) = W(y, x)$ the symmetric jump rate between two neighboring sites x and y , the generator of the SEP is then given by

$$L^{\text{SEP}} f(\eta) = \sum_{(x,y) \in E} W(x, y) (f(\eta^{x,y}) - f(\eta)),$$

where $\eta^{x,y}$ is obtained from η by exchanging the values at sites x and y . The dynamics of the SEP preserve the total number of particles, which will be the conserved quantity of interest.

Suppose now that, instead of a fixed graph, one considers a sequence of graphs $\{G_N = (V_N, E_N)\}_N$, which should be thought of as finer and finer discretizations of a macroscopic space; as N increases, the distance between neighboring sites shrinks. If one were to keep the original microscopic time scale for the evolution of the SEP, the particles would only move over a vanishing macroscopic distance, and no evolution would be visible in the limit. This is the basic idea of a scaling limit; one observes the system at a macroscopic scale by rescaling space and time together. After this rescaling, one studies the evolution of the empirical density of particles. The empirical measure is defined as

$$\hat{\pi}_t^N = C_N \sum_{x \in V_N} \eta_t(x) \delta_x$$

where η_t is the configuration at time t of the SEP with rescaled generator given by

$$C'_N L_{G_N}^{\text{SEP}} f(\eta) = C'_N \sum_{(x,y) \in E_N} W_N(x, y) (f(\eta^{x,y}) - f(\eta)).$$

The coefficient C_N will go to zero while C'_N will increase to infinity, which accounts for the acceleration of time, as N goes to infinity. If

$$\pi_t^N \longrightarrow \rho_t(x) dx$$

in probability, in a suitable weak sense, and if the limit is deterministic, then one has obtained a hydrodynamic limit. The limiting density ρ_t is then expected to solve a partial differential equation (PDE).

For the SEP on the lattice $\mathbb{Z}^d$, in order to observe diffusion, one rescales space as $1/N$, which yields an empirical measure of the form

$$\hat{\pi}_t^N = \frac{1}{N^d} \sum_{x \in \mathbb{Z}^d} \eta_t(x) \delta_{\frac{x}{N}}$$

and accelerates time by N^2 . Under this scaling, called diffusive scaling, the empirical density of the SEP converges to the solution of the heat equation

$$\partial_t \rho_t = \Delta \rho_t.$$

The SEP is special among IPS in the sense that it possesses a self-duality structure. Although the particles interact through the exclusion rule, the expected occupation variable evolves as if it were governed by a single random walk (RW). Indeed, for a fixed site x ,

$$\frac{d}{dt}\mathbb{E}[\eta_t(x)] = \mathbb{E}[L^{\text{SEP}}\eta(x)].$$

A direct computation gives

$$L^{\text{SEP}}\eta(x) = \sum_{y \sim x} W(x, y)(\eta(y) - \eta(x)).$$

The right-hand side is the discrete generator of a RW applied to the function $x \mapsto \eta(x)$. Thus, the one-point function of the exclusion process is governed by the underlying RW generator

$$L^{\text{RW}}\varphi(x) = \sum_{y \sim x} W(x, y)(\varphi(y) - \varphi(x)).$$

This identity is a first manifestation of self-duality. More generally, correlation functions of the SEP can be expressed through another exclusion process with finitely many particles. For the hydrodynamic limit, the most important consequence is that the evolution of the density can be reduced to the analysis of a single RW. This connection with RW also hints at why the hydrodynamic limit is governed by the heat equation.

The heat equation is not limited to the Euclidean space. Its solution is given by the heat semigroup, while the heat kernel encodes many analytic properties of the space in which the equation takes place. This equation is at the heart of harmonic analysis, probability theory, and PDEs. This observation motivates the following question:

Can one derive heat equations from IPS beyond the Euclidean space?

The first kind of difficulty is that a Riemannian manifold does not come with a canonical lattice, in contrast with $\mathbb{R}^d$, where the lattice $\mathbb{Z}^d$ provides a natural discretization. In general, one has to choose a discretization. Regarding the SEP, van Ginkel–Redig [99] obtained the hydrodynamic limit towards the heat equation on compact Riemannian manifolds using random neighborhood graphs (RNG) from independently sampled vertices. Whenever two vertices are less than a certain distance, called the bandwidth parameter h_N , then they are connected by an undirected edge; together these vertices and edges form the RNG.

In this thesis, we consider the SEP on complete non-compact Riemannian manifolds. Recall that the relevant macroscopic conserved quantity is the total number of particles for the SEP. The limiting equation should therefore conserve mass as well. On a non-compact Riemannian manifold, this is not automatic: heat may escape to infinity in finite time, or equivalently Brownian motion may fail to be conservative. A natural way to rule out this behavior is to impose geometric assumptions, such as a lower bound on the Ricci curvature.

The construction of the SEP is also more delicate on non-compact Riemannian manifolds. One cannot simply sample finitely many vertices independently from the Riemannian volume measure, since this measure is infinite. Rather, the sequence of RNG is built from Poisson point processes whose intensities are proportional to a finite Gibbs measure

$$Nd\mu = Ne^{-U} d\text{Vol}_M.$$

As in the compact case, the choice of bandwidth parameter h_N goes to zero slowly enough as N goes to infinity; the degree of a vertex blows up in the limit. Yet contrary to the compact case, the existence of the SEP on such RNG is not automatic and will ultimately lead to the same geometric assumption of a lower bound on the Ricci curvature.

The hydrodynamic limit of the SEP takes the form of a Fokker–Planck equation

$$\partial_t \rho_t = \nabla \cdot (\nabla \rho_t + \rho_t \nabla U)$$

if the profile ρ_t is considered with respect to the volume measure $d\text{Vol}_M$. Equivalently,

$$\partial_t \rho_t = (\Delta_M - \nabla U \cdot \nabla) \rho_t$$

if the profile ρ_t is considered with respect to the Gibbs measure $d\mu$.

The same geometric viewpoint also allows us to go beyond elliptic diffusions. If the manifold has the structure of a principal bundle, one can distinguish horizontal directions from vertical directions. By choosing jump rates at two different scales, inspired by the horizontal diffusion maps of Gao [41], one can obtain sub-elliptic diffusions as hydrodynamic limits. The second-order limiting operator is then a horizontal Laplacian rather than the Laplace–Beltrami operator Δ_M .

In Euclidean space, Brownian motion, whose generator is $\frac{1}{2}\Delta$, can be obtained as the scaling limit of a centered RW. On a Riemannian manifold, the analogous construction is due to Jørgensen [62]. Since on a manifold one cannot add points, the next location of the RW is obtained by following a geodesic in a random direction. After diffusive rescaling, under appropriate assumptions mimicking independent increments, these geodesic RWs converge to Brownian motion on the manifold, whose generator is $\frac{1}{2}\Delta_M$.

A useful way to picture Brownian motion on a manifold M is to roll the manifold along a Euclidean Brownian path without slipping, known as the Eells–Elworthy–Malliavin construction. The Euclidean noise is then converted into intrinsic random motion on the manifold. In order to keep track of the directions while rolling, one works on the orthonormal frame bundle

$$\pi : O(M) \longrightarrow M.$$

The Brownian motion is first constructed upstairs as a horizontal Brownian motion and is then projected down to M . The corresponding relation between the operators is

$$\Delta_{\mathcal{H}}(f \circ \pi) = (\Delta_M f) \circ \pi, \quad f \in C^\infty(M),$$

where $\Delta_{\mathcal{H}}$ is the horizontal Laplacian on the orthonormal frame bundle $O(M)$.

It is thus natural to lift geodesic RWs horizontally and prove that their diffusive limit is a horizontal Brownian motion, which is part of this thesis. Since the discrete walks are related by the projection π , the limiting diffusions are also related by π , and the identity between $\Delta_{\mathcal{H}}$ and Δ_M is recovered probabilistically, as a consequence of an invariance principle. The relationship between $\Delta_{\mathcal{H}}$ and Δ_M is central in stochastic calculus on manifolds [57].

In this thesis, we also study another class of IPS, this time in the continuum. The particles no longer jump on a graph, but move directly on the manifold according to stochastic differential equations. The tools are therefore those of stochastic calculus on manifolds, in particular couplings on the orthonormal frame bundle. The systems we consider are mean-field interactions of McKean–Vlasov type. A typical example is

$$dX_t^{i,N} = \sqrt{2} dB_t^{i,M} - \nabla V(X_t^{i,N}) dt - \frac{1}{N} \sum_{j=1}^N \nabla_x W(X_t^{i,N}, X_t^{j,N}) dt, \quad i = 1, \dots, N.$$

Here, $B_t^{i,M}$ are independent Brownian motions on the manifold, V is a confinement potential, and W an interaction potential.

In the limit N goes to infinity, one expects the particles of the McKean–Vlasov IPS to become asymptotically independent. This phenomenon is known as propagation of chaos, and the empirical measure

$$\hat{\mu}_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}$$

converges weakly to the common law μ_t of the limiting particles that solves a nonlinear Fokker–Planck equation, also called an aggregation–diffusion equation, of the form

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot \left(\mu_t [\nabla V + \nabla (W * \mu_t)] \right), \quad W * \mu_t(x) = \int_M W(x, y) d\mu_t(y).$$

Propagation of chaos is, however, an asymptotic statement. It says that the particles become independent in the limit $N \rightarrow \infty$, but it does not by itself quantify how close the IPS is to its limiting law for a finite number of particles. From the point of view of simulations, it is important to have explicit estimates on the distance between the empirical measure of the IPS and the limiting law μ_t .

To obtain such estimates, it is natural to compare the interacting system with a simpler non-interacting proxy. In this auxiliary system, the empirical mean-field interaction created by the particles is replaced by the averaged interaction against the limiting law μ_t . The particles are then independent copies of the same nonlinear process; at each fixed time t , they have common law μ_t . If the interacting system remains close to this proxy, then the distance between $\hat{\mu}_t^N$ and μ_t can be controlled by combining stability estimates for

the dynamics with a quantitative law of large numbers for the empirical measure of the independent copies.

In Euclidean space, quantitative concentration estimates for empirical measures of independent samples in Wasserstein distance were obtained by Bolley–Guillin–Villani [11]. The methods used in this direction are metric in nature. They rely on estimates of covering numbers of balls, and therefore on the way the underlying space grows at large scales. In this thesis, we obtain a manifold analogue of these results and prove concentration estimates for the Wasserstein distance between an empirical measure and its underlying law, with explicit dependence on curvature. This makes precise how geometry of the space enters quantitative laws of large numbers for empirical measures.

The assumption of having a lower bound on the Ricci curvature also plays a role in the stability estimates for the McKean–Vlasov dynamics. Indeed, lower Ricci curvature bounds allow one to use comparison theory, and in particular estimates on the distance between coupled diffusions on the manifold. Thus curvature enters the analysis twice: first through metric covering estimates for empirical measures, and second through comparison estimates used to control the stability of the particle dynamics.

Finally, in the last part of this thesis we study a cornerstone equation of kinetic theory; the multi-species Landau equation

$$\partial_t f_i = \sum_{j=1}^N Q_{ij}(f_i, f_j), \quad 1 \leq i \leq N,$$

where $Q_{ij}(f_i, f_j)$ is the Landau collision operator, derived by Landau in 1936 [69]. This equation describes the evolution of a mixture of several species of particles undergoing grazing collisions. The unknown is a vector of velocity distributions

$$F = (f_1, \dots, f_N),$$

where $f_i(t, v)$ denotes the density of particles of species i of mass m_i with velocity $v \in \mathbb{R}^3$ at time t . The different species may have different masses, and the equation couples the densities through both self-interactions and cross-interactions between species.

Although this part of the thesis is not about hydrodynamic limits but about well-posedness that we establish, it is still connected to the general theme of the thesis. Indeed, the Landau equation is another macroscopic equation derived from microscopic dynamics, namely from the grazing collision limit of the Boltzmann equation. Moreover, the collision operator can be interpreted as a singular spherical Laplacian in an appropriate set of physical variables.

For the single-species Landau equation, it was recently proved by Guillin–Silvestre [52] that the Fisher information is a Lyapunov functional in the Coulomb case; the Fisher information decreases along the flow of the solution. In the multi-species setting, the

situation is more delicate. The different masses break some of the symmetries that are present in the single-species case. This has two unfortunate consequences. First, a key ingredient in the single-species argument, namely a logarithmic Sobolev inequality on the sphere, becomes weaker when it is applied to non-symmetric functions. Second, for sufficiently singular cross-interactions, the Fisher information itself ceases to be a Lyapunov functional. In particular, one finds a threshold below which monotonicity of the Fisher information can fail.

We introduce a new Lyapunov functional based not only on Fisher information but also on relative entropy and entropy dissipation to establish well-posedness below the critical threshold in the ion–electron regime.

1.2 Overview of the main results of the thesis

1.2.1 Chapter 2 – Invariance principle for lifted geodesic random walks

In the lattice $\mathbb{Z}^m$, the position of a simple symmetric random walk (RW) after N steps can be written as

$$S^{(N)} = \sum_{i=1}^N X^i, \quad X^i \text{ i.i.d. with law } \text{Unif}(\{\pm e_1, \dots, \pm e_m\}),$$

where $\{e_i\}_{i=1}^m$ is the canonical basis of $\mathbb{R}^m$. Since $\mathbb{E}[X^i] = 0$, $\text{Cov}(X^i) = \mathbb{E}[X^i \otimes X^i] = \text{Id}_{\mathbb{R}^m}/m$, the covariance of $S^{(N)}$ is $N\text{Id}_{\mathbb{R}^m}/m$. At the process level, Donsker's invariance principle states that

$$\left(\sqrt{\frac{m}{N}} S^{(\lfloor Nt \rfloor)} \right)_{t \geq 0}$$

converges in law to the standard Brownian motion on $\mathbb{R}^m$. Recall that a standard Brownian motion on $\mathbb{R}^m$ is a continuous stochastic process $(B_t)_{t \geq 0}$ such that $B_0 = 0$, its increments are independent, and

$$B_t - B_s \sim \mathcal{N}(0, (t-s)\text{Id}_{\mathbb{R}^m}), \quad 0 \leq s < t.$$

Its infinitesimal generator is $\frac{1}{2}\Delta$, half the Euclidean Laplacian.

On a manifold, that is, roughly speaking, a space which locally looks Euclidean, the preceding construction cannot be used directly. In general, there is neither a notion of lattice nor a vector space structure.

In the 70's, Jørgensen [62] introduced an analog of RW on Riemannian manifolds. The construction of the RW is the following: Starting from a point $x \in M$, one chooses a tangent vector $v \in T_x M$ according to a probability measure μ_x on $T_x M$. In local orthonormal

coordinates, one may think, for instance, of μ_x as the uniform measure on the unit sphere in $T_x M$. The next position of the walk is obtained by following for one unit of time the geodesic starting from x with initial velocity v . Equivalently, the transition operator is given by

$$Pf(x) = \int_{T_x M} f(\exp_x(v)) d\mu_x(v), \quad f \in C_c^\infty(M),$$

where $\exp_x : T_x M \rightarrow M$ denotes the exponential map.

Since addition of points is not defined on a manifold, one does not rescale the position of the walk itself. Instead, one rescales the tangent increments. For $\alpha > 0$, one considers the transition operator

$$P^{(\alpha)}f(x) = \int_{T_x M} f(\exp_x(\alpha v)) d\mu_x(v),$$

and accelerates time by the factor α^{-2} . Under suitable centering and covariance assumptions on the family $(\mu_x)_{x \in M}$, the rescaled geodesic RW converges to Riemannian Brownian motion [62, Theorem 2.1]. This is the manifold analog of Donsker's invariance principle.

Stochastic calculus on Riemannian manifolds is often performed not directly on M , but rather on its orthonormal frame bundle $\pi : O(M) \rightarrow M$, where local coordinates are described as those of M augmented by a choice of orthonormal basis of the tangent space. A standard reference for stochastic calculus on manifolds is the monograph of Hsu [57].

The advantage of $O(M)$, despite its larger dimension, is that it is a parallelizable manifold and thus carries global smooth vector fields giving at each point a basis, in this case the fundamental horizontal and vertical vector fields. Each vector e_i of the canonical basis of $\mathbb{R}^m$ gives rise to a horizontal vector field $H_i \in \Gamma(TO(M))$, where $\Gamma(\cdot)$ denotes the space of vector fields, obtained by horizontally lifting the tangent vector $ue_i \in T_{\pi(u)}M$. The horizontal lift $\tilde{\gamma}$ starting from u of a smooth curve γ in M is the moving frame that satisfies

$$\tilde{\gamma}e_i = V_i, \quad \nabla_{\tilde{\gamma}'} V_i = 0, \quad V_i(0) = ue_i,$$

or equivalently in local coordinates, $\tilde{\gamma}(t) = (\gamma^j(t), V_i^j(t))_{1 \leq i, j \leq m}$ with

$$\dot{V}_i^j(t) + \Gamma_{kl}^j \dot{\gamma}^k(t) V_i^l(t) = 0.$$

These horizontal vector fields define the Bochner horizontal Laplacian

$$\Delta_{\mathcal{H}} = \sum_{i=1}^m H_i^2$$

as a sum of squares, an operator in Hörmander form. The corresponding horizontal Brownian motion on $O(M)$ with infinitesimal generator $\frac{1}{2}\Delta_{\mathcal{H}}$ projects onto Riemannian Brownian

motion on M [57, Proposition 3.1.2]. Analytically, this is expressed by the identity

$$\Delta_{\mathcal{H}}(f \circ \pi) = (\Delta f) \circ \pi, \quad f \in C^\infty(M). \quad (1.2.1)$$

This identity is a basic mechanism in stochastic analysis on manifolds: one constructs a process on the orthonormal frame bundle, where the driving noise is Euclidean and globally defined, and then projects it back to the manifold. This also allows one to construct Riemannian Brownian motion by rolling a Euclidean Brownian motion without slipping on the manifold; this is the Eells–Elworthy–Malliavin construction.

We revisit Jørgensen’s construction. Since Brownian motion on M can be obtained as a scaling limit of random walks (RW), it is therefore natural to lift these RW horizontally to the orthonormal frame bundle. We prove an invariance principle for these lifted geodesic RW by a simple expansion of the transition probability. The limiting process is the horizontal Brownian motion with generator $\frac{1}{2}\Delta_{\mathcal{H}}$. Since by construction the random walks and horizontally lifted random walks are π -related processes, the relation (1.2.1) between the horizontal Laplacian and the Laplace–Beltrami operator is then recovered as a consequence of the invariance principle we obtain on $O(M)$. This gives a probabilistic interpretation of (1.2.1). Actually, we give such an interpretation in the more general setting of Riemannian submersions with totally geodesic fibers and bracket-generating horizontal distribution, which is the setting of the recent monograph of Baudoin–Demni–Wang [7] on stochastic areas, horizontal Brownian motion, and hypoelliptic heat kernels.

1.2.2 Chapter 3 – Hydrodynamic limit of the symmetric exclusion process on manifolds

We now turn to interacting particle systems whose microscopic state space is discrete. The paradigmatic example is the symmetric exclusion process (SEP), introduced by Spitzer [93] in the 1970s. On a graph $G = (V, E)$, a configuration is a map

$$\eta : V \rightarrow \{0, 1\},$$

where $\eta(x) = 1$ means that a particle occupies the site x , while $\eta(x) = 0$ means that the site is empty. The exclusion rule imposes that at most one particle can occupy each site. Particles attempt to jump along the edges of the graph, but jumps to already occupied sites are suppressed. More precisely, if $W(x, y) = W(y, x) \geq 0$ denotes the jump rate along the edge (x, y) , the generator of the SEP is given, on local functions $f : \{0, 1\}^V \rightarrow \mathbb{R}$, by

$$L^{\text{SEP}} f(\eta) = \sum_{(x,y) \in E} W(x, y) (f(\eta^{x,y}) - f(\eta)),$$

where $\eta^{x,y}$ is the configuration obtained from η by exchanging the occupations at x and y .

A hydrodynamic limit is a law of large numbers for the empirical density of the system. On $\mathbb{Z}^m$, one observes the system at the macroscopic scale by rescaling space by N^{-1} and

time by N^2 . The empirical density field is then given by the random measure

$$\widehat{\pi}_t^N = \frac{1}{N^m} \sum_{x \in \mathbb{Z}^m} \eta_{N^2 t}(x) \delta_{x/N}.$$

A hydrodynamic limit means that, as $N \rightarrow \infty$, this random measure-valued path converges in probability, in a suitable weak sense, to a deterministic path $\rho_t(u) du$, where the limiting density ρ_t solves a macroscopic PDE. For the SEP on $\mathbb{Z}^m$ with nearest-neighbor symmetric jumps, the limiting equation is the heat equation

$$\partial_t \rho_t = \Delta \rho_t. \quad (1.2.2)$$

We refer to the monograph of Kipnis–Landim [67] for an extensive account of scaling limits of IPS.

The SEP is particularly tractable because it has a microscopic closure property. Indeed, the expected occupation variable evolves according to the generator of the underlying RW. More precisely, if

$$L^{\text{RW}} \varphi(x) = \sum_{y \sim x} W(x, y) (\varphi(y) - \varphi(x)), \quad \varphi : V \rightarrow \mathbb{R},$$

then a direct computation gives

$$L^{\text{SEP}} \eta(x) = L^{\text{RW}} \eta(x).$$

Thus the one-point function of the SEP is governed by the same operator as a single RW. This is a manifestation of the self-duality of the exclusion process; see Giardinà–Redig [43] for a recent monograph on duality from the Lie algebraic approach.

On Riemannian manifolds, on the other hand, there is no canonical lattice and one must therefore choose a discretization. We take inspiration from machine learning where high-dimensional data embedded in Euclidean space is supposed to lie on a manifold of smaller intrinsic dimension; this is the manifold hypothesis, a posit in machine learning. Random neighborhood graphs (RNG) are natural in manifold learning and used to approximate its geometry. In our setting, the vertices are given by a Poisson point process Λ^N on M with intensity $N d\mu$ and Gibbs measure

$$d\mu = e^{-U} d\text{Vol}_M,$$

where U is a smooth potential bounded from below. Two points $x, y \in \Lambda^N$ are connected whenever $d_M(x, y) \leq h_N$, where $h_N \downarrow 0$ is the bandwidth parameter.

On compact Riemannian manifolds, random neighborhood graphs provide consistent discrete approximations of the Laplace–Beltrami operator as shown in the seminal work of

Hein–Audibert–von Luxburg [56]. In our non-compact weighted setting, if the symmetric jump rates are of the form

$$W_{N;\alpha}(x, y) = \frac{1}{N} \frac{h_N^{-m} k\left(\frac{d(x, y)}{h_N}\right)}{\left(\widehat{k}_N(x) \widehat{k}_N(y)\right)^\alpha}, \quad \widehat{k}_N(x) = \frac{1}{N} \sum_{y \in \Lambda^N} h_N^{-m} k\left(\frac{d(x, y)}{h_N}\right), \quad (1.2.3)$$

for some parameter $\alpha \geq 0$, where $k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a bounded measurable kernel supported in $[0, 1)$ continuous near 0 with $k(0) > 0$; $\widehat{k}_N$ is a statistical estimator of e^{-U} .

after the diffusive rescaling by h_N^{-2} , we prove that the corresponding RW generator converges, up to a constant, uniformly on compact sets and almost surely in the realization of the graphs, to a weighted Laplacian $e^{(2\alpha-1)U} \Delta_{2(1-\alpha)}$, which is the Witten Laplacian Δ_U when $\alpha = 1/2$, given in that case by

$$\Delta_U f = e^U \nabla \cdot (e^{-U} \nabla f) = \Delta f - \nabla U \cdot \nabla f, \quad f \in C^\infty(M).$$

Combining this convergence of the RW with the duality of the SEP, we prove the quenched hydrodynamic limit. Namely, for almost every realization of the random neighborhood graphs, the empirical density field of the SEP with jump rates (1.2.3), rescaled in time by h_N^{-2} , converges in probability to the deterministic trajectory $(\rho_t d\mu)_{t \in [0, T]}$, where ρ_t weakly solves

$$\partial_t \rho_t = C_{k;\alpha} e^{(2\alpha-1)U} \Delta_{2(1-\alpha)} \rho_t = C_{k;\alpha} e^{(2\alpha-1)U} \left(\Delta - 2(1-\alpha) \nabla U \cdot \nabla \right) \rho_t. \quad (1.2.4)$$

The convergence is understood in the Skorokhod topology after testing against compactly supported smooth functions. Note the presence of the potential U in (1.2.4), which generalizes, even in the Euclidean setting, the heat equation (1.2.2) obtained as the hydrodynamic limit of the SEP on the lattice $\mathbb{Z}^m$.

The non-compact setting creates two difficulties which are absent in the compact case treated by van Ginkel–Redig [99]. First, the RNG are infinite. Therefore the SEP is not automatically well-defined; from the graphical construction, one must rule out explosion of the underlying RW. Second, the limiting diffusion on the manifold must be conservative. On a non-compact manifold, heat may escape to infinity in finite time, or equivalently the diffusion generated by, say, Δ_U may fail to have total mass one. Thus stochastic completeness of the weighted manifold is a natural assumption in the hydrodynamic limit. In terms of the heat kernel p_U associated with Δ_U , this means that

$$\int_M p_U(t, x, y) d\mu(y) = 1, \quad t > 0, x \in M.$$

We prove non-explosion of the RW on the RNG by adapting a stochastic-completeness criterion for weighted graphs due to Folz [38], which is the graph analogue of Grigor'yan's

criterion on manifolds [47]. The idea is to introduce an intrinsic graph metric whose edge lengths are shorter where the jump rates are larger. This metric is adapted to the RW and allows one to compare graph balls with Riemannian balls. Using Bishop–Gromov volume comparison under a lower Ricci curvature bound, we obtain the required volume-growth control for the graph. Folz’s criterion then yields non-explosion of the RW, and hence existence of the SEP by the graphical construction.

We also extend the construction to principal bundles $\pi : \widetilde{M} \rightarrow M$. The idea is to discretize both the base manifold and the fibers, while using the connection to compare points in different fibers. Starting from the Poisson cloud $\Lambda^N \subset M$, we sample additional points in each fiber $\widetilde{M}_x = \pi^{-1}(x)$. Thus a vertex of the lifted graph is a pair (x, u) , with $x \in \Lambda^N$ and $u \in \widetilde{M}_x$. Two vertices (x, u) and (y, q) are connected when their base points are close and their fiber components are close after identifying the fibers by parallel transport;

$$d_M(x, y) < h_N, \quad d_{\widetilde{M}_x}(u, \widetilde{P}_{y \rightarrow x} q) < h'_{N'}.$$

Here $\widetilde{P}_{y \rightarrow x} : \widetilde{M}_y \rightarrow \widetilde{M}_x$ denotes the parallel transport induced by the connection. The geometric assumption imposed on the principal bundle covers the case of the orthonormal frame bundle.

The two bandwidths have different roles. The parameter h_N controls the motion in the base manifold, while $h'_{N'}$ controls the fiber thermalization. With a suitable choice of weights and a double-scaling limit, we show that the associated single RW converges to a horizontal diffusion. By duality, the corresponding lifted SEP has a hydrodynamic limit given by a sub-elliptic heat equation driven by a weighted horizontal Laplacian. This construction is inspired by Gao’s horizontal diffusion maps [41], where pairwise correspondences between fibers are used to build discrete operators converging to horizontal diffusion operators.

1.2.3 Chapter 4 – Concentration and expectations for empirical measures on manifolds

Given independent random variables $\{X^i\}_{1 \leq i \leq N}$ with common law ν on a smooth Riemannian manifold (M, g) , the associated empirical measure

$$\widehat{\nu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{X^i}$$

is a random probability measure which encodes the distribution of the samples. By the law of large numbers, also known as the Glivenko–Cantelli theorem (and due to Varadarajan on metric spaces [100]), $\widehat{\nu}^N$ converges weakly to ν as $N \rightarrow \infty$. A natural quantitative question is then: how far is $\widehat{\nu}^N$ from ν for possibly large but finite N ? Since the Wasserstein distance metrizes weak convergence [105, Theorem 6.9], it is natural to quantify this error in the

former; we refer to the monograph of Villani [105] for an extensive account on optimal transport.

In the Euclidean space $\mathbb{R}^m$, the convergence of empirical measures in Wasserstein distance is a classical topic, already dating back to the late 1960's with the seminal work of Dudley [31]. The rate of convergence depends on the dimension m , on the value of p , and on the moments of the law. The higher the dimension, the harder it is to approximate a probability measure by finitely many points; this is referred to as the curse of dimensionality. This dimensional obstruction is already present for measures with compact support: one needs enough points to cover the support at the relevant scale.

On general Polish metric spaces (E, d) , the convergence of empirical measures is related to the covering numbers of E , namely the smallest number of balls of a certain radius needed to cover E . For instance, on a compact space, estimates on the covering number $N(E, \delta)$ lead directly to effective bounds on $\mathbb{E}[W_p(\hat{\nu}^N, \nu)]$; see Boissard–Le Gouic [10, Theorem 1.1]. On a non-compact Riemannian manifold, however, the covering number of the whole space is infinite. One must therefore combine a local covering estimate on large balls with a tail estimate controlling the probability of being outside of these balls.

We place ourselves on metric measure spaces satisfying a measure contraction property. We shall not recall the full definition here, but one should think of it as a generalization of the synthetic lower curvature bound of curvature-dimension condition of Lott [77] and Sturm–Villani [96]; if $(E, d, \mathfrak{m})$ satisfies $\text{MCP}(K, n)$ for some $K \in \mathbb{R}$ and $n > 1$, then an analog of Bishop–Gromov theorem holds [85, Theorem 5.1], and for an m -dimensional Riemannian manifold, $\text{MCP}(K, m)$ coincides with $\text{Ric} \geq K$. In particular, when $K \leq 0$, which is the non-positive curvature setting, we obtain the covering estimate

$$N(B(x, R), \delta) \leq \left(\frac{5R}{\delta}\right)^n \exp\left(2\sqrt{|K|(n-1)R}\right). \quad (1.2.5)$$

We obtain quantitative bounds in expectation of the p -Wasserstein distance between the empirical measure of i.i.d. samples and their common law. The strategy is to truncate the measure on a large ball, apply compact estimates on that ball, and then optimize the radius of the ball as a function of N . The optimal choice of the radius depends on two competing effects: the covering number of the ball estimated by (1.2.5), which grows with the radius, and the tail of the measure, which becomes smaller as the radius increases. In spaces of polynomial volume growth, this leads to Euclidean-like rates, while in negatively curved spaces, the exponential growth of covering numbers forces stronger tail assumptions, typically exponential moments, in order to recover sharp polynomial rates as in the Euclidean space. When $K \leq 0, n > 2p \geq 2$, and the probability measure ν satisfies the exponential moment condition $\int_E e^{\alpha d(o, x)} d\nu(x) < +\infty$ for some $o \in E$ and $\alpha > \frac{2p}{n-p} \sqrt{|K|(n-1)}$, we obtain for every $N \geq 1$, the sharp rate

$$\mathbb{E}[W_p(\hat{\nu}^N, \nu)] \leq CN^{-\frac{1}{n}}.$$

The second type of result concerns concentration inequalities. We obtain quantitative versions of Sanov's theorem, which describes the large deviations of empirical measures in terms of the relative entropy, generalizing the seminal work of Bolley–Guillin–Villani [11] from the Euclidean space to manifolds. Recall that if ν' is absolutely continuous with respect to ν , the relative entropy is

$$H(\nu' \mid \nu) = \begin{cases} \int_E \log \left(\frac{d\nu'}{d\nu} \right) d\nu' & \text{if } \nu' \ll \nu; \\ +\infty & \text{otherwise.} \end{cases}$$

Sanov's theorem says roughly that

$$\mathbb{P} [\text{dist}(\hat{\nu}^N, \nu) > \varepsilon] \approx e^{-NI(\varepsilon)}, \quad I(\varepsilon) := \inf \{H(\nu' \mid \nu) : \text{dist}(\nu', \nu) > \varepsilon\}$$

for every distance dist continuous with respect to the weak topology. To turn this asymptotic statement into a quantitative one, we again truncate the measure on large balls whose radius comes from optimization considerations. We choose to track each constant to highlight the precise effect of the curvature lower bound compared to the Euclidean setting; this makes the proofs rather technical. If the measure only has polynomial decay $\int_E d^q(o, x) d\nu(x) < +\infty$ for some $q > p \geq 1$ and the lower bound on Ricci curvature is $\bar{K} \leq 0$, then the same concentration estimate as in $\mathbb{R}^n$ holds provided N is sufficiently large, say, for $n' > n$, we require

$$N \geq N_0 \varepsilon^{-\frac{q(n'+2p)}{q-p}} \exp \left(2\sqrt{|K|(n-1)} \varepsilon^{-\frac{p}{q-p}} \right),$$

and $K = 0$ recovers exactly the Euclidean threshold for N . But when the measure decays fast enough, say it has Gaussian tails, then the threshold for N is proportional to the Euclidean one.

These estimates for independent samples are strong enough to study an IPS with mean-field interaction of McKean–Vlasov type on a Riemannian manifold. More precisely, we consider particles evolving according to an aggregation-diffusion dynamics of the form

$$dX_t^i = \sqrt{2} dB_t^{i,M} - \nabla V(X_t^i) dt - \frac{1}{N} \sum_{j=1}^N \nabla_x W(X_t^i, X_t^j) dt, \quad i = 1, \dots, N,$$

where $B_t^{i,M}$ are independent Riemannian Brownian motions, V is a confinement potential, and W is an interaction potential. The associated time-dependent empirical measure is

$$\hat{\mu}_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}.$$

In the many-particle limit, one expects propagation of chaos: the particles become asymptotically independent and their common law μ_t solves the weakly nonlinear aggregation-diffusion

equation

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot \left(\mu_t [\nabla V + \nabla (W * \mu_t)] \right), \quad W * \mu_t(x) = \int_M W(x, y) d\mu_t(y).$$

Since our quantitative estimates only apply to independent particles, we introduce the non-interacting proxy system where mean-field interaction through the empirical measure of the particles is replaced by the limiting law μ_t ;

$$dY_t^i = \sqrt{2} d\bar{B}_t^{i,M} - \nabla V(Y_t^i) dt - \nabla (W * \mu_t)(Y_t^i) dt, \quad i = 1, \dots, N.$$

For each fixed time t , the variables $\{Y_t^i\}_{1 \leq i \leq N}$ are independent with common law μ_t , and the concentration estimates for empirical measures on manifolds can be applied to

$$\hat{\nu}_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}.$$

We then compare the interacting system with the non-interacting proxy by coupling the particles pairwise; for each i , we couple X_t^i and Y_t^i by parallel coupling on the manifold. The idea is that the two particles should be driven by the same Brownian noise, but interpreted in their respective tangent spaces by parallel transport along the geodesic joining them.

Under technical assumptions on the potentials V and W , if μ_0 has Gaussian tails, we prove for every fixed $T > 0$ a uniform-in-time concentration estimate over $[0, T]$ of the form

$$\mathbb{P} \left[\sup_{0 \leq t \leq T} W_1(\hat{\mu}_t^N, \mu_t) > \varepsilon \right] \leq C (1 + T\varepsilon^{-2}) \exp(-C' N\varepsilon^2),$$

for all $0 < \varepsilon < 1$ and N sufficiently large but finite, depending on ε in the same fashion as in the Euclidean case given by [11, Theorem 2.9].

1.2.4 Chapter 5 – Well-posedness of the multi-species Landau equation

The last part of this thesis concerns a cornerstone kinetic equation from plasma physics. The spatially homogeneous multi-species Landau equation [69] describes the evolution of a gas mixture composed of several species undergoing grazing collisions. The unknown is a vector of velocity distributions $F = (f_1, \dots, f_N)$, where $f_i(t, v)$ denotes the density of particles of species i with velocity $v \in \mathbb{R}^3$ at time t . The species may have different masses, and the equation couples the densities through both intra-species and inter-species collision operators:

$$\partial_t f_i = \sum_{j=1}^N Q_{ij}(f_i, f_j), \quad i = 1, \dots, N. \quad (1.2.6)$$

Here Q_{ii} describes collisions between particles of the same species, while Q_{ij} , $i \neq j$, describes collisions between particles of different species;

$$Q_{ij}(f_i, f_j)(v_i) = \frac{1}{m_i} \nabla_{v_i} \cdot \left(\int_{\mathbb{R}^3} c_{ij} A(v_i - v_j) \left(\frac{\nabla_{v_i}}{m_i} - \frac{\nabla_{v_j}}{m_j} \right) (f_i \otimes f_j)(v_i, v_j) dv_j \right),$$

where m_i and m_j are the respective masses of species i and j , $c_{ij} > 0$ a symmetric physical constant, and

$$A(z) = |z|^{2+\gamma} \Pi_z^\perp, \quad \Pi_z^\perp = \text{Id}_{\mathbb{R}^3} - \frac{z \otimes z}{|z|^2}$$

with the parameter $\gamma \in [-3, 1]$ encoding the type of interactions ($\gamma = -3$ corresponds to Coulomb interactions).

The Landau equation possesses a fundamental thermodynamic structure. It conserves the natural collision invariants, namely the total mass of each species, the total momentum, and the total energy of the mixture, while it also dissipates the Boltzmann entropy

$$\mathcal{H}(F) = \sum_{i=1}^N \int_{\mathbb{R}^3} f_i \log f_i dv.$$

This entropy structure is one of the main tools in the analysis of kinetic equations. However, for the Landau equation with soft interactions, which are the most singular and physically relevant ones in plasma physics, entropy dissipation alone is not sufficient to obtain a satisfactory well-posedness theory.

In the single-species Coulomb case, Villani explains in [102, Chapter 3] that finite-time blow up may be expected from an analogy between the Landau equation and the nonlinear heat equation, since the Landau equation takes the form

$$\partial_t f = -\text{Hess}(-\Delta)^{-2} f : \text{Hess} f + f^2$$

resembling $\partial_t f = \Delta f + f^2$, which is known to blow-up in finite time. However, after seeing some numerical simulations he became convinced that blow-up should not occur.

In the single-species Coulomb Landau equation, the recent breakthrough of Guillen–Silvestre [52] shows that the Fisher information

$$I(f) = \int_{\mathbb{R}^3} |\nabla \log f|^2 f dv$$

is a Lyapunov functional; it is nonincreasing along solutions to the Landau equation. This monotonicity relies on the log-Sobolev inequality on the sphere for symmetric functions. The bound on Fisher information then yields well-posedness.

A natural question is then whether this Fisher-information mechanism holds for multi-species mixtures. We introduce the correct analog of Fisher information, that is, a mass-weighted Fisher information of each species,

$$\mathcal{I}(F; a) = \sum_{i=1}^N a_i \int_{\mathbb{R}^3} |\nabla \log f_i|^2 f_i dv, \quad a_i = \frac{1}{m_i},$$

where the choice of weights is distinguished by the equilibrium structure: it is the choice for which products of Maxwellians with a common temperature are minimizers of $\mathcal{I}(F; a)$ under the constraint of fixed total energy.

Following Guillen–Silvestre, the dissipation of Fisher information is not studied directly on the original equation, but through a lifted linear equation obtained by doubling the variables. The idea is to introduce, for each species, an independent copy of the velocity variable. More precisely, we consider

$$v_{2N} = (v_1, \dots, v_N, v_{N+1}, \dots, v_{2N}) \in (\mathbb{R}^3)^{2N},$$

with the convention

$$m_{N+i} = m_i, \quad 1 \leq i \leq N.$$

On this space, we introduce the lifted linear operators

$$\tilde{Q}_{ij}(\tilde{F}) = \left(\frac{\nabla_{v_i}}{m_i} - \frac{\nabla_{v_j}}{m_j} \right) \cdot \left[c_{ij} A(v_i - v_j) \left(\frac{\nabla_{v_i}}{m_i} - \frac{\nabla_{v_j}}{m_j} \right) \tilde{F} \right],$$

for $1 \leq i \leq N$ and $N+1 \leq j \leq 2N$ and consider the linear Cauchy problem

$$\partial_t \tilde{F} = \sum_{\substack{1 \leq i \leq N \\ N+1 \leq j \leq 2N}} \tilde{Q}_{ij}(\tilde{F}), \quad (1.2.7)$$

with initial datum

$$\tilde{F}^\circ(v_1, \dots, v_{2N}) = \prod_{i=1}^N f_i^\circ(v_i) f_i^\circ(v_{N+i}).$$

Its usefulness comes from the fact that the first N marginals of $\tilde{F}$ solve the original multi-species Landau system (at time $t = 0$). Moreover, we show that the time derivative of a lifted version of the weighted Fisher information along the linear Cauchy problem (1.2.7) controls the time derivative of the weighted Fisher information along the multi-species Landau equation (1.2.6). Thus estimates proved at the level of the lifted linear system can be projected back to estimates for the nonlinear system.

The lifted formulation is precisely the right setting to study Fisher information. Indeed, the operators $\tilde{Q}_{ij}$ act only in the relative direction between two particles, while preserving

the corresponding center-of-mass direction. After passing to relative-velocity variables and center-of-mass variables defined as

$$z_{ij} = v_i - v_j, \quad z_{ij}^* = \frac{m_i v_i + m_j v_j}{m_i + m_j}, \quad (1.2.8)$$

it turns out that $\tilde{Q}_{ij}$ is a spherical Laplacian; setting $z_{ij} = r\omega$, and $f_i(v_i)f_j(v_j) = \bar{f}_{ij}(r, \omega)$,

$$\tilde{Q}_{ij}(f_i(v_i)f_j(v_j)) = c_{ij}\alpha_{ij}^2 r^\gamma \Delta_{\mathbb{S}^2} \bar{f}_{ij}(r, \omega), \quad \alpha_{ij} = \frac{1}{m_i} + \frac{1}{m_j}.$$

This is the reason why ultimately the log-Sobolev inequality on the sphere will enter.

However, the multi-species case contains a new obstruction. In the single-species equation, the lifted density of the form $f(v)f(v_*)$ is symmetric under the exchange $v \leftrightarrow v_*$. This symmetry is crucial as the constant in the log-Sobolev inequality improves for symmetric functions (see [52, Lemma 9.1] and [60, Theorem 1.1]), and without this improvement the sign of the time derivative of Fisher information remains unclear. In the multi-species equation, the corresponding two-particle density is instead $f_i(v_i)f_j(v_j)$ which is not symmetric under the exchange $v_i \leftrightarrow v_j$ when $m_i \neq m_j$. The different masses tilt the relation between relative velocity, center-of-mass velocity, and the Euclidean gradients appearing in the Fisher information.

As a consequence, the spherical logarithmic Sobolev inequality still produces a coercive term, but the constant obtained from the non-symmetric lifted density is not always strong enough to control the terms generated by the mass asymmetry. This is the central difficulty: the weighted Fisher information is the correct quantity for well-posedness, but it need not be a Lyapunov functional for arbitrary multi-species mixtures.

We make this obstruction precise; we identify an explicit regime $\gamma \in [-\sqrt{8}, -2)$ in which the coercive spherical term dominates the error terms. This gives the a priori Fisher-information estimate needed for well-posedness.

We then show that outside this favorable regime, in particular for Coulomb interactions corresponding to $\gamma = -3$, the obstruction is genuine. For the two-species setting with large mass ratio, we construct a counterexample showing that the weighted Fisher information ceases to be a Lyapunov functional. Thus the breakdown of the single-species argument reflects the lack of symmetry of inter-species collisions in the lifted physical variables (1.2.8).

The final part of the chapter treats the ion–electron regime, where one species is much heavier than the other. In this regime, we identify a new Lyapunov functional combining the Fisher information with relative entropy and entropy dissipation along the inter-species collision operator. These additional terms compensate for the loss of coercivity caused by the asymmetry and allow us to recover a closed a priori estimate, leading to well-posedness for Coulomb interactions.

CHAPTER 2

Invariance Principle for Lifted Geodesic Random Walks

This chapter is the content of the joint work [64] with Frank Redig, and Rik Versendaal. “Invariance Principle for Lifts of Geodesic Random Walks.” *Journal of Theoretical Probability* 39.2 (2026): 21.

2.1 Introduction

In this chapter we consider geodesic random walks on a Riemannian manifold (M, g) and their horizontal lift into a manifold $(N, \tilde{g})$ such that there is a Riemannian submersion $\pi : N \rightarrow M$. A motivating example of this setting is the orthonormal frame bundle $\pi_{O(M)} : O(M) \rightarrow M$ of a Riemannian manifold. This example is the basis of the Eells–Elworthy–Malliavin construction of Brownian motion. The important point in this setting is that the horizontal Brownian motion has as a generator the horizontal Laplacian which is a sum of squares of globally defined vector fields; i.e., it is in *Hörmander form*

$$\Delta_{\mathcal{H}} = \sum_{i=1}^d H_i^2,$$

where d is the dimension of the manifold. Because of this, the Markov process $\{U(t)\}_{t \geq 0}$ generated by $\frac{1}{2} \Delta_{\mathcal{H}}$ can be constructed as the solution of a Stratonovich SDE driven by an $\mathbb{R}^d$ -valued Brownian motion $\{W(t)\}_{t \geq 0}$, namely (see the monograph of Hsu [57]),

$$dU(t) = \sum_{i=1}^d H_i(U(t)) \circ dW^i(t).$$

Moreover, the Brownian motion on the manifold is the projection of this horizontal Brownian motion, based on the fact that

$$\Delta_{\mathcal{H}}(f \circ \pi) = (\Delta_M f) \circ \pi \tag{2.1.1}$$

for all smooth $f : M \rightarrow \mathbb{R}$. The proof of identity (2.1.1) in [57] is based on computations not easily followed by probabilists, using the π -relations between the covariant derivative and its lifted counterpart. Beyond the setting of the orthonormal frame bundle, horizontal Brownian motion is extensively studied in the monograph of Baudoin–Demni–Wang [7].

Brownian motion on M can be obtained as a scaling limit of geodesic random walks as initially considered by Jørgensen [62]. It is therefore natural to lift these walks horizontally in order to obtain horizontal Brownian motion in the scaling limit. As a consequence of such a weak convergence result, the horizontal Brownian motion on the total manifold N and the Brownian motion on the base manifold M are then π -related automatically. It is precisely the aim of this chapter to prove such result for a class of geodesic random walks, in the setting of Riemannian submersions, which is the framework in the monograph [7]. In several sub-Riemannian settings, motivated by different contexts such as Carnot groups, equivalence of Laplacians, and horizontal sub-Laplacians as limits of Hamiltonian flows, several authors have studied scaling limits of geodesic random walks; see [2, 13, 45, 46, 97].

First, in Section 2.2 we introduce the horizontal random walks and recall the notion of horizontal lift. Second, in Section 2.3 we consider general submersions for which we prove the π -relation as a corollary of an invariance principle, and we provide several examples, including the orthonormal frame bundle, which we discuss in detail.

2.2 Random walks and horizontal random walks

In this section we introduce the stochastic processes we study, namely, horizontal random walks. To do so, we first introduce the analog of random walks in M , so-called geodesic random walks, following [62, 101]. Afterwards, we explain how these geodesic random walks can be lifted to the total space N along a Riemannian submersion $\pi : N \rightarrow M$.

2.2.1 Geodesic random walk

We consider a d -dimensional geodesically complete Riemannian manifold M with metric g , and denote by $T_x M$ the tangent space of M at $x \in M$. In order to describe increments of our random walks, we have to consider a collection of probability measures μ_x on $T_x M$ called a distribution of *increments*; the nomenclature being inspired from [62] where μ_x describes the direction in which the random walk follows a geodesic when starting from x .

More precisely, we define the following Markov processes based on $\{\mu_x\}_{x \in M}$:

Definition 2.1. 1. *The discrete-time unit speed random walk $\{S_k\}_{k \in \mathbb{N}}$ based on the measures $\{\mu_x\}_{x \in M}$ is defined via its transition operator*

$$Pf(x) := \mathbb{E}[f(S_{k+1}) | S_k = x] = \int_{T_x M} f(\exp_x(v)) \mu_x(dv); \quad (2.2.1)$$

2. The discrete-time random walk with speed α based on the measures $\{\mu_x\}_{x \in M}$ is denoted by $\{S_k^{(\alpha)}\}_{k \in \mathbb{N}}$ and is defined via its transition operator

$$P^{(\alpha)} f(x) := \mathbb{E} \left[f \left(S_{k+1}^{(\alpha)} \right) \middle| S_k^{(\alpha)} = x \right] = \int_{T_x M} f(\exp_x(\alpha v)) \mu_x(dv); \quad (2.2.2)$$

3. Finally, the continuous-time process $\{Z_t^{(\alpha)}\}_{t \geq 0}$ is defined via its generator

$$L^{(\alpha)} f(x) := \alpha^{-2} \left(P^{(\alpha)} f(x) - f(x) \right). \quad (2.2.3)$$

The process $\{S_k\}_{k \in \mathbb{N}}$ evolves as follows: whenever $S_k = x$, S_{k+1} is obtained by randomly choosing X_{k+1} on $T_x M$ according to the measure μ_x and following the geodesic starting at x in the direction X_{k+1} for time 1, and analogously for the walk with speed scaled by α .

In what follows, we want to prove weak convergence to (horizontal) Brownian motion for the continuous walk $\{Z_t^{(\alpha)}\}_{t \geq 0}$ and its horizontal lift (defined below) as α tends to zero. As consequence, one can obtain corresponding results for the discrete walk $\{S_{\lfloor \alpha^{-2} t \rfloor}^{(\alpha)}\}_{t \geq 0}$ as α tends to zero (see Remark 2.13).

In order to proceed, we need some conditions on the distribution of increments. Because we aim at proving convergence to Brownian motion, there is a centering and variance condition. Finally, in order to prove uniform convergence of generators, it is convenient to have an additional third moment condition. More precisely, we make the following assumptions:

Assumption 2.2 (Centering and covariance). *For every $x \in M$, the measure μ_x has zero expectation and its covariance equals the inverse metric; i.e.,*

$$\int_{T_x M} v \mu_x(dv) = 0, \quad \int_{T_x M} v \otimes v \mu_x(dv) = g^{-1}(x),$$

or equivalently, in any smooth coordinate system about x ,

$$\int_{T_x M} v^i \mu_x(dv) = 0, \quad \int_{T_x M} v^i v^j \mu_x(dv) = g^{ij}(x), \quad i, j = 1, \dots, d. \quad (2.2.4)$$

Assumption 2.3 (Third moment condition). *The third moment of the collection of measures $\{\mu_x\}_{x \in M}$ is finite, uniformly on compacts; i.e., for all $C \subset M$ compact,*

$$\sup_{x \in C} \int_{T_x M} \|v\|^3 \mu_x(dv) < +\infty.$$

Remark 2.4. *If $p \mapsto \mu_x$ is invariant under parallel transport as considered by Jørgensen [62] in order to mimic identically distributed increments, then if Assumptions 2.2 and 2.3 are satisfied for a single $x \in M$, then they are for all $x \in M$.*

2.2.2 Horizontal lift of geodesic random walks

Now that we have defined geodesic random walks on the base manifold (M, g) , we can construct a new process on the total manifold $(N, \tilde{g})$ carrying a metric $\tilde{g}$ that will be specified later on. In order to define this process, we recall some terminology from Riemannian geometry.

Definition 2.5. A Riemannian submersion $\pi : (N, \tilde{g}) \rightarrow (M, g)$ is a smooth surjective map whose differential is an isomorphism

$$d\pi_u : (\ker d\pi_u)^\perp \rightarrow T_{\pi(u)}M$$

which is also an isometry. Here $\perp$ denotes the orthogonal complement with respect to the metric $\tilde{g}$ in N .

In the setting of Riemannian submersions, the tangent space $T_u N$ of the total manifold N at a point $u \in N$ splits into the *vertical* and *horizontal subspaces* as follows:

$$\mathcal{V}_u N := \ker d\pi_u, \quad \mathcal{H}_u N := (\mathcal{V}_u N)^\perp, \quad T_u N = \mathcal{H}_u N \oplus \mathcal{V}_u N.$$

Their disjoint unions form two subbundles of TN denoted respectively by $\mathcal{V}N = \sqcup_{u \in N} \mathcal{V}_u N$ and $\mathcal{H}N = \sqcup_{u \in N} \mathcal{H}_u N$. This splits the metric $\tilde{g}$ on TN into its two factors $g_{\mathcal{V}N}$ and $g_{\mathcal{H}N}$. Denote by Γ the space of sections. A horizontal vector field $X \in \Gamma(\mathcal{H}N)$ is π -related to a vector field $\bar{X} \in \Gamma(TM)$ if for any $u \in N$ it holds that

$$d\pi_u(X_u) = \bar{X}_{\pi(u)}. \quad (2.2.5)$$

We stress that relating manifolds via a Riemannian submersion ensures that π -related tangent vectors as in (2.2.5) have the same norm because $d\pi|_{\mathcal{H}N}$ is an isometry.

Remark 2.6. In several situations, the total manifold N comes with a natural projection map $\pi : N \rightarrow M$ defining the vertical subspaces $\mathcal{V}_u N = \ker d\pi_u$ but no specification of a metric $\tilde{g}$. One can then use any connection form ω to define the horizontal subspaces $\mathcal{H}_u N = \ker \omega_u$. Now, with the help of this choice of horizontal bundle, obtained either by the Riemannian submersion or by the specification of a connection form, one can lift any smooth curve on the base manifold to the total manifold with respect to the horizontal bundle.

We can now define the horizontal lift of a curve $\gamma : I \rightarrow M$. We denote $\gamma'(t) = \frac{d}{dt}(\gamma(t))$.

Definition 2.7. The horizontal lift $\tilde{\gamma}$ with respect to $\mathcal{H}N$ starting at $u_0 \in N_{\gamma(0)} = \pi^{-1}(\{\gamma(0)\})$ of a smooth curve $\gamma : I \rightarrow M$ is the unique curve satisfying

$$\pi \circ \tilde{\gamma} = \gamma, \quad \tilde{\gamma}'(t) \in \mathcal{H}_{\tilde{\gamma}(t)} N. \quad (2.2.6)$$

The *horizontal lift* $\tilde{v}(u, v)$ with respect to $\mathcal{H}N$ starting at u of a tangent vector $v \in T_{\pi(u)}M$ is given by

$$\tilde{v} := \tilde{v}(u, v) := (d\pi_u)^{-1}v \in \mathcal{H}_uN,$$

which corresponds to differentiating (2.2.6) for any curve γ such that $\gamma(0) = \pi(u)$ and $\gamma'(0) = v \in T_{\pi(u)}M$. One can prove that this is independent of the choice of the curve, and in particular if $\tilde{v}$ is the horizontal lift of $v \in T_xM$ starting at u , then for every γ in M such that $\gamma(0) = x$, and $\gamma'(0) = v$, $\tilde{v}$ is initially tangent to the horizontal lift of γ starting at u ; $\tilde{v} = \tilde{\gamma}'(0)$.

We recall that the horizontal lift of a geodesic under a Riemannian submersion is again a geodesic (see [79, Lemma 26.11]). It is important to notice that geodesics in N with initial horizontal tangent vector, are horizontal curves; i.e., the tangent vectors remain horizontal. Moreover, by the geodesic property, the tangent vector at any point of the curve is the parallel transport of the initial tangent vector.

Given a distribution of increments $\{\mu_x\}_{x \in M}$ we define the distribution of *horizontally lifted increments* $\{\tilde{\mu}_u\}_{u \in N}$ as follows. First draw v according to $\mu_{\pi(u)}$ and then lift v to $\tilde{v}(u, v)$. Then $\{\tilde{\mu}_u\}_{u \in N}$ is the distribution of the lift $\tilde{v}$. It then follows that the (discrete or continuous-time) random walks based on $\{\mu_x\}_{x \in M}$ are horizontally lifted to the (discrete or continuous-time) random walks based on $\{\tilde{\mu}_u\}_{u \in N}$, and conversely, the projections of random walks based on $\{\tilde{\mu}_u\}_{u \in N}$ are distributed as the random walks based on $\{\mu_x\}_{x \in M}$.

As a consequence, the *horizontal lift of the rescaled continuous-time random walk* $\{Z_t^{(\alpha)}\}_{t \geq 0}$ defined via its generator (2.2.3) is the process $\{\tilde{Z}_t^{(\alpha)}\}_{t \geq 0}$ on the total manifold N with generator defined on smooth compactly supported functions $f : N \rightarrow \mathbb{R}$ by

$$\mathcal{L}_\alpha f(u) := \alpha^{-2} \int_{T_{\pi(u)}M} (f(\exp_u\{\alpha \tilde{v}(v, u)\}) - f(u)) \mu_{\pi(u)}(dv), \quad (2.2.7)$$

so that $\{\pi(\tilde{Z}_t^{(\alpha)})\}_{t \geq 0} = \{Z_t^{(\alpha)}\}_{t \geq 0}$ in distribution.

2.3 Invariance principle for Riemannian submersions

2.3.1 Main result

Let (M, g) and $(N, \tilde{g})$ be Riemannian manifolds with Riemannian submersion $\pi : N \rightarrow M$, and let d be the dimension of M . Each vector field $X \in \Gamma(TN)$ can be uniquely decomposed into its horizontal part $X_{\mathcal{H}} \in \mathcal{H}$ and vertical part $X_{\mathcal{V}} \in \mathcal{V}$ respectively. Under this setting, we consider the Laplace-Beltrami operator Δ_N on N and its decomposition in horizontal and vertical parts as follows (see [7]):

Definition 2.8. The horizontal Laplacian $\Delta_{\mathcal{H}}$ is the generator of the pre-Dirichlet form

$$\mathcal{E}_{\mathcal{H}}(f, h) = - \int_N \langle (\text{grad } f)_{\mathcal{H}}, (\text{grad } h)_{\mathcal{H}} \rangle_{\tilde{g}} d\text{Vol}_{\tilde{g}}, \quad f, h \in C_c^\infty(N).$$

Analogously, the vertical Laplacian $\Delta_{\mathcal{V}}$ is the generator of the pre-Dirichlet form

$$\mathcal{E}_{\mathcal{V}}(f, h) = - \int_N \langle (\text{grad } f)_{\mathcal{V}}, (\text{grad } h)_{\mathcal{V}} \rangle_{\tilde{g}} d\text{Vol}_{\tilde{g}}, \quad f, h \in C_c^\infty(N).$$

We have the decomposition of the total Laplacian into its horizontal and vertical parts

$$\Delta_N = \Delta_{\mathcal{H}} + \Delta_{\mathcal{V}}.$$

In local orthonormal frames $\{E_i\}_{1 \leq i \leq d}$ of $\mathcal{H}$ and $\{F_j\}_{1 \leq j \leq l}$ of $\mathcal{V}$, the horizontal Laplacian can be rewritten as

$$\Delta_{\mathcal{H}} = - \sum_{i=1}^d E_i E_i^* = \sum_{i=1}^d (E_i^2 - (\nabla_{E_i} E_i)_{\mathcal{H}}) - \sum_{j=1}^l (\nabla_{F_j} F_j)_{\mathcal{H}}, \quad (2.3.1)$$

where the adjoint is understood in $L^2(N, d\text{Vol}_{\tilde{g}})$.

The F_j 's are vertical, so they cannot be obtained as horizontal lifts. Thus, horizontally lifted geodesic random walks are not expected to produce the last term in (2.3.1) since their tangent vectors are horizontal. The following type of Riemannian submersion ensures that the last term indeed vanishes; $(\nabla_{F_j} F_j)_{\mathcal{H}} = 0$ (see [84] and [83, Proposition 4.13]).

Definition 2.9. The fibers $N_p := \pi^{-1}(\{p\})$ of a Riemannian submersion $\pi : N \rightarrow M$ are said to be totally geodesic if any geodesic in a fiber, seen as a submanifold of N with the induced metric, is also a geodesic in N .

Assuming that the submersion has totally geodesic fibers, the horizontal Laplacian (2.3.1) takes the form

$$\Delta_{\mathcal{H}} = \sum_{i=1}^d (E_i^2 - (\nabla_{E_i} E_i)_{\mathcal{H}}). \quad (2.3.2)$$

Definition 2.10. A distribution Λ of the tangent bundle TN is said to be bracket-generating if it is generated by a finite number of Lie brackets of vector fields in $\Gamma(\Lambda)$.

Whenever the horizontal subbundle $\mathcal{H}N$ of TN is bracket-generating, the subellipticity of $\Delta_{\mathcal{H}}$ is then guaranteed by Hörmander's theorem. Moreover, [7, Proposition 4.1.5] guarantees in that case its self-adjointness on $C_c^\infty(N)$, and its associated pre-Dirichlet form

has a unique closed extension. On the other hand, as $\mathcal{V}$ is never bracket-generating, we will not consider vertically lifted geodesic random walks.

We are now ready to state the invariance principle for the horizontal lift of the rescaled continuous-time random walk for these types of Riemannian submersions. As a corollary, we obtain the associated relation between the Laplace–Beltrami operator and the horizontal Laplacian.

Theorem 2.11 (Invariance principle for Riemannian submersions). *Let (M, g) and $(N, \tilde{g})$ be geodesically complete Riemannian manifolds equipped with a Riemannian submersion $\pi : N \rightarrow M$ with totally geodesic fibers such that the horizontal subbundle $\mathcal{H}N$ of TN is bracket-generating. Let $\{\mu_x\}_{x \in M}$ be a distribution of increments on M satisfying Assumption 2.2 and Assumption 2.3. Let $\{\tilde{Z}_t^{(\alpha)}\}_{t \geq 0}$ be the process with generator (2.2.7). Then as $\alpha \rightarrow 0$, this process converges to horizontal Brownian motion; i.e., the process with generator $\frac{1}{2}\Delta_{\mathcal{H}}$.*

Corollary 2.12. *Under the setting of Theorem 2.11, the following identity holds:*

$$\Delta_{\mathcal{H}}(f \circ \pi) = (\Delta_M f) \circ \pi, \quad f \in C^\infty(M). \quad (2.3.3)$$

Proof of Corollary 2.12. Consider an α -rescaled continuous-time random walk $\{Z_t^{(\alpha)}\}_{t \geq 0}$ on M that satisfies Assumption 2.2 and Assumption 2.3. By Theorem 2.11 with $\text{Id}_M : M \rightarrow M$ as submersion, this process converges to Brownian motion; i.e., the process with generator $\frac{1}{2}\Delta_M$. On the other hand, by Theorem 2.11, the horizontal lift of this α -rescaled random walk, namely, the process $\{\tilde{Z}_t^{(\alpha)}\}_{t \geq 0}$ on N , converges to horizontal Brownian motion; i.e., the process with generator $\frac{1}{2}\Delta_{\mathcal{H}}$ on N . Because the projection $\pi : N \rightarrow M$ is continuous, and because the random walks $\{Z_t^{(\alpha)}\}_{t \geq 0}$ and $\{\tilde{Z}_t^{(\alpha)}\}_{t \geq 0}$ are by construction π -related, the corresponding limiting Brownian motions must be π -related as well. This proves the identity (2.3.3) for the case of smooth compactly supported functions, and, since (2.3.3) is a pointwise identity, this is enough. $\square$

Proof of Theorem 2.11. Let $f : N \rightarrow \mathbb{R}$ be a smooth compactly supported function. We perform a Taylor expansion of $f \circ \tilde{\gamma}_\alpha$ around $u \in N_p$, where $\tilde{\gamma}_\alpha$ is the horizontal lift starting at u of the curve $\gamma_\alpha(t) = \exp_x(\alpha t v)$, where $v \in T_x M$. There is some $0 < s < 1$ such that

$$f(\tilde{\gamma}_\alpha(1)) - f(u) = \left(\frac{d}{dt} \Big|_{t=0} + \frac{1}{2} \frac{d^2}{dt^2} \Big|_{t=0} + \frac{1}{6} \frac{d^3}{dt^3} \Big|_{t=s} \right) f(\tilde{\gamma}_\alpha(t)).$$

The first time derivative is given by

$$\frac{d}{dt} f(\tilde{\gamma}_\alpha(t)) = df_{\tilde{\gamma}_\alpha(t)} \circ \left(d\pi_{\tilde{\gamma}_\alpha(t)} \right)^{-1} (\gamma'_\alpha(t)) = \langle \text{grad } f, \tilde{\gamma}'_\alpha(t) \rangle_{\tilde{g}} = \langle (\text{grad } f)_{\mathcal{H}}, \tilde{\gamma}'_\alpha(t) \rangle_{\tilde{g}},$$

where we used the fact that $\widetilde{\gamma}_\alpha$ is horizontal for the last equality.

To obtain the second time derivative, we use the Levi-Civita connection ∇ on N and use the fact that $\widetilde{\gamma}_\alpha$ is a geodesic being the horizontal lift of a geodesic under a Riemannian submersion;

$$\begin{aligned} \frac{d^2}{dt^2} f(\widetilde{\gamma}_\alpha(t)) &= \left\langle \nabla_{\widetilde{\gamma}_\alpha'(t)} (\text{grad } f)_\mathcal{H}, \widetilde{\gamma}_\alpha'(t) \right\rangle_{\widetilde{g}} + \left\langle (\text{grad } f)_\mathcal{H}, \nabla_{\widetilde{\gamma}_\alpha'(t)} \widetilde{\gamma}_\alpha'(t) \right\rangle_{\widetilde{g}} \\ &= \left\langle \nabla_{\widetilde{\gamma}_\alpha'(t)} (\text{grad } f)_\mathcal{H}, \widetilde{\gamma}_\alpha'(t) \right\rangle_{\widetilde{g}}. \end{aligned}$$

In particular, at time $t = 0$, consider the orthonormal basis $\{E_i\}_{1 \leq i \leq d}$ of $\mathcal{H}_u N \subset T_u N$ defined as the horizontal lift of an orthonormal basis $\{\overline{E}_i\}_{1 \leq i \leq d}$ of $T_x M$. Write $v = v^i \overline{E}_i \in T_x M$. By linearity of the horizontal lift, we get $\widetilde{\gamma}_\alpha'(0) = \alpha v^i E_i$, and thus

$$\left. \frac{d^2}{dt^2} \right|_{t=0} f(\widetilde{\gamma}_\alpha(t)) = \alpha^2 v^i v^j \langle \nabla_{E_i} (\text{grad } f)_\mathcal{H}, E_j \rangle_{\widetilde{g}} = \alpha^2 v^i v^j (E_i E_j - (\nabla_{E_i} E_j)_\mathcal{H}) f.$$

By Assumption 2.2 on the first and second moments, we deduce that

$$\alpha^{-2} \int_{T_x M} \left. \frac{d}{dt} \right|_{t=0} f(\widetilde{\gamma}_\alpha(t)) \mu_x(dv) = \alpha^{-1} df_u \circ (d\pi_u)^{-1} \left(\int_{T_x M} v \mu_x(dv) \right) = 0,$$

and

$$\begin{aligned} \alpha^{-2} \int_{T_x M} \left. \frac{d^2}{dt^2} \right|_{t=0} f(\widetilde{\gamma}_\alpha(t)) \mu_x(dv) &= \left(\int_{T_x M} v^i v^j \mu_x(dv) \right) (E_i E_j - (\nabla_{E_i} E_j)_\mathcal{H}) f \\ &= \sum_{i=1}^d (E_i^2 - (\nabla_{E_i} E_i)_\mathcal{H}) f. \end{aligned}$$

The last term is the horizontal Laplacian (2.3.2) for a submersion with totally geodesic fibers.

For the third time derivative, first define the horizontal Hessian

$$\text{Hess}_\mathcal{H} f(Y, Z) = \langle \nabla_Y (\text{grad } f)_\mathcal{H}, Z \rangle_{\widetilde{g}}, \quad Y, Z \in \Gamma(TN),$$

which is a symmetric covariant tensor of order 2. Its covariant derivative is thus the tensor given by

$$\nabla \text{Hess}_\mathcal{H} f(X, Y, Z) = X(\text{Hess}_\mathcal{H} f(Y, Z)) - \text{Hess}_\mathcal{H} f(\nabla_X Y, Z) - \text{Hess}_\mathcal{H} f(Y, \nabla_X Z).$$

Note that, again since $\widetilde{\gamma}_\alpha$ is a geodesic,

$$\begin{aligned} \frac{d^3}{dt^3} f(\widetilde{\gamma}_\alpha(t)) &= \nabla_{\widetilde{\gamma}'_\alpha(t)} \left\langle \nabla_{\widetilde{\gamma}'_\alpha(t)} (\text{grad } f)_\mathcal{H}, \widetilde{\gamma}'_\alpha(t) \right\rangle_{\widetilde{g}} \\ &= \nabla_{\widetilde{\gamma}'_\alpha(t)} \left\langle \nabla_{\widetilde{\gamma}'_\alpha(t)} (\text{grad } f)_\mathcal{H}, \widetilde{\gamma}'_\alpha(t) \right\rangle_{\widetilde{g}} - 2 \left\langle \nabla_{\nabla_{\widetilde{\gamma}'_\alpha(t)} \widetilde{\gamma}'_\alpha(t)} (\text{grad } f)_\mathcal{H}, \widetilde{\gamma}'_\alpha(t) \right\rangle_{\widetilde{g}} \\ &= \nabla \text{Hess}_\mathcal{H} f (\widetilde{\gamma}'_\alpha(t), \widetilde{\gamma}'_\alpha(t), \widetilde{\gamma}'_\alpha(t)). \end{aligned}$$

Locally, $(\nabla \text{Hess}_\mathcal{H} f)_u : T_u^3 N \rightarrow \mathbb{R}$ is a bounded operator being linear on a finite dimensional vector space with operator norm given by

$$C(u) = \max_{(\eta, \xi, \zeta) \in T_u^3 N : \|\eta\|_{\widetilde{g}} = \|\xi\|_{\widetilde{g}} = \|\zeta\|_{\widetilde{g}} = 1} |(\nabla \text{Hess}_\mathcal{H} f)_u (\eta, \xi, \zeta)|.$$

This constant $C(u)$ can be uniformly bounded since $|\nabla \text{Hess}_\mathcal{H} f| : N \times T^3 N \rightarrow \mathbb{R}$ is a continuous map on the compact set $\{(q, (\eta, \xi, \zeta)) : q \in \text{Supp } f, (\eta, \xi, \zeta) \in T_q^3 N, \|\eta\|_{\widetilde{g}} = \|\xi\|_{\widetilde{g}} = \|\zeta\|_{\widetilde{g}} = 1\}$, and hence attains a maximum $C > 0$. Since $\|\widetilde{\gamma}'_\alpha(t)\|_{\widetilde{g}} = \|\widetilde{\gamma}'_\alpha(0)\|_{\widetilde{g}} = \alpha \|v\|_g$, we are able to conclude by Assumption 2.3 on the third moment;

$$\begin{aligned} &\alpha^{-2} \int_{T_x M} \frac{d^3}{dt^3} \Big|_{t=s} f(\widetilde{\gamma}_\alpha(t)) \mu_x(dv) \\ &= \alpha^{-2} \int_{T_x M} \nabla \text{Hess}_\mathcal{H} f (\widetilde{\gamma}'_\alpha(s), \widetilde{\gamma}'_\alpha(s), \widetilde{\gamma}'_\alpha(s)) \mu_x(dv) \\ &\leq \alpha C \sup_{y \in \pi(\text{Supp } f)} \int_{T_y M} \|v\|_g^3 \mu_y(dv), \end{aligned}$$

which goes to 0 independently of u as $\alpha \rightarrow 0$. $\square$

Remark 2.13. The transition operator of the discrete-time random walk $\{S_k^{(\alpha)}\}_{k \in \mathbb{N}}$ is related to the generator of the continuous-time process $\{Z_t^{(\alpha)}\}_{t \geq 0}$ via (2.2.3):

$$P^{(\alpha)} = \left(\text{Id} + \alpha^2 L^{(\alpha)} \right).$$

The strong convergence of the generator $L^{(\alpha)}$ towards $\mathcal{L}$ on its core $C_c^\infty(M)$ then implies that $(P^{(\alpha)})^{\lfloor \alpha^{-2} t \rfloor}$ converges to the semigroup $e^{t\mathcal{L}}$ by the Trotter–Kurtz theorem. Analogously, the same holds for the horizontally lifted discrete-time random walk.

Remark 2.14. We obtained Corollary 2.12 as a consequence of the invariance principle, which contrasts with the classical approach. For the sake of completeness, we now briefly outline the latter. Essentially, the proof reduces to showing that the Levi-Civita connection ∇ on N is π -related to the Levi-Civita connection $\overline{\nabla}$ on M (see [84, Lemma 1]). This

follows from the fact that both the inner products for the specific metrics and the Lie brackets preserve π -relations;

$$\langle X, W \rangle_{\bar{g}} = \langle \bar{X}, \bar{W} \rangle_g \circ \pi, \quad d\pi([Y, Z]) = [\bar{Y}, \bar{Z}] \circ \pi$$

for π -related vector fields $W, X, Y, Z \in \Gamma(\mathcal{H}N)$ to $\bar{W}, \bar{X}, \bar{Y}, \bar{Z} \in \Gamma(TM)$, and hence

$$\langle X, [Y, Z] \rangle_{\bar{g}} = \langle \bar{X}, [\bar{Y}, \bar{Z}] \rangle_g \circ \pi.$$

It remains to express the Levi-Civita connection ∇ on N via Coszul's formula for any triple $X, Y, Z \in \Gamma(TN)$;

$$\begin{aligned} 2 \langle \nabla_X Y, Z \rangle_{\bar{g}} &= X \langle Y, Z \rangle_{\bar{g}} + Y \langle X, Z \rangle_{\bar{g}} - Z \langle X, Y \rangle_{\bar{g}} \\ &\quad - \langle X, [Y, Z] \rangle_{\bar{g}} - \langle Y, [X, Z] \rangle_{\bar{g}} + \langle Z, [X, Y] \rangle_{\bar{g}}. \end{aligned}$$

2.3.2 Important examples

Let us go through some important examples from [7, Sections 4.1, 4.4] where the restrictions on the Riemannian submersion, namely, that the fibers are totally geodesic and that the horizontal distribution is bracket-generating, are verified.

1. The manifold (M, g) itself, with $\text{Id}_M : M \rightarrow M$ as submersion. The horizontal distribution is the whole tangent space. Theorem 2.11 gives then a proof of the invariance principle for geodesic random walks on Riemannian manifolds.
2. The tangent bundle $\pi_{TM} : TM \rightarrow M$ equipped with the Sasaki metric [87] defined in terms of coordinates $(\{x_i\}_{1 \leq i \leq d}, \{y_j\}_{1 \leq j \leq d})$ about (p, v) in TM by

$$ds^2 = g_{ij} dx^i dx^j + g_{ij} Dy^i Dy^j,$$

where D denotes the covariant differential with respect to $\bar{\nabla}$ on M ; $Dy^k = dy^k + \bar{\Gamma}_{ij}^k y^i dx^j$.

3. The orthonormal frame bundle $\pi_{O(M)} : O(M) \rightarrow M$ which plays a central role in defining stochastic processes on manifolds by constructing them from their Euclidean counterparts. This motivated our study of horizontal random walks.

Definition 2.15. An orthonormal frame u at x is an ordered choice of orthonormal basis $\{ue_i\}_{1 \leq i \leq d}$ of $T_x M$, where $\{e_i\}_{1 \leq i \leq d}$ is the canonical basis of $\mathbb{R}^d$. The set of all orthonormal frames at x is denoted O_x and their disjoint union $O(M) := \sqcup_{x \in M} O_x$ is referred to as the orthonormal frame bundle.

The orthonormal frame bundle $O(M)$ is a manifold of dimension $d(d+1)/2$ that comes with a natural submersion $\pi_{O(M)} : O(M) \rightarrow M$ sending any orthonormal frame $u \in O_x$ to the basepoint x . If $(U, \{x^i\}_{1 \leq i \leq d})$ is a local chart in M about x , we can express the orthonormal basis of $T_x M$ as $ue_i = e_i^j \frac{\partial}{\partial x^j}$, where $e_i^j = (ue_i)^j$, and this gives a local chart $(\pi^{-1}(U), (\{x^k\}_{1 \leq k \leq d}, \{e_i^j\}_{1 \leq i, j \leq d}))$ in $O(M)$ about u . The coefficients e_i^j satisfy the orthonormality relations $g_{jk} e_i^j e_\ell^k = \delta_{i\ell}$, and local coordinates on the fiber may be chosen from local coordinates on $O(d)$. It remains to define a splitting of $TO(M)$, for instance, by specifying a notion of horizontality.

Definition 2.16. A smooth curve $u : I \rightarrow O(M)$ is horizontal if for any $e \in \mathbb{R}^d$ the tangent vector field $u(t)e \in T_{\pi(u(t))} M$ is itself parallel with respect to the Levi-Civita connection $\bar{\nabla}$ on M along the curve $\pi \circ u : I \rightarrow M$.

This notion of horizontality induces the splitting $TO(M) = \mathcal{H}O(M) \oplus \mathcal{V}O(M)$ and allows us to lift smooth curves horizontally. Given a smooth $\gamma : I \rightarrow M$ and its horizontal lift $\tilde{\gamma}$ starting at u , we recover the parallel transport of tangent vectors along γ given by

$$\tau_{\gamma; t_1 t_2} : T_{\gamma(t_1)} \rightarrow T_{\gamma(t_2)} : v \mapsto \tilde{\gamma}(t_2) (\tilde{\gamma}(t_1))^{-1} v.$$

Note that for each frame u , the collection $\{ue_i\}_{1 \leq i \leq d}$ is a basis of $T_{\pi(u)} M$, and thus one can horizontally lift each of those tangent vectors in order to obtain a new collection of horizontal vectors. In this way, one creates the globally defined canonical horizontal vector fields.

Definition 2.17. Let u be an orthonormal frame at x . The canonical horizontal vector fields

$$H_i(u) := \widetilde{ue_i}, \quad i = 1, \dots, d, \quad (2.3.4)$$

are the horizontal lifts with respect to $\mathcal{H}O(M)$ of the tangent vectors $ue_i \in T_x M$ starting at u .

The global canonical horizontal vector fields $H_i \in \Gamma(\mathcal{H}O(M))$ allow us to obtain a horizontal Laplacian for the orthonormal frame bundle as a sum of squares;

Definition 2.18. The horizontal Laplacian of $O(M)$ is given by

$$\Delta_{\mathcal{H}} := \sum_{i=1}^d H_i^2.$$

Remark 2.19. Note that $\sum_{i=1}^d H_i^2$ is in Hörmander's form, and differs from (2.3.2) in the general case of submersions with totally geodesic fibers; this is due to the fact that $O(M)$ is a parallelizable manifold. While Nash's embedding theorem allows

one to write the Laplace–Beltrami operator of M as a sum of squares of orthogonal projections (see for instance [57, Theorem 3.1.4]), this comes at cost of extra terms coming from the dimension of the isometric embedding.

This horizontal Laplacian and the Laplace–Beltrami operator satisfy the following relation, and this is a starting point in stochastic calculus on manifolds based on the rolling-without-slipping construction of Brownian motion as described in Hsu’s monograph [57]:

Proposition 2.20. *The following identity holds:*

$$\Delta_{\mathcal{H}}(f \circ \pi_{O(M)}) = \Delta_M f \circ \pi_{O(M)}, \quad f \in C^\infty(M). \quad (2.3.5)$$

One can show that Definition 2.18 coincides with (2.3.1) once a specific metric is prescribed, namely, the Sasaki–Mok metric (specified below in more detail). This means the invariance principle for horizontal geodesic random walks on $O(M)$ are a consequence of Theorem 2.11. However, it turns out that for $O(M)$ a more direct proof of this invariance principle is available without relying on this metric and the fact that the fibers are totally geodesic. We will detail this proof below. In particular, Proposition 2.20 is a corollary of that invariance principle just as before. Given $f \in C^\infty(O(M))$ and $\widetilde{\gamma}_\alpha(t)$ the horizontal lift starting at the frame $u \in \pi_{O(M)}^{-1}(\{p\})$ of the geodesic γ with $\gamma'(0) = v \in T_x M$, then by definition of horizontal lift

$$\gamma'_\alpha(t) = \tau_{\gamma_\alpha; 0t}(\alpha v) = \widetilde{\gamma}_\alpha(t)(\widetilde{\gamma}_\alpha(0))^{-1}(\alpha v) = \alpha \widetilde{\gamma}_\alpha(t) u^{-1} v = \alpha (u^{-1} v)^i \widetilde{\gamma}_\alpha(t) e_i.$$

Since u is an orthonormal frame, $\{ue_i\}_{1 \leq i \leq d}$ is an orthonormal basis of $T_x M$, and we get

$$v^i = \langle v, ue_i \rangle_g = \langle u^{-1} v, e_i \rangle_{\mathbb{R}^d} = (u^{-1} v)^i.$$

Therefore,

$$\frac{d}{dt} f(\widetilde{\gamma}_\alpha(t)) = df_{\widetilde{\gamma}_\alpha(t)} \circ \widetilde{\gamma}_\alpha'(t) = \alpha v^i H_i f(\widetilde{\gamma}_\alpha(t)),$$

and likewise,

$$\begin{aligned} \frac{d^2}{dt^2} f(\widetilde{\gamma}_\alpha(t)) &= \alpha^2 v^j v^i H_j H_i f(\widetilde{\gamma}_\alpha(t)), \\ \frac{d^3}{dt^3} f(\widetilde{\gamma}_\alpha(t)) &= \alpha^3 v^k v^j v^i H_k H_j H_i f(\widetilde{\gamma}_\alpha(t)). \end{aligned}$$

By the Assumption 2.2, the first moment vanishes, the second moment is given by

$$\begin{aligned} \alpha^{-2} \int_{T_x M} \frac{d^2}{dt^2} \Big|_{t=0} f(\widetilde{\gamma}_\alpha(t)) \mu_x(dv) &= \left(\int_{T_x M} v^j v^i \mu_x(dv) \right) H_j H_i f(u) \\ &= \sum_{i=1}^d H_i^2 f(u), \end{aligned}$$

and for the third moment we argue as follows: $\pi_{O(M)}(\text{Supp } f)$ is compact by continuity of the projection, and thanks to Assumption 2.3 we estimate

$$\begin{aligned} \alpha^{-2} \int_{T_x M} \frac{d^3}{dt^3} \Big|_{t=s} f(\widetilde{\gamma}_\alpha(t)) \mu_x(dv) &= \alpha \int_{T_x M} v^k v^j v^i H_k H_j H_i f(\widetilde{\gamma}_\alpha(s)) \mu_x(dv) \\ &\leq \alpha \sup_{y \in \pi_{O(M)}(\text{Supp } f)} \int_{T_q M} \|v\|^3 \mu_y(dv) \sum_{i,j,k=1}^d \|H_k H_j H_i f\|_\infty, \end{aligned}$$

which goes to 0 independently of the frame u as $\alpha \rightarrow 0$.

For the sake of completeness, we now describe this metric which was introduced by Mok [80] (see also Sasaki [87] and [68]). One starts by finding a coordinate expression for the canonical horizontal vector fields. Consider a horizontal lift $\tilde{\gamma} : I \rightarrow O(M)$ that starts at u with $\gamma'(0) = ue_i$. By definition of horizontal lift with respect to $\mathcal{H}O(M)$,

$$H_i(u) = \widetilde{ue_i} = \tilde{\gamma}'(0) = \dot{\gamma}^j(0) \frac{\partial}{\partial x^j} + ((\tilde{\gamma}e_l)^m)'(0) \frac{\partial}{\partial e_l^m},$$

and since the tangent vectors $\tilde{\gamma}(t)e_l$ are parallel with respect to $\overline{\nabla}$ on M along the curve $\pi \circ \tilde{\gamma} : I \rightarrow M$ whose initial tangent vector is ue_i , the geodesic equation yields

$$((\tilde{\gamma}e_l)^m)'(0) = -e_i^j e_l^k \overline{\Gamma}_{jk}^m, \quad 1 \leq l < m \leq d,$$

where $\{\overline{\Gamma}_{ij}^k\}_{1 \leq i,j,k \leq d}$ denote the Christoffel symbols of $\overline{\nabla}$. The horizontal and vertical subbundles of $TO(M)$ are thus respectively spanned by (see [57, Proposition 2.1.3])

$$H_i(u) = e_i^j \left(\frac{\partial}{\partial x^j} - e_l^k \overline{\Gamma}_{jk}^m \frac{\partial}{\partial e_l^m} \right), \quad i = 1, \dots, d, \quad (2.3.6)$$

and

$$V_j^k(u) := e_j^l \frac{\partial}{\partial e_k^l}, \quad 1 \leq j < k \leq d. \quad (2.3.7)$$

Definition 2.21. *The canonical 1-form θ and the connection form ω on $O(M)$ associated to $\overline{\nabla}$ on M are the dual forms to the vector fields (2.3.6) and (2.3.7) given by*

$$\theta^k(u) := e_l^k dx^l, \quad \omega_i^j(u) := e_k^j \left(\overline{\Gamma}_{lm}^k e_i^l dx^m + de_i^k \right).$$

The Sasaki–Mok metric $\tilde{g}$ is defined pointwise by

$$\langle \eta, \xi \rangle_{\tilde{g}} := \langle d\pi_{O(M),u}(\eta), d\pi_{O(M),u}(\xi) \rangle_g + \langle \omega_u(\eta), \omega_u(\xi) \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes an $O(d)$ -invariant inner product on $\mathfrak{o}(d)$.

With this metric and connection form, one can prove that $\pi_{O(M)} : O(M) \rightarrow M$ is a totally geodesic submersion. As consequence, the invariance principle for lifted geodesic random walk with limiting generator $\frac{1}{2}\Delta_{\mathcal{H}}$ follows from Theorem 2.11 and therefore also the equality between Definition 2.18 and (2.3.1).

4. A general class of spaces on which such invariance principle holds are the principal bundles $\pi_P : P \rightarrow M$ with fiber Lie group G . Given a G -compatible connection form ω and a G -invariant metric b on G , there is a unique metric $\tilde{g} = \pi^*g + b\omega$ on x that makes π_P into a Riemannian submersion with totally geodesic fibers such that the horizontal distribution $\mathcal{H}$ of ω is the orthogonal complement of the vertical distribution [106, Theorem 3.5]. Whenever the horizontal distribution $\mathcal{H}$ is bracket-generating, the subellipticity of $\Delta_{\mathcal{H}}$ is guaranteed and there is a unique closed extension of its associated pre-Dirichlet form. The previous examples fall under this category.

CHAPTER 3

Hydrodynamic Limit of the Symmetric Exclusion Process on Manifolds

This chapter is the content of the joint work [63] with Frank Redig and Rik Versendaal "Hydrodynamic limit of the symmetric exclusion process on complete Riemannian manifolds and principal bundles." arXiv preprint arXiv:2410.20167 (2024).

3.1 Introduction and main results

3.1.1 Introduction

Interacting particle systems serve as microscopic models from which macroscopic non-equilibrium behavior can be derived. An example is the hydrodynamic limit where one derives partial differential equations for macroscopic quantities such as the particle density by appropriate scaling of space and time in the microscopic particle dynamics, see for e.g. the monographs of Kipnis–Landim [67], DeMasi–Presutti [26], and Spohn [94]. A well-known and thoroughly studied example of an interacting system is the symmetric exclusion process (SEP) introduced by Spitzer [93] (see also Liggett [74], Chapter 8), where particles perform random walks with the restriction of having at most one particle per site. For this model in the standard setting of the d -dimensional lattice with nearest neighbour jumps, the hydrodynamic limit is the heat equation which can be derived after diffusive rescaling of space and time.

At present, hydrodynamic limits have been considered mostly on the lattice $\mathbb{Z}^d$ or on a discrete torus, and the corresponding macroscopic space is then either $\mathbb{R}^d$ or the d -dimensional torus. Let us mention a few cases where the hydrodynamic limit of the SEP was obtained beyond the Euclidean setting: Faggionato [37] considered symmetric exclusion processes in random environments, including Delaunay's triangulations starting

from a homogeneous Poisson point process for which the limiting heat equation has a homogenized diffusion matrix. Chen–Gonçalves [20, 21] studied hydrodynamic limits for particle systems on graphs approximating a fractal such as the Sierpinski gasket.

In [99], van Ginkel–Redig started investigating the setting where the macroscopic space is a compact Riemannian manifold, with the aim of deriving macroscopic equations from interacting particle systems on discrete graphs approximating the manifold. More precisely, for a general class of neighborhood graphs that approximate a compact Riemannian manifold (based on uniformly chosen points according to the normalized volume measure), in [99] the corresponding SEP is studied and a proof of the hydrodynamic limit (yielding the heat equation) is given.

In this chapter, we extend the realm of these results in several ways. First, we drop the assumption of the compactness of the manifold. Second, we consider random neighborhood graphs whose vertices are given by a Poisson point process with Gibbs intensity. Third, whereas [99] relied on the Wasserstein convergence of empirical distribution of i.i.d. points sampled from the normalized volume measure, here we rely on a simpler argument based on Bernstein’s inequality to prove convergence of the rescaled random walks on the random neighborhood graphs obtained from the Poisson point processes towards a diffusion process. Fourth, we lift the SEP to principal bundles, thus obtaining a simple sub-Riemannian setting where convergence to the heat equation corresponding to the horizontal Laplacian is obtained. As a result, starting from the corresponding SEP, we obtain several new variants of the heat equation, including the heat equation with weighted Laplacian $e^U \operatorname{div}(e^{-U} \nabla)$, where U is the potential of the Gibbs measure, replacing the Laplace–Beltrami operator. In the principal bundle setting, we obtain a horizontal heat equation with weighted horizontal Laplacian which can be considered as a first result in a sub-Riemannian setting. These results extend significantly the hydrodynamic equations that one can obtain from the SEP, both in extending the geometric setting, but even in the Euclidean setting, where the heat equation with a weighted Laplacian is new.

The SEP is special because it manifests a microscopic closure of the expected density field; i.e., one does not need replacement lemmas to derive a macroscopic equation. This property, which can be viewed as a manifestation of self-duality, is crucial in the derivation of the (expected) macroscopic equation, which reduces to the proof of the scaling limit of a single random walk (RW) to a diffusion process; the Kolmogorov forward equation of this diffusion process is then the hydrodynamic limit. Therefore, an important aspect is to consider suitable symmetric transition rates for RW on neighbourhood graphs in order to approximate a diffusion process with (horizontal) weighted Laplacian as generator. These transition rates are inspired from Hein–Audibert–von Luxburg [8] and Gao [41]; in the context of horizontal diffusions, this requires a double scaling in the base space and the fibers. To prove weak convergence in path space of the trajectory of the empirical density, one needs to additionally control the quadratic variation of the corresponding Dynkin martingale.

The rest of the chapter is organized as follows : In [Section 3.1.2](#) we introduce the SEP in

our geometric setting and state the main results. In [Section 3.2](#) we establish existence of the SEP in our setting. For the case of non-compact manifolds, the approximating neighborhood graph is infinite, and the classical conditions of existence of SEP (from [74]) might not be satisfied. However, via the graphical construction the proof of existence can be reduced to the proof of existence (i.e., non-explosion) of the random walk almost surely in the realization of the Poisson point process. In [Section 3.3](#) we prove an abstract hydrodynamic limit for the SEP on complete manifolds which encompasses our setting. In [Section 3.4](#) we prove the almost sure convergence of a single RW on a sequence of random neighborhood graphs drawn from a PPP with Gibbs intensity. In [Section 3.5](#) we extend the setting to principal bundles by a horizontal lift procedure and obtain the corresponding hydrodynamic limit.

3.1.2 Main results

We consider a geodesically complete connected m -dimensional Riemannian manifold (M, g) equipped with a Gibbs reference measure

$$d\mu := e^{-U} d\text{Vol}_M. \quad (3.1.1)$$

Here Vol_M denotes the Riemannian volume measure. We make the following assumptions on the manifold and the Gibbs measure:

- (A1) The manifold has Ricci curvature lower bound; $\text{Ric}_M \geq (m - 1)\kappa$ for some $\kappa \in \mathbb{R}$.
- (A2) The potential $U \in C^\infty(M)$ is such that $e^{-U} \in L^\infty(M)$.

In the next definition we introduce a sequence of random neighborhood graphs $(G_N)_{N \geq 1}$ which approximate the manifold with corresponding measure $d\mu$. We denote by d_M the Riemannian distance on M .

Definition 3.1 (Random neighborhood graph (RNG)). *Let Λ^N be a Poisson point process PPP of intensity $Nd\mu$. The vertices V_N consist of the points X^i of Λ^N , and an undirected edge is present between two vertices whenever they are close; i.e., $d_M(X^i, X^j) \leq h_N$ for some bandwidth parameter h_N . The sequence of random neighborhood graphs (RNG) is given by $(G_N := (V_N, E_N))_{N \geq 1}$, where E_N denotes the set of undirected edges.*

The symmetric exclusion process we consider evolves on those random neighborhood graphs. Each site can accommodate at most one particle; this is the exclusion rule. Particles hop with symmetric edge weights over edges; jumps leading to more than one particle per site are suppressed.

Definition 3.2 (Symmetric exclusion process (SEP)). *Given a graph $G = (V, E)$ with symmetric weights W , we denote the configuration of particles by $\eta \in \Omega_V = \{0, 1\}^V$ and*

the exchange of occupancies between the vertices x and y by

$$\eta^{x,y}(z) := \begin{cases} \eta(z) & \text{if } z \neq x, y, \\ \eta(y) & \text{if } z = x, \\ \eta(x) & \text{if } z = y. \end{cases}$$

The symmetric exclusion process (SEP) is the continuous-time Markov jump process on the configuration space Ω_V whose generator evaluated on functions depending only on a finite number of occupancies, namely, local functions $f : \Omega_V \rightarrow \mathbb{R}$, is given by

$$L_G^{(\text{SEP})} f(\eta) := \sum_{(x,y) \in E} W(x,y) (f(\eta^{x,y}) - f(\eta)).$$

In our setting the choice of symmetric weights take into account the geometry of the manifold M , and the choice of the potential U . More precisely, let $k : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a bounded measurable with compact support in $[0, 1)$, continuous near 0, and with $k(0) > 0$, which makes the moments $C_{k,0}$ and $C_{k,2}$ defined in Lemma 3.28 positive. We refer to k as a *kernel*. We consider the following *symmetric edge weights* indexed by a fixed parameter $\alpha \geq 0$,

$$W_{N;\alpha}(X^i, X^j) := \frac{1}{N} \frac{h_N^{-m} k\left(\frac{d_M(X^i, X^j)}{h_N}\right)}{\left(\widehat{k}_N(X^i) \widehat{k}_N(X^j)\right)^\alpha}, \quad \widehat{k}_N(x) := \frac{1}{N} \sum_{z \in \Lambda^N} h_N^{-m} k\left(\frac{d_M(x, z)}{h_N}\right). \quad (3.1.2)$$

The choice of the edge weights is motivated by the fact that the corresponding diffusively rescaled random walk with jumps over the edge (xy) with rate $h_N^{-2} W_{N;\alpha}(x, y)$ converges to a diffusion process whose generator is the α -weighted Laplacian

$$\mathcal{L}_\alpha = C_{k;\alpha} e^{(2\alpha-1)U} \Delta_{2(1-\alpha)}, \quad C_k := \frac{C_{k,2}}{2(C_{k,0})^{2\alpha}}. \quad (3.1.3)$$

When $\alpha = 1/2$, the generator is a multiple of the Witten Laplacian $\Delta - \nabla U \cdot \nabla$. Because the edge weights are symmetric, by self-duality of the SEP, the hydrodynamic limit of the associated SEP reduces to the scaling limit of a single random walk (see Section 3.3 for details).

Definition 3.3 (Weighted Laplacian). *Given a smooth function $\phi \in C^\infty(M)$, the α -weighted Laplacian is the reversible operator with respect to the α -weighted Gibbs measure $d\mu_\alpha := e^{-\alpha U} d\text{Vol}_M$; that is, the operator satisfying*

$$\forall \varphi \in C_c^\infty(M), \quad \int_M \varphi \Delta_\alpha \phi e^{-\alpha U} d\text{Vol}_M = - \int_M \nabla \varphi \cdot \nabla \phi e^{-\alpha U} d\text{Vol}_M,$$

which is given by

$$\Delta_\alpha := e^{\alpha U} \nabla \cdot (e^{-\alpha U} \nabla) = \Delta - \alpha \nabla U \cdot \nabla.$$

We use the following additional hypothesis which ensures that the hydrodynamic limit PDE is well-posed:

- (A3) The closure of $\mathcal{L}_\alpha$ on $C_c^\infty(M)$ generates a conservative Markov semigroup and $C_c^\infty(M)$ is a core for its generator.

For $\alpha = 1/2$, the conservativeness part of (A3) is precisely stochastic completeness for the weighted diffusion $\Delta - \nabla U \cdot \nabla$. In that case, it is sufficient that $\text{Hess } U \geq \beta g$ for some $\beta \in \mathbb{R}$ so that, combined with (A1), the Bakry–Émery tensor $\text{Ric} + \text{Hess } U$ is bounded from below, which implies stochastic completeness (see [47]). In general, a sufficient condition for stochastic completeness is Grigor'yan's criterion

$$\int_{r_0}^{\infty} \frac{r}{\log \mu[B_{g_\alpha}(o, r)]} = +\infty$$

for some $o \in M$ and $r_0 > 0$, where $g_\alpha = e^{-(2\alpha-1)U}g$ is the intrinsic metric for which $\mathcal{L}_\alpha$ is essentially self-adjoint on $C_c^\infty(M)$ in $L^2(M, \mu)$; B_{g_α} denotes the Riemannian ball with respect to the metric g_α . Regarding the hypothesis on the core, for the case $\alpha = 1/2$, this follows from geodesic completeness of (M, g) by [47, Theorem 2.2], which is assumed in our case, while for general $\alpha \geq 0$, it is sufficient that (M, g_α) is geodesically complete by [89, Theorem 1.1]. We emphasize that geodesic completeness of (M, g) does not imply geodesic completeness of (M, g_α) in general unless the potential U is bounded.

In order to formulate the result of the hydrodynamic limit, we introduce the empirical measure associated to a configuration of SEP. More precisely, for a configuration η^N of the SEP on the weighted graphs G^N we define the macroscopic quantity

$$\widehat{\pi}^N = \frac{1}{N} \sum_{X^i \in V_N} \eta^N(X^i) \delta_{X^i}. \quad (3.1.4)$$

The hydrodynamic limit can then be described in words as follows: if the empirical measure at time 0 converges in an appropriate sense to $\rho_0 d\mu$, then after diffusive rescaling of time by h_N^{-2} , at time t , the empirical measure converges to $\rho_t d\mu$ where ρ_t satisfies (3.1.5) below. We give below the precise formulation of this result on the hydrodynamic limit. We denote by $C_0(M)$ the space of continuous functions vanishing at infinity.

Theorem 3.4 (Quenched hydrodynamic limit of the SEP). *Let $(G_N = (V_N, E_N))_{N \geq 1}$ be a sequence of RNG on M in the sense of Definition 3.1 with symmetric weights $(W_{N;\alpha})_{N \geq 1}$ given by (3.1.2) for a fixed $\alpha \geq 0$. Assume that (A1), (A2), and (A3) hold. Let η_0^N be the initial configuration of the SEP on G_N with associated initial empirical measure $\widehat{\pi}_0^N$ that converges vaguely to $\rho_0 d\mu$ in probability, where $0 \leq \rho_0 \leq 1$; i.e.,*

$$\forall \varphi \in C_0(M), \quad \left\langle \varphi, d\widehat{\pi}_0^N \right\rangle \xrightarrow{\text{pr.}} \langle \varphi, \rho_0 d\mu \rangle \quad \text{as } N \rightarrow +\infty,$$

where

$$\widehat{\pi}_0^N := \frac{1}{N} \sum_{X^i \in V_N} \eta_0^N(X^i) \delta_{X^i}.$$

Suppose that the bandwidth parameters $(h_N)_{N \geq 1}$ are chosen in such a way that both $h_N \rightarrow 0$ and $Nh_N^{m+2}/\log N \rightarrow +\infty$ as $N \rightarrow +\infty$. Then for almost all realizations of the sequence of graphs, the trajectory of the empirical measure $t \mapsto \widehat{\pi}_t^N$ associated to the configuration η_t^N that evolves according to the diffusively rescaled SEP with generator $h_N^{-2} L_{G_N}^{(\text{SEP})}$ converges in probability in the Skorokhod topology to a deterministic trajectory $t \mapsto \rho_t d\mu$ whose density ρ_t weakly solves

$$\partial_t \rho_t = C_{k,\alpha} e^{(2\alpha-1)U} \Delta_{2(1-\alpha)} \rho_t. \quad (3.1.5)$$

For $\alpha = 1/2$, this reduces to $\partial_t \rho_t = C_{k;1/2} (\Delta - \nabla U \cdot \nabla) \rho_t$.

[Theorem 3.4](#) is a direct consequence of an abstract hydrodynamic limit ([Theorem 3.14](#)) and an almost sure pointwise consistency of a single random walk on the RNG ([Theorem 3.16](#)) that we establish in [Section 3.3](#) and [Section 3.4](#) respectively. The existence of the SEP is the content of [Section 3.2](#). We work throughout with a fixed parameter $\alpha \geq 0$.

Remark 3.5.

1. In [Theorem 3.4](#), we prepare the initial configuration to approximate $\rho_0 d\mu$ and consider the evolution of the trajectory $t \mapsto \rho_t d\mu$ for a density ρ_t with respect to the reference Gibbs measure $d\mu$. Since $\mathcal{L}_\alpha = C_{k;\alpha} e^U \operatorname{div}(e^{-2(1-\alpha)U} \nabla)$ is symmetric in $L^2(M, d\mu)$, [\(3.1.5\)](#) is the forward Kolmogorov equation for densities with respect to $d\mu$.
2. In the special case $\alpha = 1/2$, if instead, one prepares the initial configuration such that it approximates $\rho_0 d\operatorname{Vol}_M$, then it is natural to look for the evolution of the density $\rho_t d\operatorname{Vol}_M$; that is, with respect to the volume measure. Because the $L^2(M, d\operatorname{Vol}_M)$ -adjoint of $\Delta - \nabla U \cdot \nabla$ is $\Delta + \operatorname{div}(\cdot \nabla U)$, in that case, the hydrodynamic limit is given by the Fokker–Planck equation

$$\partial_t \rho_t = C_{k;1.2} \nabla \cdot (\nabla \rho_t + \rho_t \nabla U). \quad (3.1.6)$$

3.2 Existence of the SEP

In a non-compact setting, the existence of the SEP is not trivial, because the weighted graphs G_N with weights $W_{N;\alpha}$ given in [\(3.1.2\)](#) are infinite graphs as soon as the measure $d\mu$ in [\(3.1.1\)](#) is an infinite measure (i.e., when $\mu(M) = \infty$). For every fixed N and $\alpha \geq 0$, the statistical estimator $\widehat{k}_N$ of e^{-U} (see [Lemma 3.29](#)) contains the self-point and therefore

$$\widehat{k}_N(x) \geq \frac{k(0)}{Nh_N^m} > 0, \quad x \in \Lambda^N. \quad (3.2.1)$$

Consequently, the individual edge weights are uniformly bounded,

$$\|W_{N;\alpha}\|_\infty \leq \frac{\|k\|_\infty}{Nh_N^m} \left(\frac{Nh_N^m}{k(0)} \right)^{2\alpha} < \infty. \quad (3.2.2)$$

This motivates the use of $\widehat{k}_N$ instead of e^{-U} for (3.1.2) since otherwise the individual edge weights might not be uniformly bounded unless U is. Note that this does not give a uniform bound on the total jump rate, since the degrees of an infinite Poisson neighborhood graph is not uniformly bounded. Hence, the existence criterion [74, Proposition 6.1] for the process need not be satisfied.

Nevertheless, via the graphical construction of SEP, it is enough to prove that a single random walk on the graph has infinite existence time for every starting point. To this end, we rely on a general criterion of graph stochastic completeness.

Theorem 3.6 (Folz [38]). *Let (G, W) be a countable, locally finite connected weighted graph and d_G a graph metric such there exist constants $C, C' > 0$ satisfying*

$$d_G(x, y) \leq C \text{ if } x \sim y, \quad \sum_{y \sim x} W(x, y) d_G(x, y)^2 \leq C' \text{ for all } x \in G. \quad (3.2.3)$$

Then (G, W) is stochastically complete if

$$\int_{r_o}^{\infty} \frac{r}{\log \#B_G(o, r)} dr = +\infty \quad (3.2.4)$$

for some $o \in G$ and $r_o > 0$ such that $\#B_G(o, r_o) > 1$, where $B_G(o, r)$ denotes the graph ball of radius r centered at o .

The Ricci curvature lower bound $\text{Ric}_M \geq (m-1)\kappa$ we assume allows us to compare the volume of large balls with the ones of the model spaces. Ultimately, thanks to this comparison, we will be able to upper bound the denominator of (3.2.4) in a way that the integral diverges, and thus prove the existence of the process. For the comparison of the volume of large balls with the ones of the model spaces, we use the Bishop–Gromov inequality stated below.

Theorem 3.7 (Bishop–Gromov inequality [72, Theorem 11.19]). *Assume $\text{Ric}_M \geq (m-1)\kappa$ for some $\kappa \in \mathbb{R}$. Then, for every $x \in M$,*

$$\text{Vol}_M(B_M(x, r)) \leq \text{Vol}_\kappa(r), \quad (3.2.5)$$

where $\text{Vol}_\kappa(r)$ denotes the volume of a ball of radius r in the simply connected space of constant sectional curvature κ .

Remark 3.8. *Under the Ricci lower bound $\text{Ric}_M \geq (m-1)\kappa$, if $\kappa > 0$, then the manifold is compact by Myer's theorem [72, Theorem 12.24]. If $\kappa = 0$, then its volume growth is at most like r^m . Finally, if $\kappa < 0$, then its volume growth is at most like $e^{(m-1)\sqrt{-\kappa}r}$.*

We also need a basic large deviation estimate for the Poisson distribution stated below.

Proposition 3.9 (Bennett's inequality [14, Theorem 2.9]). *One has*

$$\mathbb{P}[\text{Poisson}(\lambda) \geq \lambda(1+s)] \leq \exp(-\lambda\Phi(s)) \quad (3.2.6)$$

for $s \geq 0$ and

$$\mathbb{P}[\text{Poisson}(\lambda) \leq \lambda(1+s)] \leq \exp(-\lambda\Phi(s))$$

for $-1 \leq s \leq 0$, where $\Phi(s) := (1+s) \log(1+s) - s$.

Theorem 3.10 (Quenched existence of the SEP). *Let $(G_N = (V_N, E_N))_{N \geq 1}$ be a sequence of RNG on M in the sense of Definition 3.1 with symmetric weights $(W_{N;\alpha})_{N \geq 1}$ given by (3.1.2), for a fixed $\alpha \geq 0$. Assume (A1) and (A2). Then for almost every realization of the sequence of graphs, the SEP exists for all time.*

Proof. Step 1 (Choice of an adapted graph metric). The existence is trivial if M is compact. Thus we assume $\text{Ric}_M \geq (m-1)\kappa$ for some $\kappa \leq 0$ in the non-compact case; recall that $\kappa > 0$ implies that M is compact. Set the length of positive-rate edges as

$$l_{G_N}(x, y) := \min \left\{ 1, \frac{1}{\sqrt{\lambda_N(x) \vee \lambda_N(y)}} \right\}, \quad \lambda_N(x) := \sum_{y \sim x} W_{N;\alpha}(x, y),$$

and let d_{G_N} be the associated shortest-path metric; that is,

$$d_{G_N}(x, y) := \inf \left\{ \sum_{i=1}^{\ell} l_{G_N}(X^i, X^{i+1}) : X^1 = x, X^{\ell+1} = y \text{ and } X^i \sim X^{i+1} \right\}.$$

Then we can pick $C = C' = 1$ in (3.2.3) for any connected component of the positive-rate graph since $d_{G_N}(x, y) \leq 1$ for $x \sim y$ and

$$\sum_{y \sim x} W_{N;\alpha}(x, y) d_{G_N}(x, y)^2 \leq \sum_{y \sim x} W_{N;\alpha}(x, y) \frac{1}{\lambda_N(x)} = 1$$

whenever x is not isolated; if x is isolated then there is nothing to prove.

Step 2 (Comparison of the graph and Riemannian balls). We claim that for all $x \in \Lambda^N$,

$$B_{G_N}(x, r) \subseteq \Lambda^N \cap B_M(x, C_1 r^2) \quad \text{for all } r \text{ large enough a.s.} \quad (3.2.7)$$

with some constant $C_1 > 0$ specified below. Let $y \in B_{G_N}(x, r) \setminus \{x\}$. Then there exists a path γ from x to y of the form

$$x =: X^1 \sim \dots \sim X^{\ell+1} := y$$

for some $\ell > 0$ such that

$$\sum_{i=1}^{\ell} l_{G_N}(X^i, X^{i+1}) < r. \quad (3.2.8)$$

Moreover, $X^i \in \overline{B_M}(x, R_\gamma)$ for all $1 \leq i \leq \ell + 1$, where $R_\gamma := \max\{d_M(x, X^i)\}_{i=1}^{\ell+1}$. In particular,

$$y \in \Lambda^N \cap \overline{B_M}(x, R_\gamma). \quad (3.2.9)$$

Set

$$\lambda'_N(R) := \sup_{z \in \Lambda^N \cap \overline{B_M}(x, R)} \lambda_N(z). \quad (3.2.10)$$

Since every edge has Riemannian length at most h_N , any path reaching a vertex at distance R_γ contains at least R_γ/h_N edges. All its vertices lie in $\overline{B_M}(x, R_\gamma)$, and therefore (3.2.8) gives

$$\frac{R_\gamma/h_N}{\sqrt{1 \vee \lambda'_N(R_\gamma)}} < r. \quad (3.2.11)$$

We still have to obtain an almost sure linear upper bound on $\lambda'_N(R)$, which will imply $R_\gamma \lesssim r^2$ and hence prove (3.2.7).

Fix the sampled root x and work first under its Palm law $\mathbb{P}^x$; under $\mathbb{P}^x$ the configuration is represented by an independent PPP of intensity $N d\mu$ together with the fixed point x . For $R, K > 0$, define the number of bad samples (with too much neighbors)

$$N_{\text{bad}}(R; K) := \# \{z \in \Lambda^N \cap \overline{B_M}(x, R) : \#(\Lambda^N \cap \overline{B_M}(z, h_N)) \geq K\}.$$

Note that

$$\lambda_N(z) \leq \|W_{N,\alpha}\|_\infty \#(\Lambda^N \cap \overline{B_M}(z, h_N)), \quad z \in \Lambda^N. \quad (3.2.12)$$

Hence, courtesy of (3.2.10), (3.2.12), and Markov's inequality, we get

$$\mathbb{P}^x[\lambda'_N(R) \geq \|W_{N,\alpha}\|_\infty K] \leq \mathbb{P}^x[N_{\text{bad}}(R; K) \geq 1] \leq \mathbb{E}^x[N_{\text{bad}}(R; K)]. \quad (3.2.13)$$

The multivariate Mecke formula [70, Theorem 4.4] gives

$$\begin{aligned} \mathbb{E}^x[N_{\text{bad}}(R; K)] &= \mathbb{P}^x[\#(\Lambda^N \cap \overline{B_M}(x, h_N)) \geq K] \\ &\quad + N \int_{\overline{B_M}(x, R)} \mathbb{P}^{x,z}[\#(\Lambda^N \cap \overline{B_M}(z, h_N)) \geq K] e^{-U(z)} d\text{Vol}_M(z), \end{aligned} \quad (3.2.14)$$

where $\mathbb{P}^{x,z}$ denotes the two-fold Palm law. Under $\mathbb{P}^{x,z}$, the remaining configuration is again an independent PPP of intensity $N d\mu$, while the two points x, z are fixed. By Bishop–Gromov inequality (3.2.5),

$$N\mu(\overline{B}_M(z, h_N)) = N \int_{\overline{B}_M(z, h_N)} e^{-U(y)} d\text{Vol}_M(y) \leq N \|e^{-U}\|_\infty \text{Vol}_\kappa(h_N) =: C_2(h_N). \quad (3.2.15)$$

Consequently, for $K > C_2(h_N) + 2$, Bennett's inequality (3.2.6) yields

$$\begin{aligned} \sup_{x,z} \mathbb{P}^{x,z} [\#(\Lambda^N \cap \overline{B}_M(z, h_N)) \geq K] &\leq \mathbb{P}[\text{Poisson}(C_2(h_N)) \geq K - 2] \\ &\leq \exp \left[-C_2(h_N) \Phi \left(\frac{K - 2}{C_2(h_N)} - 1 \right) \right]. \end{aligned} \quad (3.2.16)$$

The same upper bound applies to $\mathbb{P}^x[\#(\Lambda^N \cap \overline{B}_M(x, h_N)) \geq K]$. Since

$$C_2(h_N) \Phi \left(\frac{K - 2}{C_2(h_N)} - 1 \right) = (K - 2) \log \left(\frac{K - 2}{C_2(h_N)} \right) - (K - 2) + C_2(h_N),$$

there exist $K_0 < \infty$ and $c_3 > 0$, depending on N, h_N, m, κ and $\|e^{-U}\|_\infty$ but not on x, z , such that

$$\begin{aligned} \sup_{x,z} \mathbb{P}^{x,z} [\#(\Lambda^N \cap \overline{B}_M(z, h_N)) \geq K] + \sup_x \mathbb{P}^x [\#(\Lambda^N \cap \overline{B}_M(x, h_N)) \geq K] \\ \leq 2e^{-c_3 K \log K} \end{aligned} \quad (3.2.17)$$

for $K \geq K_0$. By inserting (3.2.17) into (3.2.13)–(3.2.14) and using Bishop–Gromov inequality, we obtain, for $K \geq K_0$ and $C_3 > 0$ large enough,

$$\mathbb{P}^x [\lambda'_N(R) \geq \|W_{N;\alpha}\|_\infty K] \leq C_3 (1 + \text{Vol}_\kappa(R)) e^{-c_3 K \log K}. \quad (3.2.18)$$

By Remark 3.8, for $R \geq 1$ the factor $1 + \text{Vol}_\kappa(R)$ grows at most polynomially if $\kappa = 0$ and at most exponentially if $\kappa < 0$. Consequently, choosing $R = K \in \mathbb{N}$, the right-hand side of (3.2.18) is summable in K ;

$$\sum_{K=1}^{\infty} \mathbb{P}^x [\lambda'_N(K) \geq \|W_{N;\alpha}\|_\infty K] < \infty.$$

Hence, by the Borel–Cantelli lemma, $\mathbb{P}^x$ -almost surely there exists a random $K_1 = K_1(x) < \infty$ such that

$$\lambda'_N(K) \leq \|W_{N;\alpha}\|_\infty K, \quad K \geq K_1. \quad (3.2.19)$$

Since $R \mapsto \lambda'_N(R)$ is nondecreasing, for every sufficiently large R , taking $K = \lceil R \rceil$ yields

$$\lambda'_N(R) \leq \|W_{N;\alpha}\|_\infty (1 + R). \quad (3.2.20)$$

Substituting (3.2.20) into (3.2.11), for R_γ large enough, we obtain

$$R_\gamma^2 \leq h_N^2 r^2 (1 + \lambda'_N(R_\gamma)) \leq (1 + \|W_{N;\alpha}\|_\infty) h_N^2 r^2 (1 + R_\gamma)$$

and hence $R_\gamma \leq C_1 r^2$ for all r large enough. Using the deterministic upper bound (3.2.2) on $\|W_{N;\alpha}\|_\infty$, we may choose

$$C_1 := 2 \left(1 + \frac{\|k\|_\infty}{N h_N^m} \left(\frac{N h_N^m}{k(0)} \right)^{2\alpha} \right) h_N^2.$$

Together with (3.2.9), this proves the claim (3.2.7) under $\mathbb{P}^x$.

Step 3 (Verification of the stochastic completeness criterion). Since we have proved the inclusion

$$B_{G_N}(x, r) \subseteq \Lambda^N \cap B_M(x, C_1 r^2) \quad \text{for all } r \text{ large enough, } \mathbb{P}^x\text{-a.s.},$$

then, for some sufficiently large $r_x > 0$,

$$\int_{r_x}^\infty \frac{r}{\log \# B_{G_N}(x, r)} dr \geq \int_{r_x}^\infty \frac{r}{\log \# (\Lambda^N \cap B_M(x, C_1 r^2))} dr, \quad \mathbb{P}^x\text{-a.s.} \quad (3.2.21)$$

Recall that the Bishop–Gromov inequality yields

$$N \mu(B_M(x, R)) \leq C_2(R), \quad C_2(R) := N \|e^{-U}\|_\infty \text{Vol}_\kappa(R).$$

Hence, by Bennett's inequality with $s = 1$,

$$\begin{aligned} \mathbb{P}^x [\# (\Lambda^N \cap B_M(x, R)) \geq 1 + 2C_2(R)] &\leq \mathbb{P}[\text{Poisson}(C_2(R)) \geq 2C_2(R)] \\ &\leq \exp(-C_2(R)\Phi(1)). \end{aligned} \quad (3.2.22)$$

By Remark 3.8, $\text{Vol}_\kappa(R)$ grows polynomially when $\kappa = 0$ and exponentially when $\kappa < 0$. In particular, for $\kappa \leq 0$, the right-hand side of (3.2.22) is summable in $R = K \in \mathbb{N}$;

$$\sum_{K=1}^\infty \mathbb{P}^x [\# (\Lambda^N \cap B_M(x, K)) \geq 1 + 2N \|e^{-U}\|_\infty \text{Vol}_\kappa(K)] < \infty.$$

Hence, by the Borel–Cantelli lemma, $\mathbb{P}^x$ -almost surely there exists a random $K_2 = K_2(x) < \infty$ such that

$$\# (\Lambda^N \cap B_M(x, K)) \leq 1 + 2N \|e^{-U}\|_\infty \text{Vol}_\kappa(K), \quad K \geq K_2.$$

Therefore, there exists a constant $C_4 > 0$, depending on $N, \alpha, h_N, m, \kappa, k$ and $\|e^{-U}\|_\infty$, but not on x nor on the realization of the PPP, such that, for all sufficiently large r ,

$$\# (\Lambda^N \cap B_M(x, C_1 r^2)) \leq C_4 e^{C_4 r^2}, \quad \mathbb{P}^x\text{-a.s.} \quad (3.2.23)$$

It follows from (3.2.21) that

$$\int_{r_x}^{\infty} \frac{r}{\log \#B_{G_N}(x, r)} dr \gtrsim \int_{r_x}^{\infty} \frac{dr}{r} = +\infty.$$

Hence, by Theorem 3.6, the weighted random walk is stochastically complete in the connected component of x , and therefore the graphical construction defines the SEP in that connected component for all time.

It remains only to make the preceding Palm-a.s. conclusion simultaneous over all sampled roots. Let F_x denote the event that either (3.2.19) or (3.2.23) fails to hold eventually. By the two Borel–Cantelli arguments above, $\mathbb{P}^x(F_x) = 0$ for every deterministic $x \in M$. Hence, for every compact set $S \subset M$, the Campbell–Mecke formula [70, Theorem 4.1] gives

$$\mathbb{E} \left[\sum_{x \in \Lambda^N \cap S} \mathbf{1}_{F_x} \right] = N \int_S \mathbb{P}^x(F_x) e^{-U(x)} d\text{Vol}_M(x) = 0.$$

Taking a countable compact exhaustion of M , we conclude that almost surely none of the sampled points is bad. Thus the above argument holds simultaneously for every $x \in \Lambda^N$, and we obtain the global existence of the SEP, which concludes the proof. $\square$

Remark 3.11. For $\alpha \geq 1/2$, existence can alternatively be obtained directly from Liggett’s criterion [74, Proposition 6.1]. Indeed, by continuity near 0 and $k(0) > 0$, choose $r_* \in (0, 1)$ and $k_* > 0$ such that $k(r) \geq k_*$ on $[0, r_*]$. As a consequence of the Bishop–Gromov inequality, for the fixed radius h_N , every ball $B_M(x, h_N)$ can be covered by at most $\mathcal{N}$ sets $S_i(x)$ of diameter at most $r_* h_N$; $\mathcal{N}$ is independent of x . If $n_i = \#(S_i(x) \cap \Lambda^N)$ and $y \in S_i(x) \cap \Lambda^N$, then

$$\widehat{k}_N(y) \geq \frac{1}{N} \sum_{z \in S_i(x) \cap \Lambda^N} h_N^{-m} k \left(\frac{d_M(y, z)}{h_N} \right) \geq \frac{k_* n_i}{N h_N^m},$$

while $A_N(x, y) := N^{-1} h_N^{-m} k(d_M(x, y)/h_N) \leq \|k\|_{\infty} / (N h_N^m)$. Therefore, summing over the covering set $\cup_i S_i(x) \supseteq B(x, h_N)$,

$$\sum_{y \sim x} \frac{A_N(x, y)}{\widehat{k}_N(y)} \leq \sum_i \sum_{y \in S_i(x) \cap \Lambda^N} \frac{A_N(x, y)}{\widehat{k}_N(y)} \leq \mathcal{N} \frac{\|k\|_{\infty}}{k_*}.$$

Hence, by Cauchy–Schwarz and $\sum_y A_N(x, y) = \widehat{k}_N(x)$,

$$\begin{aligned} \lambda_{N;1/2}(x) &= \sum_{y \sim x} \frac{A_N(x, y)}{\left(\widehat{k}_N(x) \widehat{k}_N(y)\right)^{1/2}} \\ &\leq \left(\frac{\sum_{y \sim x} A_N(x, y)}{\widehat{k}_N(x)} \right)^{1/2} \left(\sum_{y \sim x} \frac{A_N(x, y)}{\widehat{k}_N(y)} \right)^{1/2} \\ &\leq \sqrt{\mathcal{N} \|k\|_\infty / k_*}. \end{aligned}$$

Thus $\sup_x \lambda_{N;1/2}(x) < \infty$. If $\alpha > 1/2$, (3.2.1) gives

$$W_{N;\alpha}(x, y) = W_{N;1/2}(x, y) \left(\widehat{k}_N(x) \widehat{k}_N(y)\right)^{1/2-\alpha} \leq \left(\frac{k(0)}{N h_N^m} \right)^{1-2\alpha} W_{N;1/2}(x, y),$$

and the same criterion applies.

For $0 \leq \alpha < 1/2$, Liggett’s criterion need not hold, even though the SEP still exists by Theorem 3.10. For example, for a homogeneous PPP on $\mathbb{R}^m$, arbitrarily large isolated clusters occur almost surely. Any isolated cluster of n points contained in a ball of diameter at most $r_* h_N$ satisfies

$$\widehat{k}_N(x) \leq \frac{n \|k\|_\infty}{N h_N^m} \quad \text{for every cluster point } x,$$

while $A_N(x, y) \geq k_*/(N h_N^m)$ for any two distinct cluster points. Hence

$$\lambda_{N;\alpha}(x) \geq (n-1) \frac{k_*}{N h_N^m} \left(\frac{n \|k\|_\infty}{N h_N^m} \right)^{-2\alpha} \asymp n^{1-2\alpha}.$$

Thus, for $\alpha < 1/2$, arbitrarily large isolated clusters imply $\sup_{x \in \Lambda^N} \lambda_{N;\alpha}(x) = +\infty$ almost surely.

Remark 3.12. Another commonly used road to show the existence of the SEP is by a percolation argument, as in [54], which boils down to show the finiteness of “active clusters” of a site x , uniformly in x . In our setting this boils down to adapt arguments from [78] for the Poisson–Boolean model.

3.3 Abstract hydrodynamic limit via duality

3.3.1 Duality and hypotheses on the random neighborhood graph

The SEP satisfies a duality property with a single random walk due to the symmetry of the jump rates. In words, this property means that the SEP generator acting on the η variable

of the function $D(x, \eta) = \eta(x)$ has the same effect as the single-random walk generator acting on the x variable of that same function (which is therefore named single particle duality function).

Definition 3.13 (Random walk (RW)). *Given a graph $G = (V, E)$ with weights W , the random walk (RW) is the continuous-time Markov jump processes on V whose generator evaluated on functions $\varphi : V \rightarrow \mathbb{R}$ is given by*

$$L_G^{(\text{RW})} \varphi(x) := \sum_{y \sim x} W(x, y) (\varphi(y) - \varphi(x)). \quad (3.3.1)$$

The duality between the RW and the SEP described above manifests itself as follows: for pairings between the empirical measure of a configuration η and a local function φ on V , it holds

$$L_G^{(\text{SEP})} \left\langle \varphi, \sum_{x \in V} \eta(x) \delta_x \right\rangle = \left\langle L_G^{(\text{RW})} \varphi, \sum_{x \in V} \eta(x) \delta_x \right\rangle.$$

Before stating and proving our abstract hydrodynamic limit for the SEP via duality, we introduce three assumptions on the sequence of graphs $(G_N = (V_N, E_N))_{N \geq 1}$ with symmetric weights $(W_N)_{N \geq 1}$ and bandwidth parameters $(h_N)_{N \geq 1} \downarrow 0$:

(i) (Initial data is well-prepared). Assume that the empirical measure

$$\widehat{\pi}_0^N := \frac{1}{N} \sum_{x \in V_N} \eta_0^N(x) \delta_x$$

associated to the initial configuration η_0^N of the SEP on G_N converges vaguely to $\rho_0 d\text{Vol}_M$ in probability; that is,

$$\forall \varphi \in C_0(M), \quad \left\langle \varphi, d\widehat{\pi}_0^N \right\rangle \xrightarrow{\text{pr.}} \langle \varphi, \rho_0 d\text{Vol}_M \rangle \quad \text{as } N \rightarrow +\infty; \quad (\text{I})$$

(ii) (Convergence of the reference empirical measure). Set $\widehat{\nu}^N := \frac{1}{N} \sum_{x \in V_N} \delta_x$. Assume that, vaguely on M ,

$$\widehat{\nu}^N \longrightarrow \nu = q d\text{Vol}_M, \quad (\text{II})$$

for a locally finite measure ν .

(iii) (Convergence of the generator). Let $\mathcal{L}$ be the generator of a diffusion process on M whose core is $C_c^\infty(M)$. Assume that the diffusively rescaled generator $h_N^{-2} L_{G_N}^{(\text{RW})}$ converges towards the generator $\mathcal{L}$ in the following sense:

$$\forall \varphi \in C_c^\infty(M), \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{x \in V_N} \left| h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(x) - \mathcal{L} \varphi(x) \right| = 0. \quad (\text{III})$$

Note that assumptions (i) and (ii) imply that $0 \leq \rho_0 d\text{Vol}_M \leq \nu$ since $\widehat{\pi}_0^N \leq \widehat{\nu}^N$.

3.3.2 Hydrodynamic limit of the SEP

Theorem 3.14. (*Abstract hydrodynamic limit*) Let $(G_N = (V_N, E_N))_{N \geq 1}$ be a sequence of graphs on M with symmetric weights $(W_N)_{N \geq 1}$ and bandwidth parameters $(h_N)_{N \geq 1}$. If the assumptions (I), (II), and (III) are satisfied, then the trajectory of the empirical measure $t \mapsto \hat{\pi}_t^N$ associated to a configuration η_t^N that evolves according to the diffusively rescaled SEP with generator $h_N^{-2} L_{G_N}^{(\text{SEP})}$ converges in probability in $D([0, T], \mathcal{M}_{\text{loc}}(M))$, endowed with the Skorokhod topology associated with vague convergence, to the deterministic trajectory $t \mapsto \rho_t d\text{Vol}_M$ whose density solves weakly the forward Kolmogorov equation

$$\partial_t \rho_t = \mathcal{L}^* \rho_t \quad (3.3.2)$$

with initial condition ρ_0 of assumption (I), provided this equation has a unique continuous weak solution in the class $0 \leq \rho_t d\text{Vol}_M \leq \nu$; the adjoint is understood in $L^2(M, d\text{Vol}_M)$.

Proof. We outline the argument for complete manifolds modified from the torus case found in the monograph of Kipnis and Landim [67, Chapter 4].

Step 1 (The Dynkin martingale argument). We pair any test function $\varphi \in C_c^\infty(M)$ with the empirical measure $\hat{\pi}_t^N$ associated to the configuration η_t (with $0 \leq t \leq T$ for some time $T > 0$), and we consider its associated Dynkin martingale;

$$M_t^N[\varphi] = \langle \varphi, \hat{\pi}_t^N \rangle - \langle \varphi, \hat{\pi}_0^N \rangle - \int_0^t h_N^{-2} L_{G_N}^{(\text{SEP})} \langle \varphi, \hat{\pi}_s^N \rangle ds.$$

By duality between the SEP and the RW,

$$h_N^{-2} L_{G_N}^{(\text{SEP})} \langle \varphi, \hat{\pi}_s^N \rangle = \langle h_N^{-2} L_{G_N}^{(\text{RW})} \varphi, \hat{\pi}_s^N \rangle.$$

Therefore, convergence of generators can be established at the level of random walks for a single particle, which is guaranteed by assumption (III), and the above yields

$$M_t^N[\varphi] = \langle \varphi, \hat{\pi}_t^N \rangle - \langle \varphi, \hat{\pi}_0^N \rangle - \int_0^t \langle \mathcal{L} \varphi, \hat{\pi}_s^N \rangle ds + o(1).$$

Here the error is uniform in $0 \leq t \leq T$, because $0 \leq \eta_s^N \leq 1$ gives

$$\sup_{t \leq T} \left| \int_0^t \langle h_N^{-2} L_{G_N}^{(\text{RW})} \varphi - \mathcal{L} \varphi, \hat{\pi}_s^N \rangle ds \right| \leq T \frac{1}{N} \sum_{x \in V_N} \left| h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(x) - \mathcal{L} \varphi(x) \right| \rightarrow 0.$$

To show that the Dynkin martingale goes to 0 in probability, it is easier to consider its quadratic variation

$$\langle M^N[\varphi] \rangle_t = \int_0^t h_N^{-2} L_{G_N}^{(\text{SEP})} \langle \varphi, \hat{\pi}_s^N \rangle^2 - 2 \langle \varphi, \hat{\pi}_s^N \rangle h_N^{-2} L_{G_N}^{(\text{SEP})} \langle \varphi, \hat{\pi}_s^N \rangle ds.$$

Since an exchange across an undirected edge $\{x, y\}$ changes $\langle \varphi, \widehat{\pi}^N \rangle$ by $N^{-1}(\eta(y) - \eta(x))(\varphi(x) - \varphi(y))$, one has

$$\begin{aligned} \langle M^N[\varphi] \rangle_t &\leq \frac{t}{N^2} h_N^{-2} \sum_{\{x, y\} \in E_N} W_N(x, y) (\varphi(y) - \varphi(x))^2 \\ &= \frac{t}{N^2} \sum_{x \in V_N} \varphi(x) h_N^{-2} \left(-L_{G_N}^{(\text{RW})} \right) \varphi(x). \end{aligned}$$

Hence, by (II) and (III),

$$\mathbb{E}[\langle M^N[\varphi] \rangle_T] \leq \frac{T}{N} \left(\frac{1}{N} \sum_{x \in V_N} |h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(x) - \mathcal{L}\varphi(x)| + \langle \widehat{\nu}^N, |\mathcal{L}\varphi| \rangle \right) = O(N^{-1}).$$

By Doob's inequality,

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |M_t^N[\varphi]|^2 \right] \leq 4 \mathbb{E} [\langle M^N[\varphi] \rangle_T] \rightarrow 0,$$

hence $\sup_{t \leq T} |M_t^N[\varphi]| \rightarrow 0$ in probability; that is,

$$\sup_{0 \leq t \leq T} \left| \langle \varphi, \widehat{\pi}_t^N \rangle - \langle \varphi, \widehat{\pi}_0^N \rangle - \int_0^t \langle \mathcal{L}\varphi, \widehat{\pi}_s^N \rangle ds \right| \xrightarrow{\text{pr.}} 0 \quad \text{as } N \rightarrow +\infty.$$

Step 2 (Relative compactness of the associated measure). We equip the space of càdlàg paths with the probability measure associated to $t \mapsto \widehat{\pi}_t^N$ that we denote by Q^N . Then the convergence obtained in Step 1 implies that

$$Q^N \left[\left\{ \nu \in D([0, T], \mathcal{M}_{\text{loc}}(M)) : \sup_{0 \leq t \leq T} \left| \langle \varphi, \nu_t \rangle - \langle \varphi, \nu_0 \rangle - \int_0^t \langle \mathcal{L}\varphi, \nu_s \rangle ds \right| \leq \epsilon \right\} \right] \rightarrow 1$$

as $N \rightarrow +\infty$, for all $\epsilon > 0$, where $D([0, T], \mathcal{M}_{\text{loc}}(M))$ denotes the Skorokhod path space of càdlàg locally finite measures equipped with the vague topology. It remains to show relative compactness of the associated measures Q^N .

Since we work in the vague topology on measures, we use a fixed countable determining family in $C_c^\infty(M)$ and combine tightness of all its projections with compact containment. Namely, consider the real-valued random trajectories $\{\langle \pi_t^N, \varphi \rangle, 0 \leq t \leq T\}$, for every smooth compactly supported test function $\varphi \in C_c^\infty(M)$. Compact containment follows from $0 \leq \widehat{\pi}_t^N \leq \nu_N$: for a compact exhaustion $(K_j)_{j \geq 1}$ of M choose $\chi_j \in C_c(M)$ with $\chi_j \geq 1$ on K_j ; then $\sup_t \widehat{\pi}_t^N(K_j) \leq \nu_N(\chi_j)$, and the latter is uniformly bounded for large N by (II). Relative compactness of this sequence of trajectories is in turn a consequence of Aldous' criterion;

$$\lim_{\Theta \rightarrow 0} \limsup_{N \rightarrow +\infty} \sup_{\tau \in \mathcal{I}_T, \theta \leq \Theta} Q^N \left[\left\{ \widehat{\pi}^N : \left| \langle \varphi, \widehat{\pi}_{(\tau+\theta) \wedge T}^N \rangle - \langle \varphi, \widehat{\pi}_\tau^N \rangle \right| > \epsilon \right\} \right] = 0,$$

where $\mathcal{I}_T$ is the set of all stopping times bounded by T . To verify Aldous's criterion, we rewrite the difference of empirical measure using the Dynkin martingale;

$$\left\langle \varphi, \widehat{\pi_{(\tau+\theta) \wedge T}^N} \right\rangle - \left\langle \varphi, \widehat{\pi_\tau^N} \right\rangle = M_{(\tau+\theta) \wedge T}^N[\varphi] - M_\tau^N[\varphi] + \int_\tau^{(\tau+\theta) \wedge T} \left\langle \mathcal{L}\varphi, \widehat{\pi_s^N} \right\rangle ds + o(1).$$

The integral is of order θ uniformly in N , since $|\langle \mathcal{L}\varphi, \widehat{\pi_s^N} \rangle| \leq \langle |\mathcal{L}\varphi|, \nu_N \rangle$ and (II) bounds the latter. For the difference of the martingales, Chebyshev's inequality yields

$$\begin{aligned} Q^N \left[\left\{ \widehat{\pi^N} : \left| M_{(\tau+\theta) \wedge T}^N[\varphi] - M_\tau^N[\varphi] \right| > \epsilon \right\} \right] &\leq \frac{1}{\epsilon^2} \mathbb{E} \left[\left| M_{(\tau+\theta) \wedge T}^N[\varphi] - M_\tau^N[\varphi] \right|^2 \right] \\ &\leq \frac{1}{\epsilon^2} \mathbb{E} \left[\langle M^N[\varphi] \rangle_T \right] \\ &= o(1). \end{aligned}$$

Moreover each jump of $\langle \varphi, \widehat{\pi_t^N} \rangle$ has size at most $2\|\varphi\|_\infty/N$; hence every subsequential limit of the projected process has continuous paths. Since the countable determining family characterizes the vague topology, every subsequential limit of $\widehat{\pi^N}$ is therefore concentrated on $C([0, T], \mathcal{M}_{\text{loc}}(M))$.

Step 3 (Uniqueness of the limit). Now that the relative compactness of the sequence of measures $\{Q^N\}_{N \geq 1}$ is established, one argues that all the limits of the weakly convergent subsequences are the same, whence Q^N converges weakly to a limiting measure Q . This limiting measure is concentrated on paths that weakly solve the forward Kolmogorov equation. To see this, consider a countable dense determining family $\mathcal{D} \subset C_c^\infty(M)$, and for a given test function $\varphi \in \mathcal{D}$ the set

$$H_\epsilon := \left\{ \nu \in D([0, T], \mathcal{M}_{\text{loc}}(M)) : \sup_{0 \leq t \leq T} \left| \langle \varphi, \nu_t \rangle - \langle \varphi, \nu_0 \rangle - \int_0^t \langle \mathcal{L}\varphi, \nu_s \rangle ds \right| \leq \epsilon \right\}.$$

Pick any converging subsequence $\{Q^{N_m}\}_{m \geq 1}$. Because every limit path is continuous, the displayed functional is continuous at Q -a.e. path; the convergence from Step 1 therefore passes to the limit and gives $Q[H_\epsilon] = 1$. It follows that

$$Q[H_0] = Q[\cap_{m \geq 1} H_{1/m}] = 1 - Q[\cup_{m \geq 1} H_{1/m}^C] = 1,$$

which in turn implies that

$$\begin{aligned} Q \left[\left\{ \nu \in D([0, T], \mathcal{M}_{\text{loc}}(M)) \right. \right. \\ \left. \left. : \sup_{0 \leq t \leq T} \left| \langle \varphi, \nu_t \rangle - \langle \varphi, \nu_0 \rangle - \int_0^t \langle \mathcal{L}\varphi, \nu_s \rangle ds \right| = 0 \quad \forall \varphi \in \mathcal{D} \right\} \right] = 1. \end{aligned}$$

Moreover, for Q -a.e. path $\pi = \{\pi_t\}_{t \in [0, T]}$ one has $0 \leq \pi_t \leq \nu$. Indeed, for every non-negative $f \in C_c(M)$, $\langle f, \widehat{\pi}_t^N \rangle \leq \langle f, \widehat{\nu}^N \rangle$ for all t . Passing to a subsequential limit at continuity times and using continuity of the limit path together with (II) gives $\langle f, \pi_t \rangle \leq \langle f, \nu \rangle$ for every t . Thus $\pi_t \ll \nu = q \, d\text{Vol}_M$, and hence $\pi_t = \rho_t \, d\text{Vol}_M$ for some $0 \leq \rho_t \leq q$.

The weak identity from Step 1 and (I) show that every limit weakly solves (3.3.2) with initial datum ρ_0 by density. By the uniqueness assumed in the theorem, every subsequence has the same deterministic limit. $\square$

We now explain how to go from Theorem 3.14 towards the main result Theorem 3.4.

For the RNG of Definition 3.1, assumption (II) holds almost surely with $\nu = d\mu$; for every $f \in C_c(M)$, Poisson concentration gives $N^{-1} \sum_{x \in \Lambda^N} f(x) \rightarrow \int f \, d\mu$ almost surely, first on a countable determining family and then vaguely. Assumption (III) on the convergence of discrete Laplacian towards weighted Laplacian is the content of Theorem 3.16 for RNG; the limiting diffusion operator is $\mathcal{L} = \mathcal{L}_\alpha = C_{k; \alpha} e^{(2\alpha-1)U} \Delta_{2(1-\alpha)}$. The fact that $\mathcal{L}$ has $C_c^\infty(M)$ as core is part of the semigroup hypothesis in Theorem 3.4. Assumption (I) is made w.r.t. the reference Gibbs measure $d\mu = e^{-U} d\text{Vol}_M$ in Theorem 3.4 instead of w.r.t. the volume measure $d\text{Vol}_M$. Hence, the corresponding density w.r.t. the volume measure $d\text{Vol}_M$ is $\rho_t e^{-U}$. Therefore, Theorem 3.14, in particular (3.3.2), gives

$$\partial_t(\rho_t e^{-U}) = \mathcal{L}_\alpha^*(\rho_t e^{-U}) = e^{-U} \mathcal{L}_\alpha \rho_t,$$

because $\mathcal{L}_\alpha$ is symmetric in $L^2(M, d\mu)$. In particular, we recover the hydrodynamic limit (3.1.5);

$$\partial_t \rho_t = \partial_t \rho_t = \mathcal{L}_\alpha \rho_t.$$

The existence of the SEP is the content of Theorem 3.10.

3.4 Convergence of discrete random walks

In the proof of Theorem 3.14, a specific notion of convergence of the RW towards a limiting differential operator is used to rewrite (almost surely in the case of RNG)

$$h_N^{-2} L_{G_N}^{(\text{SEP})} \langle \varphi(x), \widehat{\pi}_t^N \rangle = \langle h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(x), \widehat{\pi}_t^N \rangle = \langle \mathcal{L} \varphi(x), \widehat{\pi}_t^N \rangle + o(1). \quad (3.4.1)$$

The almost-sure pointwise consistency of graph Laplacians was obtained by Hein–Audibert–von Luxburg [56, Theorems 28 & 30] for random neighbourhood graphs sampled from i.i.d. points (see also [8, Theorem 2.1]). We obtain their counterpart on our RNG obtained from PPP with Gibbs intensity.

To perform the RW, we consider the same weights as in the main theorem, namely

$$W_{N;\alpha}(X^i, X^j) := \frac{1}{N} \frac{h_N^{-m} k\left(\frac{d_M(X^i, X^j)}{h_N}\right)}{\left(\widehat{k}_N(X^i) \widehat{k}_N(X^j)\right)^\alpha}, \quad \widehat{k}_N(x) := \frac{1}{N} \sum_{z \in \Lambda^N} h_N^{-m} k\left(\frac{d_M(x, z)}{h_N}\right). \quad (3.4.2)$$

Note that $\widehat{k}_N(x)$ is an approximation for $e^{-U(x)}$ as proved in [Lemma 3.29](#).

Remark 3.15. *Hein–Audibert–von Luxburg consider the manifold M as a submanifold isometrically embedded in some euclidean space $\mathbb{R}^M$. They replace the Riemannian distance with the Euclidean one using the fact that for close enough points, the intrinsic and extrinsic distances match up to the third order ([91, Proposition 6]). We stick to the intrinsic setting for our purpose, namely, the hydrodynamic limit of the SEP, and we show the consistency of the graph Laplacian with (3.4.2) in our setting of RNG obtained from a PPP.*

Theorem 3.16. *Let $\varphi \in C_c^\infty(M)$ and let $(G_N)_{N \geq 1}$ be a sequence of RNG in the sense of [Definition 3.1](#) with symmetric weights $(W_{N;\alpha})_{N \geq 1}$ given by (3.4.2), for a fixed $\alpha \geq 0$, and associated RW with diffusively rescaled generator $h_N^{-2} L_{G_N}^{(\text{RW})}$. There exist constants $C_{\alpha;1}, C_{\alpha;2}, C_{\alpha;3} > 0$ that only depends on α, k, φ, M and μ , such that, for all $\varepsilon > 0$ and h_N small enough,*

$$\begin{aligned} & \mathbb{P} \left[\sup_{X^i \in \Lambda^N} \left| h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(X^i) - \left(C_{k;\alpha} e^{(2\alpha-1)U(X^i)} \Delta_{2(1-\alpha)} \varphi(X^i) + R_{N,\alpha,\varphi}(X^i) \right) \right| > \varepsilon \right] \\ & \leq C_{\alpha;1} N e^{-C_{\alpha;2} N h_N^{m+2} \varepsilon^2} \end{aligned} \quad (3.4.3)$$

with

$$\|R_{N,\alpha,\varphi}\|_\infty \leq C_{\alpha;3} h_N^2.$$

Then, if both $h_N \rightarrow 0$ and $N h_N^{m+2} / \log N \rightarrow +\infty$ as $N \rightarrow +\infty$, it holds almost surely and uniformly in the vertices X^i 's,

$$\lim_{N \rightarrow +\infty} h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(X^i) = C_{k;\alpha} e^{(2\alpha-1)U(X^i)} \Delta_{2(1-\alpha)} \varphi(X^i). \quad (3.4.4)$$

Remark 3.17. *In the compact setting, van Ginkel–Redig [99] obtain almost sure consistency of the graph Laplacian under a more restrictive condition on the bandwidth parameter h_N ; $h_N \downarrow 0$ at a much slower rate. In dimension $m \geq 3$, their argument requires $N h_N^{m(m+4)} / \log N \rightarrow +\infty$ (this is not explicitly computed in [99]). The reason is that their proof is based on the almost-sure weak convergence of empirical measures, which is metrized by the 1-Wasserstein distance. Of course requiring such convergence in W_1 is more restrictive than a general concentration argument as one has to cover Lipschitz functions by Kantorovich duality.*

We will use the following concentration inequality for Poisson integrals:

Proposition 3.18 (Bernstein inequality [86, Proposition 7]). *Let Λ be a Poisson point process on a measurable space $(\mathbb{X}, \mathcal{X})$ with intensity measure ν . Let $f : X \rightarrow \mathbb{R}$ be measurable and essentially bounded, with $\int_X f^2(x) d\nu(x) < \infty$. Then, for every $\xi > 0$,*

$$\mathbb{P} \left[\left| \int_X f(x) (d\Lambda(x) - d\nu(x)) \right| \geq \xi \right] \leq 2 \exp \left(- \frac{\xi^2}{2 \int_X f^2(x) d\nu(x) + \frac{2}{3} \xi \|f\|_\infty} \right). \quad (3.4.5)$$

Proof of Theorem 3.16. Set

$$Y(x; y) := h_N^{-m} k \left(\frac{d_M(x, y)}{h_N} \right), \quad \widehat{k}_N(x) := \frac{1}{N} \sum_{X^l \in \Lambda^N} Y(x; X^l),$$

and introduce the normalizer

$$q_{h_N}(x) := \int_M Y(x; y) d\mu(y). \quad (3.4.6)$$

Define

$$Z(x; y) := h_N^{-2} \frac{Y(x; y) (\varphi(y) - \varphi(x))}{(\widehat{k}_N(x) \widehat{k}_N(y))^\alpha}, \quad Z_{h_N}(x; y) := h_N^{-2} \frac{Y(x; y) (\varphi(y) - \varphi(x))}{(q_{h_N}(x) q_{h_N}(y))^\alpha}$$

together with

$$\overline{Z}(x) := \frac{1}{N} \sum_{X^l \in \Lambda^N} Z(x; X^l), \quad \overline{Z}_{h_N}(x) := \frac{1}{N} \sum_{X^l \in \Lambda^N} Z_{h_N}(x; X^l).$$

Note that for $x \in \Lambda^N$,

$$\overline{Z}(x) = h_N^{-2} L_{G_N}^{(\text{RW})} \varphi(x).$$

Moreover, because $Z_{h_N}(x; x) = 0$, the Palm self-point does not contribute and

$$\mathbb{E}^x[\overline{Z}_{h_N}(x)] = \int_M Z_{h_N}(x; y) d\mu(y). \quad (3.4.7)$$

Using Lemma 3.30, define the remainder

$$R_{N, \alpha, \varphi}(x) := \mathbb{E}^x[\overline{Z}_{h_N}(x)] - C_{k; \alpha} e^{(2\alpha-1)U(x)} \Delta_{2(1-\alpha)} \varphi(x), \quad (3.4.8)$$

which satisfies $\|R_{N, \alpha, \varphi}\|_\infty \leq Ch_N^2$.

Step 1 (Concentration estimate for the kernel). Let $h_N \leq h_1$. We are only interested in the Poisson points inside the h_1 -thickening or $2h_1$ -thickening of the support of φ since the empirical quantities at stake are zero otherwise, where the t -thickening is defined as

$$(\text{Supp } \varphi)_t := \left\{ z \in M; d_M(z, \text{Supp } \varphi) \leq t \right\}.$$

By [Lemma 3.29](#), after reducing h_1 if necessary there are constants $0 < c' < C' < \infty$ such that

$$c' \leq q_{h_N}(x) \leq C', \quad x \in (\text{Supp } \varphi)_{2h_1}. \quad (3.4.9)$$

Let $\delta > 0$ be a parameter that will be tuned in terms of ε later on. Fix a sampled center $x \in \Lambda^N$. Under its Palm law $\mathbb{P}^x$, the configuration is represented as an independent PPP of intensity $N\mu$ together with the fixed point x . Therefore,

$$\widehat{k}_N(x) = \frac{k(0)}{Nh_N^m} + \frac{1}{N} \sum_{z \in \Lambda^N \setminus \{x\}} Y(x; z). \quad (3.4.10)$$

Assume that

$$\frac{k(0)}{Nh_N^m} \leq \frac{\delta}{2}. \quad (3.4.11)$$

By [\(3.4.10\)](#),

$$\begin{aligned} \mathbb{P}^x \left[\left| \widehat{k}_N(x) - q_{h_N}(x) \right| > \delta \right] &\leq \mathbb{P}^x \left[\left| \frac{1}{N} \sum_{X^l \in \Lambda^N \setminus \{x\}} Y(x; X^l) - q_{h_N}(x) \right| > \frac{\delta}{2} \right] \\ &= \mathbb{P}^x \left[\left| \frac{1}{N} \int_M Y(x; z) (d\Xi^N(z) - Nd\mu(z)) \right| > \frac{\delta}{2} \right], \end{aligned}$$

where $\Xi^N := \Lambda^N \setminus \{x\}$ is a PPP of intensity $Nd\mu$. We estimate

$$\frac{1}{N} |Y(x; X^l)| \leq \frac{\|k\|_\infty}{Nh_N^m}, \quad (3.4.12)$$

and

$$\begin{aligned} \text{Var}^x \left(\frac{1}{N} \sum_{X^l \in \Lambda^N \setminus \{x\}} Y(x; X^l) \right) &= \frac{1}{N} \int_M Y(x; y)^2 d\mu(y) \\ &\leq \frac{\|k^2 e^{-U}\|_\infty}{Nh_N^{2m}} \text{Vol}_M(B(x, h_N)) \\ &\leq \frac{C_1}{Nh_N^m}, \end{aligned} \quad (3.4.13)$$

for some constant $C_1(k, \mu, M, \varphi) > 0$ uniformly for $x \in (\text{Supp } \varphi)_{2h_1}$, where in the last inequality we used the fact that, since φ is compactly supported, one can uniformly control the volume of balls in the h_1 -thickening of its support;

$$\text{Vol}_M(B(x, h_N)) = h_N^m + O(h_N^{2m}).$$

Then using (3.4.12) and (3.4.13) for Proposition 3.18, namely Bernstein's inequality, we obtain the exponential bound

$$\mathbb{P}^x \left[\left| \widehat{k}_N(x) - q_{h_N}(x) \right| > \delta \right] \leq 2 \exp \left(- \frac{N h_N^m \delta^2}{8C_1 + \frac{4}{3} \|k\|_\infty \delta} \right) \leq 2 \exp(-c_2 N h_N^m \delta^2)$$

for some constant $c_2(k, \mu, M, \varphi) > 0$, given that (3.4.11) holds for δ small enough. A union bound over the sampled points, followed by the Campbell–Mecke formula, gives

$$\begin{aligned} & \mathbb{P} \left[\sup_{x \in \Lambda^N |_{(\text{Supp } \varphi)_{2h_1}}} \left| \widehat{k}_N(x) - q_{h_N}(x) \right| > \delta \right] \\ & \leq \mathbb{E} \left[\sum_{x \in \Lambda^N \cap (\text{Supp } \varphi)_{2h_1}} \mathbf{1}_{\left\{ \left| \widehat{k}_N(x) - q_{h_N}(x) \right| > \delta \right\}} \right] \\ & = N \int_{(\text{Supp } \varphi)_{2h_1}} \mathbb{P}^x \left[\left| \widehat{k}_N(x) - q_{h_N}(x) \right| > \delta \right] d\mu(x) \\ & \leq C_3 N \exp(-c_2 N h_N^m \delta^2) \end{aligned} \tag{3.4.14}$$

for some constant $C_3(k, \mu, M, \varphi) > 0$.

Step 2 (Comparing the discrete and continuous RW). Conditioned on the event

$$E_\delta := \left(\sup_{x \in \Lambda^N |_{(\text{Supp } \varphi)_{2h_1}}} \left| \widehat{k}_N(x) - q_{h_N}(x) \right| \leq \delta \right),$$

take $\delta < c'/2$, and the elementary inequality $|a^{-\alpha} - b^{-\alpha}| \leq \alpha c^{-(\alpha+1)} |a - b|$ valid for all $a, b > c$ together with (3.4.9) yields

$$\begin{aligned} & \sup_{x, y \in \Lambda^N |_{(\text{Supp } \varphi)_{2h_1}}} \left| \frac{1}{(\widehat{k}_N(x) \widehat{k}_N(y))^\alpha} - \frac{1}{(q_{h_N}(x) q_{h_N}(y))^\alpha} \right| \\ & \leq \alpha (c' - \delta)^{-(2\alpha+2)} \sup_{x, y \in \Lambda^N |_{(\text{Supp } \varphi)_{2h_1}}} \left| \widehat{k}_N(x) \widehat{k}_N(y) - q_{h_N}(x) q_{h_N}(y) \right| \\ & \leq 2\alpha (c' - \delta)^{-(2\alpha+2)} (C' + \delta) \delta \\ & \leq C_4 \delta, \end{aligned}$$

where we used that $\widehat{k}_N \leq C' + \delta$ on $(\text{Supp } \varphi)_{2h_1}$ under E_δ . Hence, conditioned on E_δ ,

$$\begin{aligned} & \sup_{x \in \Lambda^N | (\text{Supp } \varphi)_{h_1}} |\overline{Z}(x) - \overline{Z}_{h_N}(x)| \\ & \leq \sup_{x, y \in \Lambda^N | (\text{Supp } \varphi)_{2h_1}} \left| \frac{1}{(\widehat{k}_N(x)\widehat{k}_N(y))^\alpha} - \frac{1}{(q_{h_N}(x)q_{h_N}(y))^\alpha} \right| \frac{\|\nabla \varphi\|_\infty}{h_N} \leq \frac{C_5 \delta}{h_N}. \end{aligned} \quad (3.4.15)$$

For a sampled point $x \in (\text{Supp } \varphi)_{h_1}$, we apply Bernstein's inequality to

$$\overline{Z}_{h_N}(x) - \mathbb{E}^x[\overline{Z}_{h_N}(x)] = \frac{1}{N} \int_M Z_{h_N}(x; z) (d\Lambda^N(z) - Nd\mu(z)),$$

which yields the exponential bound

$$\mathbb{P}^x [|\overline{Z}_{h_N}(x) - \mathbb{E}^x[\overline{Z}_{h_N}(x)]| > \varepsilon] \leq 2 \exp(-c_6 N h_N^{m+2} \varepsilon^2) \quad (3.4.16)$$

using that

$$\begin{aligned} \frac{1}{N} Z_{h_N}(x; X^l) &= \frac{1}{N h_N^2} \frac{h_N^{-m} k\left(\frac{d_M(x, X^l)}{h_N}\right) (\varphi(X^l) - \varphi(x))}{(q_{h_N}(x)q_{h_N}(X^l))^\alpha} \\ &\leq \frac{\|k\|_\infty \|\nabla \varphi\|_\infty (c')^{-2\alpha}}{N h_N^{m+1}} \\ &\leq \frac{C_7}{N h_N^{m+1}} \end{aligned}$$

and

$$\begin{aligned} \text{Var}^x(\overline{Z}_{h_N}(x)) &= \frac{1}{N} \int_M Z_{h_N}(x; z)^2 d\mu(z) \\ &\leq \frac{\|k^2 e^{-U}\|_\infty \|\nabla \varphi\|_\infty^2 (c')^{-4\alpha}}{N h_N^{2m+2}} \text{Vol}_M(B_M(x, h_N)) \\ &\leq \frac{C_8}{N h_N^{m+2}} \end{aligned}$$

uniformly for $x \in (\text{Supp } \varphi)_{h_1}$. A second application of a union bound together with the Campbell–Mecke formula yields

$$\mathbb{P} \left[\sup_{x \in \Lambda^N | (\text{Supp } \varphi)_{h_1}} |\overline{Z}_{h_N}(x) - \mathbb{E}^x[\overline{Z}_{h_N}(x)]| > \varepsilon \right] \leq C_9 N \exp(-c_6 N h_N^{m+2} \varepsilon^2). \quad (3.4.17)$$

Step 3 (Conclusion). Set $\delta := h_N \varepsilon$ in (3.4.14); under $N h_N^{m+2} / \log N \rightarrow \infty$, $k(0)/(N h_N^m)$ is $o(h_N)$, so the condition (3.4.11) holds for all large N . Combining (3.4.14),

(3.4.15), (3.4.17), and (3.4.8) yields the concentration estimate (3.4.3) in the theorem. This probability is summable in N in the considered regime $Nh_N^{m+2}/\log N \rightarrow +\infty$ with $h_N \rightarrow 0$ as $N \rightarrow +\infty$. The Borel–Cantelli lemma then yields almost sure convergence which concludes the proof. $\square$

3.5 Lifting the SEP to principal bundles

We extend the construction of RNG and the SEP to a total $\widetilde{m}$ -dimensional manifold $\widetilde{M}$ lying above the m -dimensional manifold M with $\text{Ric}_M \geq (m-1)\kappa$ for some $\kappa \leq 0$ (or assume that M is compact). We work in a bundle setting which allows for a canonical comparison of points in different fibers.

3.5.1 Geometric setting

Let G be a compact Lie group and $\pi : \widetilde{M} \rightarrow M$ be a principal G -bundle with right action $R_a(u) = u \cdot a$ ($u \in \widetilde{M}$, $a \in G$) and Lie algebra $\mathfrak{g} = T_e G$ equipped with a principal connection ω . We endow G with a bi-invariant Riemannian metric which yields a G -invariant metric $\widetilde{g}^\mathcal{V}$ on the vertical bundle $\mathcal{V} := \ker(d\pi)$; that is

$$\widetilde{g}^\mathcal{V}(u \cdot a)((R_a)_*w, (R_a)_*w') = \widetilde{g}^\mathcal{V}(u)(w, w'), \quad w, w' \in \mathcal{V}_u,$$

as follows: Given an Ad-invariant inner product $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$ on $\mathfrak{g}$ (equivalently, a bi-invariant Riemannian metric on G) and $\xi \in \mathfrak{g}$, define the *fundamental vertical vector field*

$$\xi^\#(u) := \left. \frac{d}{dt} \right|_{t=0} u \cdot \exp(t\xi) \in \mathcal{V}_u.$$

The map $\xi \mapsto \xi^\#(u) \in \mathcal{V}_u$ is a linear isomorphism $\mathfrak{g} \simeq \mathcal{V}_u$, hence it induces a metric $\widetilde{g}^\mathcal{V}$ on the *vertical bundle* $\mathcal{V} := \ker d\pi$ by

$$\widetilde{g}^\mathcal{V}(\xi^\#(u), \eta^\#(u)) := \langle \xi, \eta \rangle_{\mathfrak{g}}, \quad \xi, \eta \in \mathfrak{g}.$$

By Ad-invariance, $g_\mathcal{V}$ is G -invariant, namely, the right action $R_{(\cdot)}$ acts by fiberwise isometries on $\mathcal{V}$. We denote by $d_{\widetilde{M}_x}$ the distance in the fiber $\widetilde{M}_x := \pi^{-1}(x)$ above $x \in M$ and $m' := \dim \mathcal{V}$ so that $\widetilde{m} = m + m'$.

The principal connection ω defines the horizontal distribution $\mathcal{H} := \ker(\omega)$ and splits the tangent bundles as

$$T\widetilde{M} = \mathcal{H} \oplus \mathcal{V}.$$

With this splitting, we define the bundle metric,

$$\widetilde{g} := \pi^*g \oplus \widetilde{g}^\mathcal{V}. \tag{3.5.1}$$

If the compact Lie group G is semi-simple, then the (negative of the) killing form B defined by

$$B(\xi, \eta) := \text{Tr}(\text{Ad}(\xi)\text{Ad}(\eta)), \quad \xi, \eta \in \mathfrak{g},$$

also yields a bi-invariant metric $-B$ on $\mathfrak{g}$ [4, Corollary 2.33], and the total manifold $\widetilde{M}$ admits the metric

$$\widetilde{g} := \pi^*g \oplus (-B)\omega := \widetilde{g}^{\mathcal{H}} \oplus \widetilde{g}^{\mathcal{V}}.$$

This encompasses, for instance, the orthonormal frame bundle $\pi_{O(M)} : O(M) \rightarrow M$ (and $O(m)$ is compact semi-simple for $m \geq 3$) used in the construction of Riemannian Brownian motion by Eells–Elworthy–Malliavin (see [57, Section 3.1]),

The connection also allows to lift curves and tangent vectors from the base manifold M to the total manifold $\widetilde{M}$ in a coherent way.

Definition 3.19 (Horizontal lift). *Given a smooth curve $\gamma \in C^\infty(I; M)$ on an interval $I \ni 0$ and $u_0 \in \pi^{-1}(\gamma(0))$, the horizontal lift $\widetilde{\gamma}$ of γ with respect to $\mathcal{H}$ starting at u_0 is the unique curve satisfying*

$$\pi \circ \widetilde{\gamma} = \gamma, \quad \widetilde{\gamma}'(t) \in \mathcal{H}_{\widetilde{\gamma}(t)}.$$

The same procedure applies to tangent vectors: given $v \in T_{\pi(u)}M$, its horizontal lift at $u \in \widetilde{M}$ with respect to $\mathcal{H}$ is the unique vector $\widetilde{v} \in \mathcal{H}_u$ such that $d\pi(\widetilde{v}) = v$.

This yields a canonical transport between fibers: for $x, y \in M$ sufficiently close so that y is not in the cut-locus of x and $u \in \widetilde{M}_x$, we denote by $\widetilde{P}_{yx}u \in \widetilde{M}_y$ the endpoint of the horizontal lift of the (unique minimizing) geodesic from x to y starting at u . This map $\widetilde{P}_{yx} : \widetilde{M}_x \rightarrow \widetilde{M}_y$ is a diffeomorphism between the fibers. Because the fiber metric is G -invariant, $\widetilde{P}_{yx}$ acts by fiber isometries,

$$d_{\widetilde{M}_y}(\widetilde{P}_{yx}u_1, \widetilde{P}_{yx}u_2) = d_{\widetilde{M}_x}(u_1, u_2), \quad u_1, u_2 \in \widetilde{M}_x,$$

and in particular, for $u \in \widetilde{M}_x$, $q \in \widetilde{M}_y$,

$$d_{\widetilde{M}_y}(\widetilde{P}_{yx}u, q) = d_{\widetilde{M}_x}(\widetilde{P}_{xy}\widetilde{P}_{yx}u, \widetilde{P}_{xy}q) = d_{\widetilde{M}_x}(u, \widetilde{P}_{xy}q). \quad (3.5.2)$$

Definition 3.20 (Totally geodesic fibers). *The fibers of a Riemannian submersion are said to be totally geodesic if any geodesic in a fiber (for the induced metric) is also a geodesic in $\widetilde{M}$.*

In our setting, the fibers are totally geodesic for the Riemannian submersion $\pi : \widetilde{M} \rightarrow M$ with the Kaluza–Klein bundle metric (3.5.1) (see e.g. [82, Theorem 1]). Note that since (M, g) is geodesically complete by assumption, then so is $(\widetilde{M}, \widetilde{g})$. For totally geodesic fibers, the horizontal Laplacian coincides with the Bochner horizontal Laplacian in sub-Riemannian

geometry with the decomposition $\Delta_{\widetilde{M}} := \Delta^{\mathcal{H}} + \Delta^{\mathcal{V}}$ into horizontal and vertical parts [7, Section 4.1]. We denote by $X^{\mathcal{H}}$ and $X^{\mathcal{V}}$ the horizontal and vertical parts of a vector field $X \in \Gamma(T\widetilde{M})$ and by $\widetilde{\nabla}$ the Levi-Civita connection on $\widetilde{M}$.

Definition 3.21 (Horizontal Laplacian). *Given a smooth function $\phi \in C^\infty(\widetilde{M})$, the horizontal Laplacian $\Delta^{\mathcal{H}}$ is the trace of the horizontal Hessian; i.e.,*

$$\Delta^{\mathcal{H}}\phi := \text{Tr} \left(\widetilde{g}^{-1} \otimes \text{Hess}^{\mathcal{H}}\phi \right) := \widetilde{g}^{ij} \text{Hess}_{ij}^{\mathcal{H}}\phi,$$

where the $\widetilde{g}^{ij}$'s are the components of the inverse metric on $\widetilde{M}$, and the horizontal part of the Hessian is the bilinear form given by

$$\forall X, Y \in \Gamma(T\widetilde{M}), \quad \text{Hess}^{\mathcal{H}}\phi(X, Y) := \left(\widetilde{\nabla}_X \left(\widetilde{\nabla}_Y \phi \right)^{\mathcal{H}} \right) \cdot Y^{\mathcal{H}}.$$

By construction, since the distributions $\mathcal{H}$ and $\mathcal{V}$ are orthogonal to each other, the trace of the horizontal Hessian can be taken with respect to the inverse of the horizontal metric $\widetilde{g}^{\mathcal{H}} := \pi^*g$. In local orthonormal frames $(Z_i)_{1 \leq i \leq m}$ of $\mathcal{H}$ and $(F_j)_{1 \leq j \leq m'}$ of $\mathcal{V}$, the horizontal Laplacian can be rewritten as

$$\Delta^{\mathcal{H}} = \sum_{i=1}^m \left(Z_i^2 - \left(\widetilde{\nabla}_{Z_i} Z_i \right)^{\mathcal{H}} \right) - \sum_{j=1}^{m'} \left(\widetilde{\nabla}_{F_j} F_j \right)^{\mathcal{H}}.$$

In our setting, since the fibers are totally geodesic, the last term vanishes and

$$\Delta^{\mathcal{H}} = \sum_{i=1}^m \left(Z_i^2 - \left(\widetilde{\nabla}_{Z_i} Z_i \right)^{\mathcal{H}} \right).$$

3.5.2 Hydrodynamic limit of the lifted SEP

Random neighborhood graphs on the base manifold M can be lifted in a consistent manner to the total manifold $\widetilde{M}$ via the submersion.

Definition 3.22 (Lifted random neighborhood graph (lifted RNG)). *Conditionally on the PPP $\Lambda_N = \sum_i \delta_{X^i}$ of intensity $N d\mu$ in Definition 3.1, for each i , let $(\Lambda')_i^{N'}$ be an independent PPP on the fiber $\widetilde{M}_{X^i}$ with intensity $N' d\text{Vol}_{\widetilde{M}_{X^i}}$; the fiber volume is induced by $g^{\mathcal{V}}$. The vertices $V_{N, N'}$ consist of the points (X^i, X_a^i) where $X^i \in \Lambda^N$ and $X_a^i \in (\Lambda')_i^{N'}$. An undirected edge is present between two vertices whenever they are close on the base manifold and in the fibers; i.e., $d_M(X^i, X^j) \leq h_N$ and $d_{\widetilde{M}_{X^j}}(\widetilde{P}_{X^j X^i} X_a^i, X_b^j) \leq h_{N'}$ for some bandwidth parameters $h_N, h_{N'}$. The sequence of lifted random neighborhood graphs (lifted RNG) is given by $(G_{N, N'} := (V_{N, N'}, E_{N, N'}))_{N, N' \geq 1}$, where $E_{N, N'}$ denotes the set of undirected edges.*

Following Gao [41], we consider weights mimicking the ones of the SEP on the base manifold M given by (3.1.2). We take into account the additional geometrical information of the total manifold $\widetilde{M}$. Given a bounded continuous function $k_{\widetilde{M}} : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ with compact support in $[0, 1) \times [0, 1)$ and $k_{\widetilde{M}}(0, 0) > 0$, which makes the moments $C_{k_{\widetilde{M}}, 0}$ and $C_{k_{\widetilde{M}}, 2}$ defined in Lemma 3.31 positive, we set the following *weights*,

$$\begin{aligned} \widetilde{W}_{N, N'; \alpha}((X^i, X_a^i), (X^j, X_b^j)) \\ := \frac{1}{NN'} \frac{h_N^{-m} (h'_{N'})^{-m'} k_{\widetilde{M}} \left(\frac{d_M(X^i, X^j)}{h_N}, \frac{d_{\widetilde{M}_{X^j}}(\widetilde{P}_{X^j X^i} X_a^i, X_b^j)}{h'_{N'}} \right)}{\left(\widehat{k}_{\widetilde{M}}(X^i, X_a^i) \widehat{k}_{\widetilde{M}}(X^j, X_b^j) \right)^\alpha} \end{aligned} \quad (3.5.3)$$

with

$$\begin{aligned} \widehat{k}_{\widetilde{M}}(x, u) &:= \frac{1}{N} \frac{1}{N'} \sum_{(y, q) \in (\Lambda^N, (\Lambda')^{N'})} \left[h_N^{-m} (h'_{N'})^{-m'} \right. \\ &\quad \times k_{\widetilde{M}} \left(\frac{d_M(x, y)}{h_N}, \frac{d_{\widetilde{M}_y}(\widetilde{P}_{yx} u, q)}{h'_{N'}} \right) \Big]. \end{aligned}$$

For the SEP on the sequence $(G_{N, N'})_{N, N' \geq 1}$ of lifted RNG, we consider the weights given by (3.5.3). These weights are in fact *symmetric* in light of (3.5.2) due to the fiber isometry property. The existence of the SEP can be adapted from Section 3.2 to this setting; the condition $k_{\widetilde{M}}(0, 0) > 0$ ensures positivity of the self-point and the weights are uniformly bounded, and the proof follows the same stochastic completeness argument applied successively to the base and fiber point processes.

Remark 3.23. *The two bandwidth parameters h_N and $h'_{N'}$ tune the horizontal and vertical parts respectively. In the horizontal diffusive scaling, $h'_{N'}$ goes to zero faster than h_N when both $(N, N') \rightarrow +\infty$ in order to thermalize in the fiber while having diffusion only along the base manifold.*

Definition 3.24 (Lifted weighted horizontal Laplacian). *Given a smooth function $\phi \in C^\infty(\widetilde{M})$, the lifted α -weighted horizontal Laplacian with respect to the Gibbs reference measure $d\mu = e^{-U} d\text{Vol}_M$ on M is the reversible operator with respect to $e^{-\alpha \widetilde{U}} d\text{Vol}_{\widetilde{M}}$; that is, the operator satisfying*

$$\forall \varphi \in C_c^\infty(\widetilde{M}), \quad \int_{\widetilde{M}} \varphi \Delta_\alpha^\mathcal{H} \phi e^{-\alpha \widetilde{U}} d\text{Vol}_{\widetilde{M}} = - \int_{\widetilde{M}} (\widetilde{\nabla} \varphi)^\mathcal{H} \cdot (\widetilde{\nabla} \phi)^\mathcal{H} e^{-\alpha \widetilde{U}} d\text{Vol}_{\widetilde{M}},$$

which is given by

$$\Delta_\alpha^\mathcal{H} := e^{\alpha \widetilde{U}} \widetilde{\nabla} \cdot \left(e^{-\alpha \widetilde{U}} \widetilde{\nabla}^\mathcal{H} \right) = \Delta^\mathcal{H} - \alpha \widetilde{\nabla} \widetilde{U} \cdot \widetilde{\nabla}.$$

Analogously to (3.1.3), we set

$$\tilde{\mathcal{L}}_\alpha := C_{k_{\sim};\alpha} e^{(2\alpha-1)\tilde{U}} \Delta_{2(1-\alpha)}^{\mathcal{H}}, \quad C_{k_{\sim};\alpha} := \frac{C_{k_{\sim},2}}{2(C_{k_{\sim},0})^{2\alpha}}$$

and assume

(A3') The closure of $\tilde{\mathcal{L}}_\alpha$ on $C_c^\infty(\tilde{M})$ generates a conservative Markov semigroup and $C_c^\infty(\tilde{M})$ is a core for its generator.

Theorem 3.25 (Quenched hydrodynamic limit of the lifted SEP). *Let $(\tilde{G}_{N,N'})_{N,N' = (\tilde{V}_{N,N'}, \tilde{E}_{N,N'})_{N,N' \geq 1}}$ be a sequence of lifted RNG on $\tilde{M}$ in the sense of Definition 3.22 with symmetric weights $(\tilde{W}_{N,N';\alpha})_{N,N' \geq 1}$ given by (3.5.3) for a fixed $\alpha \geq 0$. Let $\tilde{\eta}_0$ be an initial configuration of the SEP on $G_{N,N'}$ with associated initial empirical measure $\widehat{\pi}_0^{N,N'}$ that converges weakly to $\tilde{\rho}_0 e^{-\tilde{U}} d\text{Vol}_{\tilde{M}}$ in probability; that is,*

$$\forall \varphi \in C_0(\tilde{M}), \quad \left\langle \varphi, d\widehat{\pi}_0^{N,N'} \right\rangle \xrightarrow{\text{pr.}} \left\langle \varphi, \tilde{\rho}_0 e^{-\tilde{U}} d\text{Vol}_{\tilde{M}} \right\rangle \quad \text{as } (N, N') \rightarrow +\infty,$$

where $\tilde{U} := U \circ \pi$ is the lift of U and

$$\widehat{\pi}_0^{N,N'} := \frac{1}{N} \frac{1}{N'} \sum_{(X^i, X_a^i) \in V_{N,N'}} \eta_0(X^i, X_a^i) \delta_{X^i, X_a^i}.$$

Suppose that the bandwidth parameters $(h_N, h'_{N'})_{N,N' \geq 1}$ are chosen in such a way that both

$$h_N \rightarrow 0, \quad h'_{N'} \rightarrow 0, \quad h'_{N'}/h_N \rightarrow 0$$

and that, in the joint limit $(N, N') \rightarrow (+\infty, +\infty)$,

$$\frac{N h_M^{m+2}}{\log(NN')} \rightarrow +\infty, \quad \frac{NN' h_N^{m+2} (h'_{N'})^{m'}}{\log(NN')} \rightarrow +\infty.$$

Then for almost all realizations of the sequence of graphs, the trajectory of the empirical measure $t \mapsto \widehat{\pi}_t^{N,N'}$ associated to the configuration $\tilde{\eta}_t^{N,N'}$ that evolves according to the rescaled SEP with generator $h_N^{-2} L_{\tilde{G}_{N,N'}}^{(\text{SEP})}$ converges in probability in the Skorohod topology to a deterministic trajectory $t \mapsto \tilde{\rho}_t e^{-\tilde{U}} d\text{Vol}_{\tilde{M}}$ whose density weakly solves

$$\partial_t \tilde{\rho}_t = C_{k_{\sim};\alpha} e^{(2\alpha-1)\tilde{U}} \Delta_{2(1-\alpha)}^H \tilde{\rho}_t.$$

Remark 3.26. Note that when the initial profile $\tilde{\rho}_0$ on the total manifold $\tilde{M}$ is π -related to some initial profile ρ_0 on the base manifold M ($\tilde{\rho}_0 = \rho_0 \circ \pi$), then the evolution of the latter is given by

$$\partial_t \rho_t = C_{k_{\tilde{M}}; \alpha} e^{(2\alpha-1)\tilde{U}} \Delta_{2(1-\alpha)} \rho_t.$$

This follows by the π -relation between the horizontal Laplacian and the Laplacian of the underlying base manifold; for all smooth functions $\phi \in C^\infty(M)$, it holds

$$\Delta^{\mathcal{H}}(\phi \circ \pi) = (\Delta_M \phi) \circ \pi.$$

This identity can be found, for instance, in the monographs of Hsu [57, Proposition 3.1.2] for the orthonormal frame bundle and of Baudoin–Demni–Wang [7, Proposition 4.2.7] for submersions with totally geodesic fibers. For a probabilistic proof, we refer to [64, Proposition 3.13].

Again, the hydrodynamic limit is a consequence of duality between the SEP and RW. Gao obtained the almost-sure pointwise consistency of (normalized) graphs Laplacians for neighborhood graphs lying on general fibre bundles [41, Theorems 2 & 4] and obtained rates for the unit tangent bundle $UTM \rightarrow M$ [41, Theorem 9].

The proof of the RW consistency, with the same weights as in Theorem 3.25, follows the same philosophy as the proof of Theorem 3.16, but takes into account the two-level structure of the lifted point process; one first controls the fluctuations of the base PPP and then, conditionally on the realization of the base points, applies the same concentration argument to the independent PPP sampled on the corresponding fibers. The

Theorem 3.27. Let $\varphi \in C_c^\infty(\tilde{M})$ and let $(\tilde{G}_{N,N'})_{N,N' \geq 1}$ be a sequence of lifted RNG in the sense of Definition 3.22 with symmetric weights $(\tilde{W}_{N,N'; \alpha})_{N,N' \geq 1}$ given by (3.5.3), for a fixed $\alpha \geq 0$, and associated RW with diffusively rescaled generator $h_N^{-2} L_{\tilde{G}_{N,N'}}^{(RW)}$. There exist constants $\tilde{C}_{\alpha;1}, \tilde{C}_{\alpha;2}, \tilde{C}_{\alpha;3}, \tilde{C}_{\alpha;4} > 0$, depending only on $\alpha, k_{\tilde{M}}, \varphi, \mu, m$ and m' such that, for all $\varepsilon > 0$ and $h_N, h'_{N'}$ small enough,

$$\begin{aligned} \mathbb{P} \left[\sup_{(X^i, X_a^i) \in \tilde{V}_{N,N'}} \left| \left(h_N^{-2} L_{\tilde{G}_{N,N'}}^{(RW)} \varphi - C_{k_{\tilde{M}}; \alpha} e^{(2\alpha-1)\tilde{U}} \Delta_{2(1-\alpha)}^{\mathcal{H}} \varphi - R_{N,N', \alpha, \varphi} \right) (X^i, X_a^i) \right| > \varepsilon \right] \\ \leq \tilde{C}_{\alpha;1} N N' \left[\exp \left(-\tilde{C}_{\alpha;2} N h_N^{m+2} \varepsilon^2 \right) + \exp \left(-\tilde{C}_{\alpha;3} N N' h_N^{m+2} (h'_{N'})^{m'} \varepsilon^2 \right) \right] \end{aligned} \quad (3.5.4)$$

with

$$\|R_{N,N', \alpha, \varphi}\|_\infty \leq \tilde{C}_{\alpha;4} (h_N + h_N^{-2} (h'_{N'})^2). \quad (3.5.5)$$

Then, if both

$$h_N \rightarrow 0, \quad h'_{N'} \rightarrow 0, \quad h'_{N'}/h_N \rightarrow 0$$

and that, in the joint limit $(N, N') \rightarrow (+\infty, +\infty)$,

$$\frac{N h_M^{m+2}}{\log(NN')} \rightarrow +\infty, \quad \frac{NN' h_N^{m+2} (h'_{N'})^{m'}}{\log(NN')} \rightarrow +\infty,$$

it holds almost surely and uniformly in the vertices $(X^i, X_a^i)'s$,

$$\lim_{(N, N') \rightarrow +\infty} h_N^{-2} L_{\tilde{G}_{N, N'}}^{(\text{RW})} \varphi(X^i, X_a^i) = C_{k_{\sim}; \alpha} e^{(2\alpha-1)\tilde{U}(X^i, X_a^i)} \Delta_{2(1-\alpha)}^{\mathcal{H}} \varphi(X^i, X_a^i).$$

3.6 Appendix

In this appendix, we gather several computational lemmas characterizing the limiting operator associated to the rescaled RW. The limits of the integral operators can be found in Hein–Audibert–von Luxburg [56, Proposition 22 & Theorem 25] and Gao [41, Lemmas 13, 14, 15 & Theorem 4] among others.

Lemma 3.28. *Denote by $S_{m-1} := 2\pi^{m/2}/\Gamma(m/2)$ the surface of the $(m-1)$ -dimensional unit sphere in $\mathbb{R}^m$ and by $B_m(0, 1)$ its unit ball. The first three moments of k are given by*

$$\begin{aligned} M_0^k &:= \int_{B_m(0, 1)} k(|v|) dv = S_{m-1} \int_0^1 k(r) r^{m-1} dr := C_{k, 0}, \\ M_1^k &:= \int_{B_m(0, 1)} k(|v|) v^i dv = 0, \\ M_2^k &:= \int_{B_m(0, 1)} k(|v|) v^i v^j dv = \left(\frac{S_{m-1}}{m} \int_0^1 k(r) r^{m+1} dr \right) \delta^{ij} := C_{k, 2} \delta^{ij}, \\ M_3^k &:= \int_{B_m(0, 1)} k(|v|) v^i v^j v^l dv = 0. \end{aligned}$$

Proof. Regarding the odd moments, the reflection $v \mapsto -v$ changes the sign of the integrated quantity; its integral must be zero. For the second moment, reflecting a single coordinate, we deduce that $\int k(|v|) v^i v^j dv \propto \delta^{ij}$. The proportionality constant $C_{k, 2}$ is then obtained noting that

$$\text{Tr}(C_{k, 2}(\delta^{ij})_{ij}) = C_{k, 2} m = \sum_{i=1}^m \int_{B_m(0, 1)} k(|v|) (v^i)^2 dv = S_{m-1} \int_0^1 k(r) r^{m+1} dr.$$

□

Lemma 3.29. *Let $\varphi \in C_c^\infty(M)$ be a test function. Then uniformly in $x \in (\text{Supp } \varphi)_{h_1}$,*

$$\int_M h^{-m} k \left(\frac{d_M(x, y)}{h} \right) d\mu(y) = C_{k, 0} e^{-U(x)} + O(h^2). \quad (3.6.1)$$

In particular, there are two constants $C, C' > 0$ such that

$$\forall x \in (\text{Supp } \varphi)_{h_1}, \quad C' \leq \int_M h^{-m} k \left(\frac{d_M(x, y)}{h} \right) d\mu(y) \leq C.$$

Proof. Going to normal coordinates about x yields

$$\begin{aligned} & \int_M h^{-m} k \left(\frac{d_M(x, y)}{h} \right) d\mu(y) \\ &= e^{-U(x)} \int_{B_{T_x M}(0, 1)} k(|v|) e^{-(U(\exp_x(hv)) - U(x))} \sqrt{\det g(\exp_x(hv))} dv \\ &= e^{-U(x)} \int_{B_m(0, 1)} k(|v|) (1 - E_i U v^i h + O(h^2)) dv \\ &= C_{k,0} e^{-U(x)} + O(h^2), \end{aligned}$$

where we used in the last equality that the odd moment vanishes by [Lemma 3.28](#). $\square$

Lemma 3.30. *Let $\varphi \in C_c^\infty(M)$ be a test function. Then uniformly in $x \in (\text{Supp } \varphi)_{h_1}$,*

$$\begin{aligned} h^{-2} \int_M \frac{h^{-m} k \left(\frac{d_M(x, y)}{h} \right) (\varphi(y) - \varphi(x))}{(q_h(x) q_h(y))^\alpha} d\mu(y) \\ = \frac{C_{k,2}}{2(C_{k,0})^{2\alpha}} e^{(2\alpha-1)U(x)} \Delta_{2(1-\alpha)} \varphi(x) + O(h^2), \end{aligned} \quad (3.6.2)$$

where

$$q_h(x) := \int_M h^{-m} k \left(\frac{d_M(x, y)}{h_N} \right) d\mu(y).$$

Proof. We expand the denominator to get

$$\begin{aligned} & h^{-2} \int_M \frac{h^{-m} k \left(\frac{d_M(x, y)}{h} \right) (\varphi(y) - \varphi(x))}{(q_h(x) q_h(y))^\alpha} d\mu(y) \\ &= h^{-2} \frac{e^{(2\alpha-1)U(x)}}{(C_{k,0})^{2\alpha}} \int_{B_{T_x M}(0, 1)} k(|v|) (\varphi(\exp_x(hv)) - \varphi(x)) \\ &\quad \times e^{-(1-\alpha)(U(\exp_x(hv)) - U(x))} \sqrt{\det g(\exp_x(hv))} (1 + O(h^2)) dv \end{aligned}$$

Expanding the right-hand side yields

$$\begin{aligned}
 & h^{-2} \frac{e^{(2\alpha-1)U(x)}}{(C_{k,0})^{2\alpha}} \int_{B_{T_x M}(0,1)} k(|v|) \left\{ \left(d_x \varphi(v) h + \frac{1}{2} \text{Hess}_x \varphi(v, v) h^2 \right. \right. \\
 & \quad \left. \left. + \frac{1}{6} \nabla \text{Hess}_x \varphi(v, v, v) h^3 \right) \times \left(1 + (\alpha - 1) d_x U(v) h + O(h^2) \right) \right. \\
 & \quad \left. \times \left(1 - \frac{1}{6} \text{Ric}_x(v, v) h^2 \right) (1 + O(h^2)) \right\} \\
 & \quad + O(h^4) dv \\
 & = h^{-2} \frac{e^{(2\alpha-1)U(x)}}{(C_{k,0})^{2\alpha}} \int_{B_m(0,1)} k(|v|) \left\{ E_i \varphi v^i h (1 + O(h^2)) \right. \\
 & \quad + \left(\frac{1}{2} E_i E_j \varphi + (\alpha - 1) E_i U E_j \varphi \right) v^i v^j h^2 \\
 & \quad + \left(\frac{1}{6} (E_i E_j E_l \varphi - \text{Ric}_{ij} E_l \varphi) + \frac{\alpha - 1}{2} E_i U E_j E_l \varphi \right) v^i v^j v^l h^3 \left. \right\} \\
 & \quad + O(h^4) dv.
 \end{aligned}$$

The odd moments vanish by [Lemma 3.28](#) while the second moment recovers the weighted Laplacian;

$$\begin{aligned}
 \left(\frac{1}{2} E_i E_j \varphi + (\alpha - 1) E_i U E_j \varphi \right) \int_{B_m(0,1)} k(|v|) v^i v^j dv &= \frac{C_{k,2}}{2} (\Delta - 2(1 - \alpha) \nabla U \cdot \nabla) \varphi \\
 &:= \frac{C_{k,2}}{2} \Delta_{2(1-\alpha)} \varphi.
 \end{aligned}$$

□

Lemma 3.31. *The first moments of $k_{\widetilde{M}}$ are given by*

$$\begin{aligned}
M_0^{k_{\widetilde{M}}} &:= \int_{B_m(0;1)} \int_{B_{m'}(0;1)} k_{\widetilde{M}}(|v|, |w|) \, dw dv \\
&= S_{m-1} S_{m'-1} \int_0^1 \int_0^1 k_{\widetilde{M}}(r, r') r^{m-1} (r')^{m'-1} \, dr' dr := C_{k_{\widetilde{M}},0}, \\
M_{\cdot,1}^{k_{\widetilde{M}}} &:= \int_{B_{m'}(0,1)} k_{\widetilde{M}}\left(\frac{d_M(x,y)}{h}, |w|\right) w^j \, dw = 0, \\
M_{1,0}^{k_{\widetilde{M}}} &:= \int_{B_m(0,1)} \int_{B_{m'}(0,1)} k_{\widetilde{M}}(|v|, |w|) v^i \, dw dv = 0, \\
M_{2,0}^{k_{\widetilde{M}}} &:= \int_{B_m(0,1)} \int_{B_{m'}(0,1)} k_{\widetilde{M}}(|v|, |w|) v^i v^j \, dw dv, \\
&= \left(\frac{S_{m-1}}{m} S_{m'-1} \int_0^1 \int_0^1 k_{\widetilde{M}}(r, r') r^{m+1} (r')^{m'-1} \, dr' dr \right) \delta^{ij} := C_{k_{\widetilde{M}},2} \delta^{ij}.
\end{aligned}$$

Proof. The proof is similar to the one of [Lemma 3.28](#). □

Lemma 3.32. *Let $\varphi \in C_c^\infty(\widetilde{M})$ be a test function. Then uniformly in (x, u) on the (h_1, h'_1) -thickening of the support of φ given by*

$$(\text{Supp } \varphi)_{h_1, h'_1} := \left\{ (z, p) \in \widetilde{M}; \, d_M(z, \pi \circ \text{Supp } \varphi) \leq h_1, d_{\widetilde{M}_z}(p, \text{Supp } \varphi_z) \leq h'_1 \right\},$$

it holds

$$\begin{aligned}
\int_{\widetilde{M}} h^{-m} (h')^{-m'} k_{\widetilde{M}} \left(\frac{d_M(x, y)}{h}, \frac{d_{\widetilde{M}_y}(\widetilde{P}_{yx} u, q)}{h'} \right) e^{-\widetilde{U}(y, q)} \, d\text{Vol}_{\widetilde{M}}(y, q) \\
= C_{k_{\widetilde{M}},0} e^{-\widetilde{U}(x, u)} + O(h^2) + O((h')^2).
\end{aligned}$$

Proof. For h and h' small enough, the restriction of both exponential maps, namely, $\exp_x : B_{T_x M}(0, h) \simeq B_m(0, h) \rightarrow M$ on the base manifold and $\exp_{\widetilde{M}_y}^{\widetilde{P}_{yx}} : B_{T_{(y, \widetilde{P}_{yx} u)} \widetilde{M}_y}(0, h') \simeq$

$B_{m'}(0, h') \rightarrow \widetilde{M}_y$ on the fiber are well-defined diffeomorphisms, so that

$$\begin{aligned} & \int_M \int_{\widetilde{M}_y} h^{-m} (h')^{-m'} k_{\widetilde{M}} \left(\frac{d_M(x, y)}{h}, \frac{d_{\widetilde{M}_y}(\widetilde{P}_{yx}u, q)}{h'} \right) d\text{Vol}_{\widetilde{M}_y}(q) d\mu(y) \\ &= \int_{B_M(x, h)} \int_{B_{T_{(y, \widetilde{P}_{yx}u)} \widetilde{M}_y(0,1)}} h^{-m} k_{\widetilde{M}} \left(\frac{d_M(x, y)}{h}, |w| \right) \\ & \quad \times \sqrt{\det \widetilde{g}^{\mathcal{V}}(\exp_{(y, \widetilde{P}_{yx}u)}^{\widetilde{M}_y}(h'w))} dw d\mu(y) \end{aligned} \quad (3.6.3)$$

by going to normal coordinates about $(y, \widetilde{P}_{yx}u)$. As for the base manifold, the expansion of the fiber's metric is given by

$$\sqrt{\det \widetilde{g}^{\mathcal{V}} \left(\exp_{(y, \widetilde{P}_{yx}u)}^{\widetilde{M}_y}(h'w) \right)} = 1 - \frac{1}{6} \text{Ric}_{\widetilde{M}}^{\mathcal{V}}(h'w, h'w) + O((h')^3) = 1 + O((h')^2)$$

with an error uniform in both (x, u) and $(y, \widetilde{P}_{yx}u)$ by compactness of $(\text{Supp } \varphi)_{h_1, h'_1}$. We further expand on the base manifold by going to normal coordinates about x and (3.6.3) becomes

$$\begin{aligned} & \int_{B_M(x; h)} \int_{B_{m'}(0; 1)} h^{-m} k_{\widetilde{M}} \left(\frac{d_M(x, y)}{h}, |w| \right) (1 + O((h')^2)) dw d\mu(y) \\ &= e^{-U(x)} \int_{B_m(0; 1)} \int_{B_{m'}(0; 1)} k_{\widetilde{M}}(|v|, |w|) (1 - E_i U v^i h + O(h^2)) (1 + O((h')^2)) dw dv \\ &= C_{k_{\widetilde{M}}, 0} e^{-\widetilde{U}(x, u)} + O(h^2) + O((h')^2) \end{aligned}$$

by Lemma 3.31. □

Lemma 3.33. *Let $\varphi \in C_c^\infty(\widetilde{M})$ be a test function. Then locally uniformly in $(x, u) \in \widetilde{M}$,*

$$\begin{aligned} & h^{-2} \int_{\widetilde{M}} \frac{h^{-m} (h')^{-m'} k_{\widetilde{M}} \left(\frac{d_M(x, y)}{h}, \frac{d_{\widetilde{M}_y}(\widetilde{P}_{yx}u, q)}{h'} \right) (\varphi(y, q) - \varphi(x, u))}{(\widetilde{q}_{h, h'}(x, u) \widetilde{q}_{h, h'}(y, q))^\alpha} e^{-\widetilde{U}(y, q)} d\text{Vol}_{\widetilde{M}}(y, q) \\ &= C_{k_{\widetilde{M}}; \alpha} e^{(2\alpha-1)\widetilde{U}(x, u)} \Delta_{2(1-\alpha)}^{\mathcal{H}} \varphi(x, u) + O(h) + O(h^{-2}(h')^2), \end{aligned}$$

where

$$\widetilde{q}_{h, h'}(x, u) := \int_M \int_{\widetilde{M}_z} h^{-m} (h')^{-m'} k_{\widetilde{M}} \left(\frac{d_M(x, z)}{h}, \frac{d_{\widetilde{M}_z}(\widetilde{P}_{zx}u, q)}{h'} \right) d\text{Vol}_{\widetilde{M}_z}(q) d\mu(z).$$

Proof. We decompose the difference of test functions with respect to the geometric structure;

$$\varphi(y, q) - \varphi(x, u) = \left(\varphi(y, q) - \varphi(y, \tilde{P}_{yx}u) \right) + \left(\varphi(y, \tilde{P}_{yx}u) - \varphi(x, u) \right).$$

Step 1 (Expansion in the fibers). We expand the first term of the splitting in normal coordinates about $(y, \tilde{P}_{yx}u)$. We replace the denominator, which is computed in [Lemma 3.32](#), to obtain

$$\begin{aligned} & h^{-2} \int_{\tilde{M}} \frac{h^{-m}(h')^{-m'} k_{\tilde{M}} \left(\frac{d_M(x, y)}{h}, \frac{d_{\tilde{M}_y}(\tilde{P}_{yx}u, q)}{h'} \right) \left(\varphi(y, q) - \varphi(y, \tilde{P}_{yx}u) \right)}{(\tilde{q}_{h, h'}(x, u) \tilde{q}_{h, h'}(y, q))^\alpha} e^{-\tilde{U}(y, q)} \\ & \quad \times d\text{Vol}_{\tilde{M}}(y, q) \\ &= h^{-2} \frac{e^{\alpha U(x)}}{(C_{k_{\tilde{M}}, 0})^{2\alpha}} \int_{B_M(x, h)} \int_{B_{T_{(y, \tilde{P}_{yx}u)} \tilde{M}_y}(0, 1)} h^{-m} \left\{ k_{\tilde{M}} \left(\frac{d_M(x, y)}{h}, |w| \right) \right. \\ & \quad \times (1 + O(h^2) + O((h')^2)) \times \left(\varphi \left(\exp_{(y, \tilde{P}_{yx}u)}^{\tilde{M}_y}(h'w) \right) - \varphi(y, \tilde{P}_{yx}u) \right) \\ & \quad \times e^{-(1-\alpha)U(y)} \sqrt{\det \tilde{g}^\mathcal{V} \left(\exp_{(y, \tilde{P}_{yx}u)}^{\tilde{M}_y}(h'w) \right)} \Big\} dw d\text{Vol}_M(y) \end{aligned} \quad (3.6.4)$$

Denote by $(F_j)_{1 \leq j \leq m'}$ the orthonormal basis of $T_{(y, \tilde{P}_{yx}u)} \tilde{M}_y = \mathcal{V}_{(y, \tilde{P}_{yx}u)}$ so that

$$\begin{aligned} \left(\varphi \left(\exp_{(y, \tilde{P}_{yx}u)}^{\tilde{M}_y}(h'w) \right) - \varphi(y, \tilde{P}_{yx}u) \right) &= d_{(y, \tilde{P}_{yx}u)} \varphi(w) h' + O((h')^2) \\ &= F_j \varphi w^j h' + O((h')^2). \end{aligned}$$

and (3.6.4) becomes

$$\begin{aligned} & \left(h^{-2} \frac{e^{\alpha U(x)}}{(C_{k_{\tilde{M}}, 0})^{2\alpha}} \int_{B_M(0, h)} e^{-(1-\alpha)U(y)} h' F_j \varphi \right. \\ & \quad \times \left. \int_{B_{m'}(0, 1)} h^{-m} k_{\tilde{M}} \left(\frac{d_M(x, y)}{h}, |w| \right) w^j dw dy \right) + O(h^{-2}(h')^2), \end{aligned}$$

whose integrated term is zero by [Lemma 3.31](#). Thus

$$\begin{aligned} & h^{-2} \int_M \int_{\tilde{M}_y} \frac{h^{-m}(h')^{-m'} k_{\tilde{M}} \left(\frac{d_M(x, y)}{h}, \frac{d_{\tilde{M}_y}(\tilde{P}_{yx}u, q)}{h'} \right) \left(\varphi(y, q) - \varphi(y, \tilde{P}_{yx}u) \right)}{(\tilde{q}_{h, h'}(x, u) \tilde{q}_{h, h'}(y, q))^\alpha} e^{-\tilde{U}(y, q)} \\ & \quad = O(h^{-2}(h')^2). \end{aligned} \quad (3.6.5)$$

Step 2 (Expansion along the base manifold). We expand the second term of the splitting in normal coordinates about both x and $(y, \tilde{P}_{yx}u)$. As in Step 1, by replacing the denominator, we obtain

$$\begin{aligned}
 & h^{-2} \int_{\tilde{M}} \frac{h^{-m}(h')^{-m'} k_{\tilde{M}} \left(\frac{d_M(x,y)}{h}, \frac{d_{\tilde{M}_y}(\tilde{P}_{yx}u,q)}{h'} \right) \left(\varphi(y, \tilde{P}_{yx}u) - \varphi(x,u) \right)}{(\tilde{q}_{h,h'}(x,u) \tilde{q}_{h,h'}(y,q))^\alpha} e^{-\tilde{U}(y,q)} \\
 &= h^{-2} \frac{e^{(2\alpha-1)U(x)}}{(C_{k_{\tilde{M}},0})^{2\alpha}} \int_{B_M(x,h)} \int_{B_{T_{(y,\tilde{P}_{yx}u)}\tilde{M}_y}(0,1)} h^{-m} \left\{ k_{\tilde{M}} \left(\frac{d_M(x,y)}{h}, |w| \right) \right. \\
 &\quad \times \left(\varphi(y, \tilde{P}_{yx}u) - \varphi(x,u) \right) \times e^{-(1-\alpha)(U(y)-U(x))} (1 + O(h^2) + O((h')^2)) \\
 &\quad \times \sqrt{\det \tilde{g}^\nu} \left(\exp_{(y,\tilde{P}_{yx}u)}^{\tilde{M}_y}(h'w) \right) dw d\text{Vol}_M(y). \tag{3.6.6}
 \end{aligned}$$

Denote by $\gamma \in C^\infty(I; M)$ the geodesic with $\gamma(0) = x$ and $\gamma'(0) = hv$ so that $\gamma(1) = y$. Its horizontal lift $\tilde{\gamma}$ starting at (x, u) is the unique horizontal geodesic satisfying

$$\pi \circ \tilde{\gamma} = \gamma, \quad \tilde{\gamma}'(t) \in \mathcal{H}_{\tilde{\gamma}(t)}$$

with $\tilde{\gamma}(0) = (x, u)$ and $\tilde{\gamma}'(0) = h\tilde{v}$ so that $\tilde{\gamma}(1) = (y, \tilde{P}_{yx}u)$. Then

$$\left(\varphi(y, \tilde{P}_{yx}u) - \varphi(x, u) \right) = \varphi \circ \tilde{\gamma}(1) - \varphi \circ \tilde{\gamma}(0) = \frac{d}{dt} \Big|_{t=0} \varphi \circ \tilde{\gamma}(t) + \frac{1}{2} \frac{d^2}{dt^2} \Big|_{t=0} \varphi \circ \tilde{\gamma}(t) + O(h^3),$$

where

$$\frac{d}{dt} (\varphi \circ \tilde{\gamma}(t)) = d_{\tilde{\gamma}(t)} \varphi \circ \left(d\pi_{\tilde{\gamma}(t)} \right)^{-1} (\gamma'(t)) = d_{\tilde{\gamma}(t)} \varphi (\tilde{\gamma}'(t)) = \tilde{\nabla} \varphi \cdot \tilde{\gamma}'(t) = \left(\tilde{\nabla} \varphi \right)^{\mathcal{H}} \cdot \tilde{\gamma}'(t),$$

and, since $\tilde{\gamma}$ is a horizontal geodesic and the inner product is compatible with the connection $\tilde{\nabla}$,

$$\frac{d^2}{dt^2} (\varphi \circ \tilde{\gamma}(t)) = \left(\tilde{\nabla}_{\tilde{\gamma}'(t)} \left(\tilde{\nabla} \varphi \right)^{\mathcal{H}} \right) \cdot \tilde{\gamma}'(t) + \left(\tilde{\nabla} \varphi \right)^{\mathcal{H}} \cdot \tilde{\nabla}_{\tilde{\gamma}'(t)} \tilde{\gamma}'(t) = \left(\tilde{\nabla}_{\tilde{\gamma}'(t)} \left(\tilde{\nabla} \varphi \right)^{\mathcal{H}} \right) \cdot \tilde{\gamma}'(t).$$

The expansion is thus given by

$$\left(\varphi(y, \tilde{P}_{yx}u) - \varphi(x, u) \right) = d_{(x,u)} \varphi(\tilde{v}) h + \frac{1}{2} \text{Hess}_{(x,u)}^{\mathcal{H}} \varphi(\tilde{v}, \tilde{v}) h^2 + O(h^3). \tag{3.6.7}$$

We write $v = v^i E_i$ with orthonormal basis $(E_i)_{1 \leq i \leq m}$ of $T_x M$ and horizontally lift it by linearity to $\tilde{v} = v^i \tilde{E}_i$ with the orthonormal basis $(\tilde{E}_i)_{1 \leq i \leq m}$ of $\mathcal{H}_{(x,u)}$; the above expansion

then becomes

$$\begin{aligned}
 \left(\varphi \left(y, \tilde{P}_{yx} u \right) - \varphi(x, u) \right) &= \tilde{E}_i \varphi v^i h + \frac{1}{2} \left(\left(\tilde{\nabla}_{\tilde{E}_i} \left(\tilde{\nabla} \varphi \right)^{\mathcal{H}} \right) \cdot \tilde{E}_j \right) v^i v^j h^2 + O(h^3) \\
 &= \tilde{E}_i \varphi v^i h + \frac{1}{2} \left(\tilde{E}_i \tilde{E}_j - \left(\tilde{\nabla}_{\tilde{E}_i} \tilde{E}_j \right)^{\mathcal{H}} \right) \varphi v^i v^j h^2 + O(h^3).
 \end{aligned} \tag{3.6.8}$$

Substituting (3.6.8) into (3.6.6), we obtain

$$\begin{aligned}
 &h^{-2} \frac{e^{(2\alpha-1)U(x)}}{\left(C_{k_{\tilde{M}},0} \right)^{2\alpha}} \int_{B_m(0,1)} \int_{B_{m'}(0,1)} k_{\tilde{M}}(|v|, |w|) \left\{ \left(\tilde{E}_i \varphi v^i h \right. \right. \\
 &\quad \left. \left. + \frac{1}{2} \left(\tilde{E}_i \tilde{E}_j - \left(\tilde{\nabla}_{\tilde{E}_i} \tilde{E}_j \right)^{\mathcal{H}} \right) \varphi v^i v^j h^2 + O(h^3) \right) \times (1 - (1 - \alpha) E_i U v^i h + O(h^2)) \right. \\
 &\quad \left. \times (1 + O(h^2) + O((h')^2)) (1 + O((h')^2)) (1 + O(h^2)) \right\} dw dv \\
 &= h^{-2} \frac{e^{(2\alpha-1)U(x)}}{\left(C_{k_{\tilde{M}},0} \right)^{2\alpha}} \int_{B_m(0,1)} \int_{B_{m'}(0,1)} k_{\tilde{M}}(|v|, |w|) \\
 &\quad \times \left\{ \tilde{E}_i \varphi v^i h + \frac{1}{2} \left(\tilde{E}_i \tilde{E}_j - \left(\tilde{\nabla}_{\tilde{E}_i} \tilde{E}_j \right)^{\mathcal{H}} - 2(1 - \alpha) E_i U \tilde{E}_j \right) \varphi v^i v^j h^2 \right\} dw dv + O(h).
 \end{aligned} \tag{3.6.9}$$

The first moment in (3.6.9) vanishes, and the second yields a weighted horizontal Laplacian by Lemma 3.31;

$$\begin{aligned}
 &\frac{1}{2} \left(\tilde{E}_i \tilde{E}_j - \left(\tilde{\nabla}_{\tilde{E}_i} \tilde{E}_j \right)^{\mathcal{H}} - 2(1 - \alpha) E_i U \tilde{E}_j \right) \varphi \int_{B_m(0,1)} \int_{B_{m'}(0,1)} k_{\tilde{M}}(|v|, |w|) v^i v^j dw dv \\
 &= \frac{C_{k_{\tilde{M}},2}}{2} \left(\tilde{E}_i \tilde{E}_j - \left(\tilde{\nabla}_{\tilde{E}_i} \tilde{E}_j \right)^{\mathcal{H}} - 2(1 - \alpha) E_i U \tilde{E}_j \right) \varphi \delta^{ij} \\
 &= \frac{C_{k_{\tilde{M}},2}}{2} \sum_{i=1}^m \left(\tilde{E}_i^2 - \left(\tilde{\nabla}_{\tilde{E}_i} \tilde{E}_i \right)^{\mathcal{H}} - 2(1 - \alpha) \tilde{E}_i U \tilde{E}_i \right) \varphi \\
 &:= \frac{C_{k_{\tilde{M}},2}}{2} \Delta_{2(1-\alpha)}^{\mathcal{H}} \varphi.
 \end{aligned}$$

Thus,

$$\begin{aligned}
 & h^{-2} \int_M \int_{\tilde{M}_y} \frac{h^{-m}(h')^{-m'} k_{\tilde{M}} \left(\frac{d_M(x,y)}{h}, \frac{d_{\tilde{M}_y}(\tilde{P}_{yx}u,q)}{h'} \right) \left(\varphi(y, \tilde{P}_{yx}u) - \varphi(x,u) \right)}{(\tilde{q}_{h,h'}(x,u) \tilde{q}_{h,h'}(y,q))^\alpha} \\
 & \quad \times d\text{Vol}_{\tilde{M}_y}(q) d\mu(y) \\
 & = \frac{C_{k_{\tilde{M}},2}}{2 \left(C_{k_{\tilde{M}},0} \right)^{2\alpha}} e^{(2\alpha-1)\tilde{U}(x,u)} \Delta_{2(1-\alpha)}^{\mathcal{H}} \varphi(x,u) + O(h) + O(h^{-2}(h')^2). \quad (3.6.10)
 \end{aligned}$$

We conclude by putting together (3.6.5) and (3.6.10). □

CHAPTER 4

Concentration and Expectation for Empirical Measures on Manifolds

This chapter is the result of an ongoing work “Quantitative bounds on expectations and concentration inequalities of empirical measures on complete manifolds”. In preparation (2026+).

4.1 Introduction

The goal of this chapter is to obtain quantitative estimates for the distance between the empirical measure of a finite number of independent samples and its underlying common law on manifolds. We are interested both in concentration inequalities and bounds in expectation, and we measure the error in Wasserstein distance.

This question is classical in Euclidean spaces and a large number of works have identified the optimal or near-optimal rates of convergence, depending on the dimension, the Wasserstein distance under consideration, and the integrability of the underlying probability measure. As the literature is vast on this topic (see [25, 42, 76] among others), we choose to highlight the works of Fournier–Guillin [40, 39] giving a precise picture of the Euclidean case under moment assumptions and the seminal work of Bolley–Guillin–Villani [11] for concentration inequalities.

Beyond Euclidean spaces, Boissard–Le Gouic [9, 10] developed a general metric-space approach based on covering numbers, which was then subsequently improved on compact spaces by Weed–Bach [111]. Regarding concentration inequalities, the recent preprint of Dedecker–Fischer–Michel [24] improves on [11]. The setting of these works is the one of metric spaces with polynomial volume growth, leaving Riemannian manifolds with negative Ricci lower bound aside.

There is also a substantial literature on empirical measures on Riemannian manifolds. In particular, Feng-Yu Wang and coauthors have obtained convergence rates and limit theorems for empirical measures associated with diffusion processes on manifolds, including compact manifolds, non-compact manifolds, manifolds with boundary, and subordinated diffusions [108, 109, 110]. These works use a semigroup approach.

The perspective which we adopt here is different but complementary. We first consider empirical measures of independent samples on general metric measure spaces, and then apply the resulting estimates to a noninteracting particle system coupled to a mean-field interacting particle system in the setting of a Riemannian manifold with Ricci curvature lower bound. Moreover, only linear diffusions on manifolds were treated in the above results while our analysis allows for interactions and covers the McKean–Vlasov particle system approximating an aggregation-diffusion as in Bolley–Guillin–Villani [11].

The chapter is organized as follows:

1. In [Section 4.2](#), we recall transportation inequalities and state our main results, namely bounds in expectation and respectively concentration inequalities for the empirical measure of a finite number of independent samples from a common law satisfying a moment condition or respectively a transportation-cost inequality, and the application to aggregation-diffusion equation from McKean–Vlasov particle system.
2. In [Section 4.3](#), we give the proofs of bounds in expectation and concentration inequalities for independent samples.
3. In [Section 4.4](#), we study the McKean–Vlasov interacting particle system by introducing a proxy noninteracting particle system whose mean-field interaction is replaced by the limiting law that evolves according to the aggregation-diffusion equation. We then couple this particle system with the interacting one by parallel coupling on the Riemannian manifold.

4.2 Definitions and main results

4.2.1 Wasserstein distance and relative entropy

We first recall the notions of Wasserstein distance, relative entropy, and transport-entropy inequalities; we refer to Villani [105] for a detailed account. Let (E, d) be a Polish metric space. For $p \geq 1$, we denote by $\mathcal{P}_p(E)$ the set of probability measures on E with finite p -th moment.

Definition 4.1 (Wasserstein distance). *Let $\mu, \nu \in \mathcal{P}_p(E)$. The p -Wasserstein distance*

between μ and ν is defined by

$$W_p(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{E \times E} d^p(x, y) \, d\pi(x, y) \right)^{1/p},$$

where $\Pi(\mu, \nu)$ is the set of probability measures on $E \times E$ with first marginal μ and second marginal ν .

In this paper, $W_p(\hat{\nu}^N, \nu)$ measures how far the empirical measure

$$\hat{\nu}^N := \frac{1}{N} \sum_{i=1}^N \delta_{X^i}, \quad X^1, \dots, X^N \sim \nu \text{ i.i.d.},$$

is from its common law ν . Large deviations for empirical measures are governed by relative entropy. If $\mu, \nu \in \mathcal{P}(E)$, the relative entropy of μ with respect to ν is

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

Sanov's theorem says, roughly speaking, that deviations of $\hat{\nu}^N$ from ν have probability of order

$$\mathbb{P}[\text{dist}(\hat{\nu}^N, \nu) > \varepsilon] \approx \exp \left(-N \inf_{\text{dist}(\mu, \nu) > \varepsilon} H(\mu|\nu) \right),$$

for distances compatible with weak convergence. In particular, to obtain concentration estimates in Wasserstein distance, one needs lower bounds on $H(\mu|\nu)$ in terms of $W_p(\mu, \nu)$. This is precisely the role of transportation-cost inequalities.

Definition 4.2 (Talagrand inequality). *Let $p \geq 1$ and $\lambda > 0$. We say that $\nu \in \mathcal{P}_p(E)$ satisfies the transport-cost inequality, or Talagrand inequality, $T_p(\lambda)$ if*

$$W_p(\mu, \nu) \leq \sqrt{\frac{2}{\lambda} H(\mu|\nu)} \quad (4.2.1)$$

for every probability measure μ on E .

Equivalently, $T_p(\lambda)$ gives the entropy lower bound

$$H(\mu|\nu) \geq \frac{\lambda}{2} W_p^2(\mu, \nu).$$

Inserted into the Sanov rate function, this suggests the Gaussian concentration scale

$$\mathbb{P}[W_p(\hat{\nu}^N, \nu) > \varepsilon] \lesssim \exp \left(-\frac{\lambda}{2} N \varepsilon^2 \right), \quad (4.2.2)$$

up to the metric-entropy corrections needed to make this estimate non-asymptotic.

The condition $T_p(\lambda)$ becomes stronger as p increases, since $W_p \geq W_q$ whenever $p \geq q$ on probability measures with finite moments. In particular, T_2 implies T_1 . On a Riemannian manifold, if the probability measure $\nu = e^{-U} d\text{Vol}$ with C^2 potential is such that $\text{Ric} + \text{Hess } U \geq K$ for some $K > 0$, then ν satisfies $T_2(K)$. T_1 is equivalent to the existence of a square-exponential moment: there exist $o \in E$ and $\alpha > 0$ such that

$$\int_E e^{\alpha d^2(o,x)} d\nu(x) < +\infty.$$

The inequality $T_1(\lambda)$ can also be read through its dual concentration form. Namely, if $f : E \rightarrow \mathbb{R}$ is 1-Lipschitz, then

$$\int_E \exp\left(t\left(f - \int_E f d\nu\right)\right) d\nu \leq \exp\left(\frac{t^2}{2\lambda}\right), \quad t \geq 0.$$

Therefore, for independent random variables $X^1, \dots, X^N$ with common law ν ,

$$\sup_{\text{Lip}(f) \leq 1} \mathbb{P}\left[\frac{1}{N} \sum_{i=1}^N f(X^i) - \int_E f d\nu > \varepsilon\right] \leq \exp\left(-\frac{\lambda}{2} N \varepsilon^2\right).$$

By Kantorovich duality,

$$W_1(\widehat{\nu}^N, \nu) = \sup_{\text{Lip}(f) \leq 1} \left\{ \int_E f d\widehat{\nu}^N - \int_E f d\nu \right\}.$$

Thus, the desired concentration estimate (4.2.2) for W_1 is not an immediate consequence of the preceding bound. Indeed, the latter controls each fixed 1-Lipschitz function separately, whereas W_1 requires controlling the supremum over all such functions inside the probability:

$$\sup_{\text{Lip}(f) \leq 1} \mathbb{P}(A_f) \quad \text{does not control} \quad \mathbb{P}\left(\sup_{\text{Lip}(f) \leq 1} A_f\right).$$

Quantitatively, this passage from pointwise concentration to concentration of the supremum is achieved by combining the transport-entropy inequality with a covering argument for the space of probability measures. The price is a small loss in the constant: the final non-asymptotic estimate holds with any $\lambda' < \lambda$, once N is large enough in terms of the deviation scale.

4.2.2 Geometric setting

On a metric space with finite diameter, Broissard–Le Gouic [10] obtained a quantitative bound on the expected value of the Wasserstein distance between the empirical and the sampling measure on bounded metric spaces in terms of covering numbers.

Definition 4.3 (Covering and packing numbers). *Let (E, d) be a metric space. The covering number $N(S, \delta)$ of a set $S \subset E$ denotes the minimal number of balls needed to cover S by balls of radius $\delta > 0$ with respect to the metric d . The packing number $P(S, \delta)$ is the maximum cardinality of any δ -packing, namely, any collection of points in S of distance at least δ from each other.*

Theorem 4.4 ([10, Theorem 1.1]). *Let (E, d) be a metric space with finite diameter D and ν be a probability measure on E . Choose $\varepsilon > 0$ such that $N(E, \varepsilon) < +\infty$. Then*

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq C \left(\varepsilon + N^{-\frac{1}{2p}} \int_{\varepsilon}^{D/4} N(E, \delta)^{\frac{1}{2p}} d\delta \right), \quad (4.2.3)$$

with $C \leq 64/3$.

For an unbounded space metric space (E, d) , the covering number $N(E, \delta)$ is infinite for every $\delta > 0$. In that setting, the estimate (4.2.3) can only be applied to the probability measure truncated on some possibly large ball. Most of the samples lie in a large compact under some moment assumptions. A typical way to estimate the covering number of a large ball is to control its volume growth. Such a control holds on metric measure spaces that satisfy the measure contraction property due to Ohta [85].

Definition 4.5 (Measure contraction property). *Let $(E, d, \mathfrak{m})$ be a metric measure space. Let $e_t(\gamma) := \gamma(t)$ be the evaluation map on the set of minimal geodesics $\Gamma := \{\gamma : [0, 1] \rightarrow E : \gamma \text{ is minimal}\}$, and denote the length by $\ell(\gamma) := d(\gamma(0), \gamma(1))$. A dynamical transference plan Π is a Borel probability measure on Γ and $\mathfrak{m}_t := (e_t)_\# \Pi$ is the associated displacement interpolation. For $K \in \mathbb{R}$ and $n > 1$, the space $(E, d, \mathfrak{m})$ is said to satisfy a measure contraction property MCP(K, n) if for every $x \in E$ and measurable set $A \subseteq E$, there exists a displacement interpolation $\mathfrak{m}_t$ associated with a transference plan Π satisfying the following:*

(i) *We have $\mathfrak{m}_0 = \delta_x$ and $\mathfrak{m}_1 = \mathfrak{m}[A]^{-1} \mathfrak{m}|_A$ as measures;*

(ii) *For every $t \in [0, 1]$,*

$$d\mathfrak{m} \geq (e_t)_\# \left(t \frac{s_{K,n}(t\ell(\gamma))}{s_{K,n}(\ell(\gamma))} \mathfrak{m}[A] d\Pi(\gamma) \right)$$

holds as measures with the convention $0/0 := 1$, where

$$s_{K,n}(t) := \begin{cases} \sin^{n-1} \left(\sqrt{\frac{K}{n-1}} t \right) & \text{if } K > 0; \\ t^{n-1} & \text{if } K = 0; \\ \sinh^{n-1} \left(\sqrt{\frac{|K|}{n-1}} t \right) & \text{if } K < 0. \end{cases}$$

The Bonnet–Myers theorem holds for metric measure spaces satisfying the measure contraction property [85, Theorem 4.3]; if $(E, d, \mathfrak{m})$ satisfies $\text{MCP}(K, n)$ for some $K > 0$ and $n > 1$, then

$$\text{diam}(E) \leq \pi \sqrt{\frac{n-1}{K}}.$$

In the unbounded setting, we will thus restrict ourselves to $K \leq 0$.

Remark 4.6. *The measure contraction property is weaker notion than the synthetic notion of curvature-dimension condition $\text{CD}(K, n)$ due to Sturm [96] and Lott–Villani [77]; we refer to the monograph of Villani [105] for an account on the subject. More precisely, if the weighted m -dimensional Riemannian manifold $(M, g, \mathfrak{m})$ with reference measure $\mathfrak{m} = e^{-U} d\text{Vol}$ satisfies $\text{CD}(K, n)$ with $n \geq m \geq 2$,*

$$\text{Ric}_{K,n} := \text{Ric} + \text{Hess } U - \frac{\nabla U \otimes \nabla U}{n-m} \geq K$$

with the convention $\text{Ric}_{K,m} := \text{Ric}$ in the case $m = n$, then it satisfies $\text{MCP}(K, n)$ [85, Theorem 3.2]

Metric measure spaces satisfying the measure contraction property also satisfy the Bishop–Gromov volume comparison theorem.

Theorem 4.7 (Bishop–Gromov [85, Theorem 5.1]). *Let $(E, d, \mathfrak{m})$ be a metric measure space satisfying $\text{MCP}(K, n)$ for some $K \in \mathbb{R}$ and $n > 1$. Then for any $x \in E$, the function*

$$r \mapsto \frac{\mathfrak{m}[B(x, r)]}{V_{K,n}(r)}$$

is monotone non-increasing; $V_{K,n}(r) := \int_0^r s_{K,n}(l) \, dl$ denotes the volume of the ball in the reference space.

As a consequence, the following simple type of Gromov’s packing lemma, which estimates the growth of the covering number of balls under the metric contraction property, holds. The proof is deferred to the appendix.

Lemma 4.8 (Covering number estimate). *Let $(E, d, \mathfrak{m})$ be a metric measure space satisfying $\text{MCP}(K, n)$ for some $K \leq 0$ and $n > 1$. Then for any $x \in E$, $0 < \delta \leq R/2$ with $R \leq \text{diam}(E)$,*

$$N(B(x, R), \delta) \leq \left(\frac{5R}{\delta} \right)^n \exp \left(2\sqrt{|K|(n-1)}R \right) \quad (4.2.4)$$

4.2.3 Upper bounds for the expected Wasserstein distance

By combining the result of Boissard–Le Gouic, [Theorem 4.4](#), on a single large ball with the covering estimate [Lemma 4.8](#) and a tail error from a finite polynomial or exponential moment, one derives simple non-optimal bounds on the expected p -Wasserstein distance between the empirical and true measure.

Proposition 4.9. *Let $(E, d, \mathfrak{m})$ be an unbounded metric measure space satisfying the measure contraction property $\text{MCP}(K, n)$ for some $K \leq 0$ and $n > 1$, and let ν be a probability measure on E . Then:*

(i) *If ν satisfies the polynomial moment condition*

$$M_q := \int_E d^q(o, x) d\nu(x) < +\infty$$

for some $q > p \geq 1$ and $o \in E$, then there exists a constant $N_0 > 0$ depending only on K, n, p, q and on M_q and a constant $C > 0$ depending only on K, n, p, q, M_q , such that for all $N \geq N_0$,

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq \begin{cases} CN^{-\frac{q-p}{2pq}} & \text{if } K = 0, n < 2p; \\ CN^{-\frac{q-p}{2pq}} (\log N)^{\frac{q-p}{q}} & \text{if } K = 0, n = 2p; \\ CN^{-\frac{q-p}{nq}} & \text{if } K = 0, n > 2p; \\ C (\log N)^{-\frac{q-p}{p}} & \text{if } K < 0. \end{cases}$$

(ii) *If ν satisfies the exponential moment condition*

$$\mathcal{E}_\alpha := \int_E e^{\alpha d(o, x)} d\nu(x) < +\infty$$

for some $\alpha > 0$ and $o \in E$, then there exist constants $N_1, C > 0$ depending only on K, n, p, α and on $\mathcal{E}_\alpha$, such that for all $N \geq N_1$,

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq \begin{cases} CN^{-\frac{1}{2p}} \log N & \text{if } K = 0, n < 2p; \\ CN^{-\frac{1}{2p}} (\log N)^2 & \text{if } K = 0, n = 2p; \\ CN^{-\frac{1}{n}} \log N & \text{if } K = 0, n > 2p; \\ CN^{-\frac{\alpha}{2p(\alpha+\kappa)}} \log N & \text{if } K < 0, n < 2p; \\ CN^{-\frac{\alpha}{2p(\alpha+\kappa)}} (\log N)^{1+\frac{\alpha}{\alpha+\kappa}} & \text{if } K < 0, n = 2p; \\ CN^{-\frac{\alpha}{n\alpha+2p\kappa}} \log N & \text{if } K < 0, n > 2p, \end{cases}$$

where we denote $\kappa := \sqrt{|K|(n-1)}$.

If instead of truncating on a single large ball we consider a sequence of annuli with increasing radii, then the rate can improve. In particular, when $n > 2p$ and $(E, d, \mathfrak{m})$ has uniform polynomial volume growth, which corresponds to $K = 0$ in our setting, then [10, Corollary 1.3] recovers the quantitative estimate

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq C_{E,n,p,M_q} N^{-\frac{1}{n}}$$

of sharp rate. But when $K < 0$ and the measure only decays polynomially, a single cutoff gives the same logarithmic rate as a truncation by annuli; the exponential growth of the volume overcomes the slower polynomial decay of the tail.

This phenomenon is similar to *quantization of measures* on non-compact Riemannian manifolds investigated by Iacobelli [58]. Quantization of measures seeks the best possible approximation of the probability measure ν among all probability measures ν^N supported in at most N points; the quantization error is given in p -Wasserstein distance as

$$e_{N,p}(\nu) := \inf \{W_p(\nu^N, \nu) ; \#(\text{Supp } \nu^N) \leq N\}.$$

Theorem 4.10 ([58, Corollary 1.6]). *Let M be a complete connected Riemannian manifold of dimension $m \geq 2$ with sectional curvature bounded from below; i.e., $\text{Sec} \geq K/(m-1)$ from some $K \leq 0$. Let $p \geq 1$ and let ν be a probability measure on M with absolutely continuous component ρ with respect to the Riemannian volume such that*

$$\int_M d^{p+\delta}(o, x) + e^{p\sqrt{\frac{|K|}{m-1}} d(o, x)} d\nu(x) < +\infty$$

for some $o \in M$ and $\delta > 0$. Then

$$\lim_{N \rightarrow \infty} N^{\frac{1}{m}} e_{N,p}(\nu) = Q_p^{1/p}([0, 1]^m) \left(\int_M \rho(x)^{\frac{m}{m+p}} d\text{Vol}(x) \right)^{\frac{m+p}{mp}}, \quad (4.2.5)$$

where $Q_p([0, 1]^m)$ denotes the quantization coefficient.

[58] provides an example of probability measure on the hyperbolic plane where all the polynomial moments are finite but no exponential moment is, and in particular the condition $\int_M e^{p\sqrt{|K|/(m-1)} d(o, x)} d\nu(x)$ coming from the sectional curvature is not satisfied. $N^{1/m} e_{N,p}(\nu)$ diverges as $N \rightarrow +\infty$ for this measure. The lower bound on the sectional curvature was then replaced by general covering growth estimates in [6].

This suggests that we can obtain a quantitative estimate

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim N^{-\frac{1}{n}}$$

whenever the measure satisfies some sufficiently high exponential moment. This is exactly the content of the following theorem.

Theorem 4.11. *Let $(E, d, \mathfrak{m})$ be an unbounded metric measure space satisfying $\text{MCP}(K, n)$ for some $K \leq 0$ and $n > 2p \geq 2$, and let ν be a probability measure on E that satisfies the exponential moment condition*

$$\mathcal{E}_\alpha := \int_E e^{\alpha d(o, x)} d\nu(x) < +\infty$$

for some $o \in E$ and

$$\alpha > \frac{2p}{n-p} \sqrt{|K|(n-1)}. \quad (4.2.6)$$

Then there exists a constant C depending only on K, n, p, α , and on $\mathcal{E}_\alpha$ such that, for every $N \geq 1$,

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq C N^{-\frac{1}{n}}. \quad (4.2.7)$$

Remark 4.12. Even though the condition (4.2.6) on the exponential moment of [Theorem 4.11](#) is stronger than the one of [Theorem 4.10](#), (4.2.7) is non-asymptotic while (4.2.5) is asymptotic. Note that a lower bound $\text{Sec} \geq K/(m-1)$ on the sectional curvature implies $\text{Ric} \geq K$ for Riemannian manifolds of dimension $m \geq 2$; then for $m > 2p$ and $K < 0$,

$$\left(\frac{2p}{m-p} \sqrt{|K|(m-1)} \right) \div \left(p \sqrt{|K|/(m-1)} \right) = 2 \frac{m-1}{m-p} \geq 2. \quad (4.2.8)$$

The factor 2 in the ratio (4.2.8) comes from the exponential $e^{2\sqrt{|K|(n-1)}R}$ in the covering estimate (4.2.4), as it will become apparent in the proof. On homogeneous spaces, one can improve the covering estimate, replacing the 2 in the exponential factor by 1. Then in that case $\frac{m-1}{m-p} \geq 1$, and for $p = 1$ and dimension $m \geq 3$, (4.2.6) becomes $\alpha > \sqrt{|K|/(m-1)}$, which is only stronger than the condition of [Theorem 4.10](#) for the inequality being strict.

Finally, the constant C in (4.2.7) depends continuously on K and the condition (4.2.6) becomes simply $\alpha > 0$ as $K \uparrow 0$. In the case $K = 0$, it is sufficient to require $M_p < +\infty$ as previously discussed.

4.2.4 Quantitative Sanov's theorem

The covering number of a metric space is also related to the covering number of probability measures supported on that space with respect to the Wasserstein distance, as the following proposition shows.

Proposition 4.13 ([11, Theorem A.1]). *Let (E, d) be a Polish metric space with finite diameter D . Then for all $\delta \in (0, D)$, the space $\mathcal{P}(E)$ of probability measure on E can be covered by $\mathcal{N}_p(E, \delta) := N(\mathcal{P}(E), \delta)$ balls of radius δ in W_p distance, with*

$$\mathcal{N}_p(E, \delta) \leq \left(\frac{8eD}{\delta} \right)^{pN(E, \frac{\delta}{2})}. \quad (4.2.9)$$

Thanks to this inequality, Bolley–Guillin–Villani derived the following quantitative version of Sanov’s theorem on bounded spaces.

Proposition 4.14 ([11, Proof of theorem 2.1 – Step 2]). *Let (E, d) be a metric space with finite diameter D , $\nu \in \mathcal{P}(E)$ and $\mathcal{A}$ a measurable subset of $\mathcal{P}(E)$. Then for all $\delta \in (0, D)$,*

$$\mathbb{P}[\hat{\nu}^N \in \mathcal{A}] \leq \left(\frac{8eD}{\delta} \right)^{pN(E, \frac{\delta}{2})} \exp \left(-N \inf_{\nu' \in \mathcal{A}_\delta} H(\nu' | \nu) \right),$$

where $\mathcal{A}_\delta$ denotes the δ -thickening in W_p distance of $\mathcal{A}$ given by

$$\mathcal{A}_\delta := \left\{ \nu' \in \mathcal{P}(E) : \exists \nu_A \in \mathcal{A}, W_p(\nu', \nu_A) \leq \delta \right\}.$$

For unbounded spaces, the diameter is infinite. But if the measure probability measure decays sufficiently fast, then most of the samples lie in some large ball for which this quantitative version of Sanov’s theorem holds. The idea is to truncate the space into a large ball for some optimal radius balancing the metric entropy contribution with the tail error.

We generalize the quantitative concentration inequality [11, Theorem 2.1] of Bolley–Guillin–Villani for independent random variables satisfying transport-cost inequalities from Euclidean spaces $\mathbb{R}^n$ to the general MCP(K, n) setting. More precisely, we have the following result.

Theorem 4.15. *Let $(E, d, \mathfrak{m})$ be an unbounded metric measure space satisfying MCP(K, n) for some $K \leq 0$ and $n > 1$. Let $1 \leq p \leq 2$ and ν be a probability measure on E that satisfies a $T_p(\lambda)$ inequality (4.2.1). Then, for every $n' > n$ and $0 < \lambda' < \lambda$, there exists a constant N_0 depending only on $K, \lambda, \lambda', n, n', p$ and on the square-exponential moment of ν , such that for all $0 < \varepsilon < 1$ and for all $N \geq N_0 \varepsilon^{-(n'+2)}$,*

$$\mathbb{P}[W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq e^{-\gamma_p \frac{\lambda'}{2} N \varepsilon^2}, \quad (4.2.10)$$

where

$$\gamma_p = \begin{cases} 1 & \text{if } 1 \leq p < 2; \\ 3 - 2\sqrt{2} & \text{if } p = 2. \end{cases}$$

We also generalize [11, Theorem 2.7] which covers the case of measure only satisfying polynomial moments.

Theorem 4.16. *Let $(E, d, \mathfrak{m})$ be an unbounded metric measure space satisfying MCP(K, n) for some $K \leq 0$ and $n > 1$. Let $q \geq 1$ and ν be a probability measure that satisfies the polynomial moment condition*

$$M_q := \int_E d^q(x_0, x) d\nu(x) < +\infty.$$

Set, for $p \in [1, q)$, the constant

$$C_{p,q} := \max \left\{ (16M_q^{1/p})^{\frac{p}{q-p}}, (2^{2p+1}M_q)^{\frac{1}{q-p}}, 1 \right\}.$$

Then:

- (i) For all $p \in [1, q/2)$, $0 < \delta < q/p - 2$, and $n' > n$, there is a constant N_0 depending only on δ, p, q, n, n' and M_q , such that

$$\mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq \varepsilon^{-q} N^{-\frac{q}{2p} + \frac{\delta}{2}}$$

for all $0 < \varepsilon < 1$ and

$$N \geq N_0 \varepsilon^{-\frac{q(n'+2p)}{q-p}} \exp \left(2C_{p,q} \sqrt{|K|(n-1)} \varepsilon^{-\frac{p}{q-p}} \right).$$

- (ii) For all $p \in [q/2, q)$, $0 < \delta < q/p - 1$, and $n' > n$, there is a constant N_0 depending only on δ, p, q, n, n' and M_q such that

$$\mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq \varepsilon^{-q} N^{1-\frac{q}{p}+\delta}$$

for all $0 < \varepsilon < 1$ and

$$N \geq N_0 \varepsilon^{-\frac{q(n'+2p)}{q-p}} \exp \left(2C_{p,q} \sqrt{|K|(n-1)} \varepsilon^{-\frac{p}{q-p}} \right).$$

Remark 4.17. The regime in [Theorem 4.15](#) is the same as in [11, Theorem 2.1], namely

$$N \geq N_0 \varepsilon^{-(n'+2)},$$

up to replacing the Euclidean dimension by any effective dimension $n' > n$ on $(E, d, \mathfrak{m})$ satisfying MCP(K, n).

In contrast, under only a polynomial moment assumption, [Theorem 4.16](#) holds in the regime

$$N \geq N_0 \varepsilon^{-\frac{q(n'+2p)}{q-p}} \exp \left(2C_{p,q} \sqrt{|K|(n-1)} \varepsilon^{-\frac{p}{q-p}} \right). \quad (4.2.11)$$

When $K = 0$, the exponential factor disappears and one recovers the exact same polynomial regime in $1/\varepsilon$, as in the Euclidean case [11, Theorem 2.7]. When $K < 0$, however, the exponential volume growth of balls forces the additional exponential factor in (4.2.11). This is consistent with the quantization phenomena discussed above: on negatively curved unbounded spaces, polynomial moments alone are in general not sufficient to compensate for exponential volume growth. In particular, as suggested by the counterexamples to the asymptotic quantization formula without suitable exponential moments, one should not expect estimates of the same strength as in the Euclidean case to hold in a purely polynomial regime in $1/\varepsilon$ under only polynomial moment assumptions.

4.2.5 Concentration for aggregation-diffusion

As an important application, we study the interacting particle system (IPS) of [11], namely McKean–Vlasov type with mean-field interacting, on Riemannian manifolds with Ricci curvature lower bound. The (IPS) is as follows:

$$dX_t^i = \sqrt{2}dB_t^M - \nabla V(X_t^i)dt - \frac{1}{N} \sum_{j=1}^N \nabla_x W(X_t^i, X_t^j)dt, \quad i = 1, \dots, N, \quad (4.2.12)$$

where dB_t^M denotes the Riemannian Brownian motion on M , V a confinement potential and W an interaction potential. The associated empirical measure

$$\hat{\mu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}$$

has limiting law μ_t whose evolution is governed by a non-linear Fokker–Planck equation, also called the aggregation-diffusion equation,

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot \left(\mu_t [\nabla V + \nabla (W * \mu_t)] \right),$$

where we define the convolution as

$$W * \mu_t(x) := \int_M W(x, y) d\mu_t(y).$$

As N becomes large, the error of replacing the pairwise interaction by a mean-field limit becomes small and we can apply our main concentration result to the empirical measure of the proxy non interacting particle system. We obtain the following result generalizing [11, Theorem 2.9].

Theorem 4.18. *Let (M, g) be a complete Riemannian manifold of dimension $m \geq 2$ satisfying $\text{Ric} \geq K$ for some $K \leq 0$. Let μ_0 be a probability measure on M with a finite square exponential moment; i.e., there exists some $\alpha_0 > 0$ and $o \in M$ such that*

$$E_{\alpha_0} := \int_M e^{\alpha_0 d^2(o, x)} d\mu_0(x).$$

Let $\{X_0^i\}_{i=1}^N$ be a N independent random variables with common law μ_0 and $\{X_t^i\}_{i=1}^N$ the solution to (4.2.12) with initial condition $\{X_0^i\}_{i=1}^N$, where V and W are smooth and satisfy Hessian lower bounds,

$$\exists \beta, \tilde{\beta} > 0 : \quad \text{Hess } V \geq \beta, \quad \forall y \in M, \text{ Hess } W(\cdot, y) \geq \tilde{\beta},$$

and gradient bounds; there exist $A_V, L_V \geq 0$ and $L_{W,1}, L_{W,2} > 0$ such that for all $x, y, z \in M$, it holds

$$|\nabla V(x)| \leq A_V + L_V d(o, x), \quad |\nabla_x W(x, y)| \leq A_W + L_{W,1}(d(o, x) + d(o, y)),$$

and for every unit vector $u \in T_x M$,

$$|(\nabla_x W(x, y) - \nabla_x W(x, z)) \cdot u| \leq L_{W,2} d(y, z).$$

Let

$$\hat{\mu}_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}$$

be the empirical measure associated to the process $\{X_t^i\}_{i=1}^N$. Then for every $T \geq 0$, there exist constants $C, C' > 0$ such that for all $m' > m$, there exists some constant N_0 such that for all $0 < \varepsilon < 1$,

$$\mathbb{P} \left[\sup_{0 \leq t \leq T} W_1(\hat{\mu}_t^N, \mu_t) > \varepsilon \right] \leq C (1 + T\varepsilon^{-2}) \exp(-C' N \varepsilon^2)$$

whenever $N \geq N_0 \varepsilon^{-(m'+2)}$.

4.3 Proofs for independent random variables

4.3.1 Expected Wasserstein distance

Before detailing the proof of [Theorem 4.11](#), we give some outline. [Theorem 4.4](#) yields bounds in expectation on spaces with finite diameter depending on their covering number. In order to obtain a sharp rate when the diameter is infinite, one cannot simply use a single truncation and compare the rate on the compact with the error term outside of that compact which depends on exponential decay of the tail. Instead, one uses an increasing sequence of truncations covering the whole space and rewrite the empirical measure as a series of empirical measure on each annuli up to some probability measure. On each of these truncations one applies [Theorem 4.4](#). The choice of the sequence of annuli depends on optimization considerations.

Proof of [Theorem 4.11](#). We decompose the space into annuli as in the proof of Boissard–Le Gouic [10, Corollary 1.3].

Set $r_0 := 0$, and

$$A_i := B(o, r_i) \setminus B(o, r_{i-1}), \quad i \geq 1 \text{ and } r_i > r_{i-1}.$$

Let $X^1, \dots, X^N$ be i.i.d. with common law ν and denote by

$$\widehat{\nu}^N := \frac{1}{N} \sum_{j=1}^N \delta_{X^j}$$

the associated empirical measure. For each $i \geq 1$, we define $N_i := N\widehat{\nu}^N(A_i)$, and whenever $N_i \neq 0$, we let

$$\widehat{\nu}_i^N := \frac{\mathbf{1}_{A_i} \widehat{\nu}^N}{\widehat{\nu}^N(A_i)}, \quad \nu_i := \frac{\mathbf{1}_{A_i} \nu}{\nu(A_i)}.$$

Notice that $N_i \in \mathbb{N}$ and $\widehat{\nu}_i^N$ is a condition empirical measure. Then there exist probability measures μ_1 and μ_2 such that

$$\widehat{\nu}^N = \sum_{i \geq 1} \lambda_i \widehat{\nu}_i^N + \lambda \mu_1, \quad \nu = \sum_{i \geq 1} \lambda_i \nu_i + \lambda \mu_2, \quad (4.3.1)$$

where

$$\lambda_i := \widehat{\nu}^N(A_i) \wedge \nu(A_i), \quad \lambda := \sum_{i \geq 1} (\widehat{\nu}^N(A_i) - \nu(A_i))_+.$$

Indeed, one can check, under the convention that whenever $\widehat{\nu}^N(A_i) = 0$ then $0 \times \widehat{\nu}_i^N := 0$, that

$$\begin{aligned} \mu_1 &:= \frac{1}{\lambda} \sum_{i \geq 1} (\widehat{\nu}^N(A_i) - \nu(A_i))_+ \widehat{\nu}_i^N, \\ \mu_2 &:= \frac{1}{\lambda} \sum_{i \geq 1} (\nu(A_i) - \widehat{\nu}^N(A_i))_+ \nu_i. \end{aligned}$$

By [10, Lemma 2.4], namely convexity of W_p^p and the triangle inequality, and the decomposition (4.3.1), we therefore have

$$W_p(\widehat{\nu}^N, \nu) \leq \sum_{i \geq 1} \lambda_i^{1/p} W_p(\widehat{\nu}_i^N, \nu_i) + \lambda^{1/p} W_p(\nu, \mu_1) + \lambda^{1/p} W_p(\nu, \mu_2). \quad (4.3.2)$$

We estimate separately the shell contribution and the two remainder terms.

Step 1 (Estimate on the remainder terms). We only treat the term involving μ_1 , the one involving μ_2 being identical. Again by [10, Lemma 2.4],

$$\lambda^{1/p} W_p(\nu, \mu_1) \leq \sum_{i \geq 1} (\widehat{\nu}^N(A_i) - \nu(A_i))_+^{1/p} W_p(\nu, \widehat{\nu}_i^N).$$

By triangle inequality, $W_p(\nu, \widehat{\nu}_i^N) \leq W_p(\nu, \delta_o) + W_p(\delta_o, \widehat{\nu}_i^N)$, and we deduce that

$$W_p(\nu, \widehat{\nu}_i^N) \leq \left(\int_E d^p(o, x) d\nu(x) \right)^{\frac{1}{p}} + \left(\frac{1}{\widehat{\nu}^N(A_i)} \int_{A_i} d^p(o, x) d\widehat{\nu}^N(x) \right)^{\frac{1}{p}} \leq M_p^{1/p} + r_i,$$

where we denote by $M_p := \int_E d^p(o, x) d\nu(x)$ the p -th moment of ν which is finite, since ν admits some exponential moment $\mathcal{E}_\alpha$. Taking expectations, we obtain

$$\mathbb{E} \left[\lambda^{1/p} W_p(\nu, \mu_1) \right] \leq \sum_{i \geq 1} \mathbb{E} \left[(\widehat{\nu}^N(A_i) - \nu(A_i))_+^{1/p} \right] (M_p^{1/p} + r_i). \quad (4.3.3)$$

By Jensen's inequality,

$$\mathbb{E} \left[(\widehat{\nu}^N(A_i) - \nu(A_i))_+^{1/p} \right] \leq \mathbb{E} \left[|\widehat{\nu}^N(A_i) - \nu(A_i)|^{1/p} \right] \leq \mathbb{E} \left[|\widehat{\nu}^N(A_i) - \nu(A_i)|^2 \right]^{\frac{1}{2p}}.$$

Since $N_i = N\widehat{\nu}^N(A_i) \sim \text{Bin}(N, \nu(A_i))$, $\text{Var} [\widehat{\nu}^N(A_i)] \leq \frac{\nu(A_i)}{N}$, and thus (4.3.3) becomes

$$\mathbb{E} \left[\lambda^{1/p} W_p(\nu, \mu_1) \right] \leq N^{-\frac{1}{2p}} \sum_{i \geq 1} \nu(A_i)^{\frac{1}{2p}} (M_p^{1/p} + r_i). \quad (4.3.4)$$

The same estimate holds with μ_2 in place of μ_1 . We now estimate the shell masses using the exponential moment. For every $i \geq 1$,

$$\nu(A_i) \leq \nu(E \setminus B(o, r_{i-1})) \leq e^{-\alpha(r_{i-1})} \int_E e^{\alpha d(o, x)} d\nu(x) = \mathcal{E}_\alpha e^{-\alpha(r_{i-1})}.$$

Therefore,

$$\sum_{i \geq 1} \nu(A_i)^{\frac{1}{2p}} (M_p^{1/p} + r_i) \leq \mathcal{E}_\alpha^{\frac{1}{2p}} \sum_{i \geq 1} e^{-\alpha \frac{r_{i-1}}{2p}} (M_p^{1/p} + r_i) := C_1, \quad (4.3.5)$$

which is finite for sequences $\{r_i\}_{i \geq 0}$ increasing sufficiently fast. We deduce from (4.3.4) that for such sequences,

$$\mathbb{E} \left[\lambda^{1/p} W_p(\nu, \mu_1) + \lambda^{1/p} W_p(\nu, \mu_2) \right] \leq 2C_1 N^{-\frac{1}{2p}}. \quad (4.3.6)$$

Step 2 (Conditional estimate on one shell). Fix $i \geq 1$ and condition on the event $\{N_i = k\}$ for some $k \geq 1$. Conditionally on this event, $\widehat{\nu}_i^N$ is the empirical measure of k independent random variables distributed according to ν_i .

Since ν_i is supported on $A_i \subset B(o, r_i)$, the uniform-in- K covering estimate (4.2.4) yields, for every $0 < \delta \leq r_i/2$,

$$N(A_i, \delta) \leq N(B(o, r_i), \delta) \leq \left(\frac{5r_i}{\delta} \right)^n \exp \left(2\sqrt{|K|(n-1)r_i} \right). \quad (4.3.7)$$

Applying [Theorem 4.4](#) with sample size k together with (4.3.7), we obtain that for every $0 < \delta < r_i/2$,

$$\begin{aligned} \mathbb{E} [W_p(\hat{\nu}_i^N, \nu_i) \mid N_i = k] &\leq \frac{64}{3} \left(\delta + k^{-\frac{1}{2p}} \int_{\delta}^{r_i/2} N(A_i, s)^{\frac{1}{2p}} ds \right) \\ &\leq \frac{64}{3} \left(\delta + \frac{5^{\frac{n}{2p}}}{\frac{n}{2p} - 1} k^{-\frac{1}{2p}} r_i^{\frac{n}{2p}} e^{\sqrt{|K|(n-1)}r_i/p} \delta^{1-\frac{n}{2p}} \right). \end{aligned} \quad (4.3.8)$$

We now choose δ to equate the two terms of the right-hand side;

$$\delta := \delta_{i,k} := \frac{5}{\left(\frac{n}{2p} - 1\right)^{2p/n} r_i} e^{\frac{2\sqrt{|K|(n-1)}r_i}{n}} k^{-\frac{1}{n}}.$$

If $\delta_{i,k} \leq r_i/2$, then (4.3.8) yields

$$\mathbb{E} [W_p(\hat{\nu}_i^N, \nu_i) \mid N_i = k] \leq C_2 r_i e^{\frac{2\sqrt{|K|(n-1)}r_i}{n}} k^{-\frac{1}{n}} \quad (4.3.9)$$

with $C_2 := \frac{640}{3} \times (\frac{n}{2p} - 1)^{-2p/n}$. Note that when $\delta_{i,k} > r_i/2$, we can estimate

$$W_p(\hat{\nu}_i^N, \nu_i) \leq \text{diam}(A_i) \leq 2r_i < 4\delta_{i,k},$$

and thus the bound (4.3.9) is valid for every $k \geq 1$.

Step 3 (Averaging the shell terms). Combining $\lambda_i \leq \hat{\nu}^N(A_i) = N_i/N$ with the estimate (4.3.9), we obtain

$$\mathbb{E} \left[\lambda_i^{1/p} W_p(\hat{\nu}_i^N, \nu_i) \mid N_i = k \right] \leq C_2 N^{-\frac{1}{p}} r_i e^{\frac{2\sqrt{|K|(n-1)}r_i}{n}} k^{\frac{1}{p} - \frac{1}{n}}.$$

By the tower property,

$$\begin{aligned} \mathbb{E} \left[\lambda_i^{1/p} W_p(\hat{\nu}_i^N, \nu_i) \right] &\leq C_2 N^{-\frac{1}{p}} r_i e^{\frac{2\sqrt{|K|(n-1)}r_i}{n}} \mathbb{E} \left[N_i^{\frac{1}{p} - \frac{1}{n}} \right] \\ &\leq C_2 N^{-\frac{1}{n}} r_i e^{\frac{2\sqrt{|K|(n-1)}r_i}{n}} \nu(A_i)^{\frac{1}{p} - \frac{1}{n}}; \end{aligned}$$

we used Jensen's inequality in the second line and $N_i = N\hat{\nu}^N(A_i)$. Summing over $i \geq 1$ and recalling that $\nu(A_i) \leq \mathcal{E}_\alpha e^{-\alpha r_{i-1}}$ with $\mathcal{E}_\alpha = \int_E e^{\alpha d(o,x)} d\nu(x)$,

$$\begin{aligned} \sum_{i \geq 1} \mathbb{E} \left[\lambda_i^{1/p} W_p(\hat{\nu}_i^N, \nu_i) \right] &\leq C_2 \mathcal{E}_\alpha^{\frac{1}{p} - \frac{1}{n}} N^{-\frac{1}{n}} \sum_{i \geq 1} r_i \exp \left(\frac{2\sqrt{|K|(n-1)}}{n} r_i - \alpha \left(\frac{1}{p} - \frac{1}{n} \right) r_{i-1} \right). \end{aligned} \quad (4.3.10)$$

Step 4 (Choice of the radii). Denoting $h_i := r_i - r_{i-1}$, the exponent in the last sum can be rewritten as

$$-\beta r_{i-1} + \frac{2\sqrt{|K|(n-1)}}{n} h_i, \quad \beta := \alpha \left(\frac{1}{p} - \frac{1}{n} \right) - \frac{2\sqrt{|K|(n-1)}}{n}.$$

In order for (4.3.10) to be summable, we therefore require $\beta > 0$, which is the condition (4.2.6).

A convenient choice of radii is $r_i = i$ so that $h_i = 1$ for $i \geq 1$, and

$$\sum_{i \geq 1} r_i \exp \left(\frac{2\sqrt{|K|(n-1)}}{n} r_i - \alpha \left(\frac{1}{p} - \frac{1}{n} \right) r_{i-1} \right) = e^{\frac{2\sqrt{|K|(n-1)}}{n}} \frac{1}{(1 - e^{-\beta})^2}.$$

Inserting this into (4.3.10), we obtain

$$\sum_{i \geq 1} \mathbb{E} \left[\lambda_i^{1/p} W_p(\hat{\nu}_i^N, \nu_i) \right] \leq C_3 N^{-\frac{1}{n}}, \quad (4.3.11)$$

where

$$C_3 := C_2 \mathcal{E}_\alpha^{\frac{1}{p} - \frac{1}{n}} e^{\frac{2\sqrt{|K|(n-1)}}{n}} \frac{1}{(1 - e^{-\beta})^2}.$$

Combining (4.3.2), (4.3.6), and (4.3.11), we obtain

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 2C_1 N^{-\frac{1}{2p}} + C_3 N^{-\frac{1}{n}},$$

where C_1 is defined in (4.3.5);

$$C_1 := \mathcal{E}_\alpha^{\frac{1}{2p}} \sum_{i \geq 1} e^{-\alpha \frac{r_{i-1}}{2p}} \left(M_p^{1/p} + r_i \right) = \mathcal{E}_\alpha^{\frac{1}{2p}} \left(\frac{M_p^{1/p}}{1 - e^{-\alpha/(2p)}} + \frac{1}{(1 - e^{-\alpha/(2p)})^2} \right).$$

Since $n > 2p$, one has $N^{-1/(2p)} \leq N^{-1/n}$ for every $N \geq 1$. Hence

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq (2C_1 + C_3) N^{-\frac{1}{n}}.$$

This concludes the proof. □

4.3.2 Concentration inequalities

Proof of Theorem 4.15. We adapt the strategy of Bolley, Guillin and Villani [11, Theorem 2.1].

Step 1 (Truncation). Denote

$$\nu_R := \frac{1_{B_R(o)}\nu}{\nu[B_R(o)]}.$$

Then we sample independent random variables $(X^i)_{1 \leq i \leq N}$ and $(Y^i)_{1 \leq i \leq N}$ according to ν and its truncation ν_R respectively in order to define

$$X_R^i = \begin{cases} X^i & \text{if } d(o, X^i) \leq R; \\ Y^i & \text{if } d(o, X^i) > R. \end{cases}$$

The p -Wasserstein distance between the measure and its truncation can be estimated as

$$W_p^p(\nu_R, \nu) \leq \mathbb{E} [d^p(X^1, Y^1) 1_{d(o, X^1) > R}] \leq 2^p \int_{E \setminus B_R(o)} d^p(o, x) d\nu(x) \leq 2^p E_\alpha R^p e^{-\alpha R^2} \quad (4.3.12)$$

for R large enough, say $R \geq R_0 := \sqrt{p/(2\alpha)}$, since $r \mapsto r^p e^{-\alpha r^2}$ is nonincreasing for $r \geq R$; recall that ν satisfies in particular $T_1(\lambda)$, and hence there is some $\alpha > 0$ such that

$$E_\alpha := \int_E e^{\alpha d^2(o, x)} d\nu(x) < +\infty.$$

We now compare the p -Wasserstein distance between the empirical measure and its truncation defined as

$$\hat{\nu}^N := \frac{1}{N} \sum_{i=1}^N \delta_{X^i}, \quad \hat{\nu}_R^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_R^i}.$$

Set

$$Z^i := 2^p d^p(o, X^i) 1_{d(o, X^i) > R} 1_{1 \leq p < 2} + d^2(X^i, X_R^i) 1_{p=2}$$

and let $\varepsilon, \theta > 0$. The exponential Chebyshev inequality yields

$$\begin{aligned} \mathbb{P} [W_p(\hat{\nu}^N, \hat{\nu}_R^N) > \varepsilon] &\leq \mathbb{P} \left[\frac{1}{N} \sum_{i=1}^N d^p(X^i, X_R^i) > \varepsilon^p \right] \\ &\leq \exp \left(-N \left\{ \theta \varepsilon^p - \log \mathbb{E} [\exp(\theta Z^1)] \right\} \right) \end{aligned} \quad (4.3.13)$$

by independence of the Z^i 's.

In the case $1 \leq p < 2$, let $0 < \alpha_1 < \alpha$. Then for R large enough, say $R \geq R_1 := (2^p \theta / \alpha_1)^{\frac{1}{2-p}}$, using the inequalities $2^p \theta r^p \leq \alpha_1 r^2$ valid for $r \geq R$ and $\log(1+x) \leq x$, we

obtain

$$\log (\mathbb{E} [\exp (\theta Z^1)]) \leq \log \left(1 + \int_{E \setminus B_R(o)} \exp (2^p \theta d^p(o, x)) \, d\nu(x) \right) \leq E_\alpha e^{-(\alpha-\alpha_1)R^2}. \quad (4.3.14)$$

Inserting (4.3.14) into (4.3.13), we deduce that

$$\mathbb{P} [W_p (\hat{\nu}^N, \hat{\nu}_R^N) > \varepsilon] \leq \exp \left(-N \left\{ \theta \varepsilon^p - E_\alpha e^{-(\alpha-\alpha_1)R^2} \right\} \right). \quad (4.3.15)$$

In the case $p = 2$, set $\theta := \alpha_1/2$ for some $0 < \alpha_1 < \alpha$. Then for R large enough, say $R \geq R_2 := \sqrt{\log(2E_\alpha)/\alpha}$, so that $\nu[B_R(o)] \geq 1 - E_\alpha e^{-\alpha R^2} \geq \frac{1}{2}$, it holds

$$\begin{aligned} \mathbb{E} [\exp (\theta Z^1)] &\leq \nu[B_R(o)] + \frac{1}{\nu[B_R(o)]} \int_{B_R(o)} \int_{E \setminus B_R(o)} e^{\frac{\alpha_1}{2} d^2(x, y)} \, d\nu(x) \, d\nu(y) \\ &\leq 1 + 2E_\alpha^2 e^{-(\alpha-\alpha_1)R^2}. \end{aligned} \quad (4.3.16)$$

Inserting (4.3.16) into (4.3.13), we deduce that

$$\mathbb{P} [W_2 (\hat{\nu}^N, \hat{\nu}_R^N) > \varepsilon] \leq \exp \left(-N \left\{ \frac{\alpha_1}{2} \varepsilon^2 - 2E_\alpha^2 e^{-(\alpha-\alpha_1)R^2} \right\} \right). \quad (4.3.17)$$

We now gather the estimates obtained so far. The triangle inequality yields

$$\mathbb{P} [W_p (\hat{\nu}^N, \nu) > \varepsilon] \leq \mathbb{P} [W_p (\hat{\nu}^N, \hat{\nu}_R^N) + W_p (\hat{\nu}_R^N, \nu_R) + W_p (\nu_R, \nu) > \varepsilon].$$

Let $\eta \in (0, 1)$ and $R \geq \max\{R_0, R_1, R_2\}$. In the case $1 \leq p < 2$, combining (4.3.12) with (4.3.15), we obtain

$$\begin{aligned} \mathbb{P} [W_p (\hat{\nu}^N, \nu) > \varepsilon] &\leq \mathbb{P} \left[W_p (\hat{\nu}_R^N, \nu_R) > \eta \varepsilon - 2E_\alpha^{1/p} R e^{-\frac{\alpha}{p} R^2} \right] \\ &\quad + \exp \left(-N \left\{ \theta (1 - \eta)^p \varepsilon^p - E_\alpha e^{-(\alpha-\alpha_1)R^2} \right\} \right). \end{aligned} \quad (4.3.18)$$

In the case $p = 2$, combining (4.3.12) with (4.3.17), we similarly obtain

$$\begin{aligned} \mathbb{P} [W_2 (\hat{\nu}^N, \nu) > \varepsilon] &\leq \mathbb{P} \left[W_2 (\hat{\nu}_R^N, \nu_R) > \eta \varepsilon - 2E_\alpha^{1/2} R e^{-\frac{\alpha}{2} R^2} \right] \\ &\quad + \exp \left(-N \left\{ \frac{\alpha_1}{2} (1 - \eta)^2 \varepsilon^2 - 2E_\alpha^2 e^{-(\alpha-\alpha_1)R^2} \right\} \right). \end{aligned} \quad (4.3.19)$$

Step 2 (Covering number estimate). Bolley–Guillin–Villani obtain the following quantitative Sanov estimate on compacts (see Step 2 in the proof of [11, Theorem 2.1]). For every measurable subset $\mathcal{A}$ of $\mathcal{P}(B_R(o))$ and $0 < \delta < R$,

$$\mathbb{P} [\hat{\nu}_R^N \in \mathcal{A}] \leq \left(\frac{16eR}{\delta} \right)^{pN(B_R(o), \frac{\delta}{2})} \exp \left(-N \inf_{\nu' \in \mathcal{A}_\delta} H(\nu' | \nu_R) \right), \quad (4.3.20)$$

where $\mathcal{A}_\delta$ denotes the δ -thickening in W_p distance of $\mathcal{A}$ given by

$$\mathcal{A}_\delta := \left\{ \nu' \in \mathcal{P}(B_R(o)) : \exists \nu_{\mathcal{A}} \in \mathcal{A}, W_p(\nu', \nu_{\mathcal{A}}) \leq \delta \right\}.$$

In light of the last two estimates (4.3.18) and (4.3.19) of Step 1, we set

$$\mathcal{A} := \left\{ \nu' \in \mathcal{P}(B_R(x_0)) : W_p(\nu', \nu_R) \geq \eta\varepsilon - 2E_\alpha^{1/p} R e^{-\frac{\alpha}{p} R^2} \right\}.$$

Now, it remains to find a lower bound to the relative entropy in (4.3.20) with respect to the truncated measure. For every $\nu' \in \mathcal{P}(B_R(o))$ and $\nu[B_R(o)] \geq 1 - E_\alpha e^{-\alpha R^2}$, it holds

$$\begin{aligned} H(\nu' | \nu_R) &= H(\nu' | \nu) + \log(\nu[B_R(o)]) \\ &\geq H(\nu' | \nu) - 2E_\alpha e^{-\alpha R^2} \\ &\geq \frac{\lambda}{2} W_p^2(\nu', \nu) - 2E_\alpha e^{-\alpha R^2}, \end{aligned}$$

where in the last inequality we used the fact that ν satisfies a $T_p(\lambda)$ inequality. Moreover, by the reverse triangle inequality and (4.3.12),

$$W_p(\nu', \nu) \geq W_p(\nu', \nu_R) - W_p(\nu_R, \nu) \geq W_p(\nu', \nu_R) - 2E_\alpha^{1/p} R e^{-\frac{\alpha}{p} R^2}.$$

Therefore,

$$H(\nu' | \nu_R) \geq \frac{\lambda}{2} \left(W_p(\nu', \nu_R) - 2E_\alpha^{1/p} R e^{-\frac{\alpha}{p} R^2} \right)^2 - 2E_\alpha e^{-\alpha R^2}.$$

Fix now $\lambda_1 \in (0, \lambda)$. Using the elementary inequality

$$(a - b)^2 \geq \frac{\lambda_1}{\lambda} a^2 - \frac{\lambda_1}{\lambda - \lambda_1} b^2, \quad (4.3.21)$$

which follows from

$$(a - b)^2 - \frac{\lambda_1}{\lambda} a^2 + \frac{\lambda_1/\lambda}{1 - \lambda_1/\lambda} b^2 = \left(1 - \frac{\lambda_1}{\lambda}\right) \times \left(a - \frac{b}{1 - \lambda_1/\lambda}\right)^2 \geq 0,$$

we deduce that

$$\begin{aligned} H(\nu' | \nu_R) &\geq \frac{\lambda_1}{2} W_p^2(\nu', \nu_R) - \frac{2\lambda\lambda_1}{\lambda - \lambda_1} E_\alpha^{2/p} R^2 e^{-\frac{2\alpha}{p} R^2} - 2E_\alpha e^{-\alpha R^2} \\ &\geq \frac{\lambda_1}{2} W_p^2(\nu', \nu_R) - C_1 R^2 e^{-\alpha R^2} \end{aligned} \quad (4.3.22)$$

with

$$C_1 := \frac{2\lambda\lambda_1}{\lambda - \lambda_1} E_\alpha^{2/p} + 2E_\alpha.$$

From (4.3.22), we have, for any $\nu \in \mathcal{A}_\delta$,

$$H(\nu'|\nu_R) \geq \frac{\lambda_1}{2} W_p^2(\nu', \nu_R) - C_1 R^2 e^{-\alpha R^2} \geq \frac{\lambda_1}{2} \xi^2 - C_1 R^2 e^{-\alpha R^2},$$

where $\xi := (\eta\varepsilon - 2E_\alpha^{1/p} R e^{-\frac{\alpha}{p} R^2} - \delta)_+$, and recalling the elementary inequality (4.3.21), we find

$$H(\nu'|\nu_R) \geq \frac{\lambda_1}{2} \xi^2 - C_1 R^2 e^{-\alpha R^2} \geq \frac{\lambda_2}{2} \varepsilon^2 - C_2 R^2 e^{-\alpha R^2} \quad (4.3.23)$$

in the regime $R \geq \max\{1, R_0, R_1, R_2\}$, for every $\lambda_2 < \lambda_1 \leq \lambda$ such that $\lambda_2 < \lambda_1(\eta_1 - \delta_1)^2$, with the constant

$$C_2 := C_1 + \frac{2\lambda_1\lambda_2 E_\alpha^{2/p}}{\lambda_1(\eta_1 - \delta_1)^2 - \lambda_2},$$

for the choice $\delta := \delta_1\varepsilon$ and $\eta := \eta_1$. While (4.3.23) is valid for $p = 2$, in that case, we instead use

$$H(\nu'|\nu_R) \geq \frac{\lambda_1}{2} \xi^2 - C_1 R^2 e^{-\alpha R^2} \geq \frac{\lambda_2}{2} \eta^2 \varepsilon^2 - C_2^{(2)} R^2 e^{-\alpha R^2} \quad (4.3.24)$$

for $\lambda_2 < \lambda_1(\eta - \delta_1)^2/\eta^2$ with the constant

$$C_2^{(2)} := C_1 + \frac{2\lambda_1\lambda_2\eta^2 E_\alpha}{\lambda_1(\eta - \delta_1)^2 - \lambda_2\eta^2}.$$

for the choice $\delta := \delta_1\varepsilon$.

We now estimate the metric entropy term in (4.3.20). Use the covering estimate (4.2.4) with $\delta/2$ instead of δ , we get

$$pN\left(B_R(o), \frac{\delta}{2}\right) \log\left(\frac{16eR}{\delta}\right) \leq p\left(\frac{10R}{\delta}\right)^n \exp\left(2\sqrt{|K|(n-1)}R\right) \log\left(\frac{16eR}{\delta}\right).$$

Let now $n_1 > n$. Since $\log r \leq \frac{r^{n_1-n}}{e(n_1-n)}$ for $r \geq 1$, we deduce that

$$pN\left(B_R(o), \frac{\delta}{2}\right) \log\left(\frac{16eR}{\delta}\right) \leq C_3 \exp\left(2\sqrt{|K|(n-1)}R\right) \left(\frac{R}{\delta}\right)^{n_1} \quad (4.3.25)$$

with

$$C_3 := \frac{p 10^n (16e)^{n_1-n}}{e(n_1-n)}.$$

We insert this estimate into (4.3.20). In the case $1 \leq p < 2$, combining (4.3.20), (4.3.23), and (4.3.25), we obtain

$$\mathbb{P} [\hat{\nu}_R^N \in \mathcal{A}] \leq \exp \left(\frac{C_3}{\delta_1^{n_1}} \exp \left(2\sqrt{|K|(n-1)} R \right) \left(\frac{R}{\varepsilon} \right)^{n_1} - N \left(\frac{\lambda_2}{2} \varepsilon^2 - C_2 R^2 e^{-\alpha R^2} \right) \right). \quad (4.3.26)$$

In the case $p = 2$, combining (4.3.20), (4.3.24), and (4.3.25), we similarly obtain

$$\begin{aligned} \mathbb{P} [\hat{\nu}_R^N \in \mathcal{A}] \\ \leq \exp \left(\frac{C_3}{\delta_1^{n_1}} \exp \left(2\sqrt{|K|(n-1)} R \right) \left(\frac{R}{\varepsilon} \right)^{n_1} - N \left(\frac{\lambda_2}{2} \eta^2 \varepsilon^2 - C_2^{(2)} R^2 e^{-\alpha R^2} \right) \right). \end{aligned} \quad (4.3.27)$$

Finally, recalling (4.3.18) and (4.3.19), we conclude that in the case $1 \leq p < 2$,

$$\begin{aligned} \mathbb{P} [W_p (\hat{\nu}^N, \nu) > \varepsilon] \\ \leq \exp \left(\frac{C_3}{\delta_1^{n_1}} \exp \left(2\sqrt{|K|(n-1)} R \right) \left(\frac{R}{\varepsilon} \right)^{n_1} - N \left(\frac{\lambda_2}{2} \varepsilon^2 - C_2 R^2 e^{-\alpha R^2} \right) \right) \\ + \exp \left(-N \left\{ \theta(1 - \eta_1)^p \varepsilon^p - E_\alpha e^{-(\alpha - \alpha_1) R^2} \right\} \right), \end{aligned} \quad (4.3.28)$$

while in the case $p = 2$,

$$\begin{aligned} \mathbb{P} [W_2 (\hat{\nu}^N, \nu) > \varepsilon] \\ \leq \exp \left(\frac{C_3}{\delta_1^{n_1}} \exp \left(2\sqrt{|K|(n-1)} R \right) \left(\frac{R}{\varepsilon} \right)^{n_1} - N \left(\frac{\lambda_2}{2} \eta^2 \varepsilon^2 - C_2^{(2)} R^2 e^{-\alpha R^2} \right) \right) \\ + \exp \left(-N \left\{ \frac{\alpha_1}{2} (1 - \eta)^2 \varepsilon^2 - 2E_\alpha e^{-(\alpha - \alpha_1) R^2} \right\} \right). \end{aligned} \quad (4.3.29)$$

Step 3 (Choice of parameters).

The last step consists of choosing parameters and is quite technical. Let us first recall the parameters at play in the case $1 \leq p < 2$. The number $n_1 > n$ is an auxiliary exponent introduced in the metric-entropy estimate, after absorbing the logarithmic factor in the covering bound. The constants $0 < \lambda_2 < \lambda_1 < \lambda$ come from the entropy lower bound after truncation, where ν satisfies $T_p(\lambda)$. The parameters $\eta_1, \delta_1 \in (0, 1)$ are chosen so that

$$\lambda_2 < \lambda_1(\eta_1 - \delta_1)^2.$$

Here η_1 determines how much of the deviation ε is assigned to the compact empirical-measure term, while δ_1 is the relative thickening parameter in the quantitative Sanov estimate, namely $\delta = \delta_1 \varepsilon$. Finally, $\theta > 0$, $\alpha_1 \in (0, \alpha)$, and the cutoff radius $R \gtrsim 1$ are

still free. They will be chosen as functions of ε so that the truncation errors and the metric-entropy contribution are all absorbed into the leading Gaussian term.

Now, fix $\lambda' \in (0, \lambda)$, and choose

$$\lambda_1 := \frac{\lambda + \lambda'}{2}, \quad \lambda_2 := \frac{\lambda_1 + \lambda'}{2}, \quad \Delta := \lambda_2 - \lambda' > 0. \quad (4.3.30)$$

We treat separately the cases $1 \leq p < 2$ and $p = 2$.

Case 1: $1 \leq p < 2$. Set

$$\delta_1 := \frac{1}{4} \left(1 - \sqrt{\frac{\lambda_2}{\lambda_1}} \right), \quad \eta_1 := 1 - \delta_1.$$

Then $0 < \delta_1 < \eta_1 < 1$ and $\lambda_2 < \lambda_1(\eta_1 - \delta_1)^2$; the constant in (4.3.23) is well-defined. Choose

$$\theta(\varepsilon) := \frac{\lambda_2}{(1 - \eta_1)^p} \varepsilon^{2-p}, \quad \alpha_1 := \frac{\alpha}{2},$$

and, denoting by $\log_+$ the nonnegative part of the logarithm,

$$R(\varepsilon) := \max \left\{ R_0, R_1, R_2, 1, \sqrt{\frac{2}{\alpha} \log_+ \left(\frac{8E_\alpha}{\Delta \varepsilon^2} \right)}, \sqrt{\frac{2}{\alpha} \log_+ \left(\frac{16C_2}{\alpha e \Delta \varepsilon^2} \right)} \right\},$$

so that $\theta(1 - \eta_1)^p \varepsilon^p = \lambda_2 \varepsilon^2$,

$$C_2 R^2 e^{-\alpha R^2} \leq \frac{2C_2}{\alpha e} e^{-\alpha R^2/2} \leq \frac{\Delta}{8} \varepsilon^2,$$

and

$$E_\alpha e^{-(\alpha - \alpha_1)R^2} = E_\alpha e^{-\alpha R^2/2} \leq \frac{\Delta}{8} \varepsilon^2.$$

Therefore, the right-hand side of (4.3.28) is bounded by

$$\exp \left(A_K(\varepsilon) - N \left(\frac{\lambda_2}{2} - \frac{\Delta}{8} \right) \varepsilon^2 \right) + \exp \left(-N \left(\lambda_2 - \frac{\Delta}{8} \right) \varepsilon^2 \right), \quad (4.3.31)$$

where

$$A_K(\varepsilon) := \frac{C_3}{\delta_1^{n_1}} \exp \left(2\sqrt{|K|(n-1)} R(\varepsilon) \right) \left(\frac{R(\varepsilon)}{\varepsilon} \right)^{n_1}. \quad (4.3.32)$$

If in addition we require

$$A_K(\varepsilon) \leq \frac{\Delta}{8} N \varepsilon^2, \quad (4.3.33)$$

then both terms of (4.3.31) are bounded by $e^{-\frac{\lambda_2 + \lambda'}{4} N \varepsilon^2}$. Whenever

$$N \varepsilon^2 \geq \frac{4 \log 2}{\Delta}, \quad (4.3.34)$$

this yields

$$\mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq \exp \left(-\frac{\lambda'}{2} N \varepsilon^2 \right)$$

recalling (4.3.30).

It remains to make the condition on N explicit so that both regimes (4.3.33) and (4.3.34) hold. Fix $n' > n_1 > n$ and set $\sigma := n' - n_1 > 0$. We choose a constant C_4 such that

$$R(\varepsilon) \leq C_4 \sqrt{\log(e/\varepsilon)} \quad \text{for all } 0 < \varepsilon < 1. \quad (4.3.35)$$

Such a constant exists, for example

$$C_4 := \max \left\{ \sqrt{\frac{p}{2\alpha}}, \left(\frac{2^p \lambda_2}{\alpha_1 (1 - \eta_1)^p} \right)^{\frac{1}{2-p}}, \sqrt{\frac{\log_+(2E_\alpha)}{\alpha}}, \right. \\ \left. 1, \sqrt{\frac{2}{\alpha} \left(2 + \log_+ \frac{8E_\alpha}{\Delta} \vee \log_+ \frac{16C_2}{\alpha e \Delta} \right)} \right\}.$$

Using

$$\exp \left(2\sqrt{|K|(n-1)} C_4 \sqrt{\log(e/\varepsilon)} \right) \leq \exp \left(\frac{2|K|(n-1)C_4^2}{\sigma} + \frac{\sigma}{2} \right) \varepsilon^{-\frac{\sigma}{2}}$$

and

$$\log(e/\varepsilon)^{n_1/2} \leq \left(\frac{n_1}{e\sigma} \right)^{n_1/2} \left(\frac{e}{\varepsilon} \right)^{\frac{\sigma}{2}}$$

from the elementary inequality $u^a e^{-bu} \leq (\frac{a}{be})^a$ valid for $u \geq 0, a, b, > 0$, recalling (4.3.35), we obtain

$$A_K(\varepsilon) = \frac{C_3}{\delta_1^{n_1}} e^{2\sqrt{|K|(n-1)} R(\varepsilon)} \left(\frac{R(\varepsilon)}{\varepsilon} \right)^{n_1} \leq B_K \varepsilon^{-n'} \quad (4.3.36)$$

with

$$B_K := C_3 \left(\frac{C_4}{\delta_1} \right)^{n_1} \exp \left(\frac{2|K|(n-1)C_4^2}{\sigma} + \sigma \right) \left(\frac{n_1}{e\sigma} \right)^{n_1/2}.$$

Therefore, using (4.3.36), the condition $A_K(\varepsilon) \leq \frac{\Delta}{8} N \varepsilon^2$, namely (4.3.33), is implied by

$$N \geq \frac{8B_K}{\Delta} \varepsilon^{-(n'+2)}.$$

For the condition $N\varepsilon^2 \geq 4\log 2/\Delta$, namely (4.3.34), to be verified as well, since $0 < \varepsilon < 1$, it is enough to take

$$N \geq N_0 \varepsilon^{-(n'+2)}, \quad N_0 := \max \left\{ \frac{8B_K}{\Delta}, \frac{4\log 2}{\Delta} \right\}.$$

Case 2: $p = 2$. Choose

$$\lambda' < \lambda_3 < \lambda_2 < \lambda_1 < \lambda, \quad \Delta_2 := \lambda_3 - \lambda' > 0.$$

Recall that since ν satisfies $T_2(\lambda)$, it has a square-exponential moment of every order $\alpha < \lambda/2$. Set $\alpha_1 := \lambda_3/2$ for $\alpha_1 < \alpha < \lambda/2$ and

$$\eta := \sqrt{2} - 1, \quad \delta_1 := \frac{\eta}{4} \left(1 - \sqrt{\frac{\lambda_2}{\lambda_1}} \right).$$

Then $0 < \delta_1 < \eta < 1$ and $\lambda_2\eta^2 < \lambda_1(\eta - \delta_1)^2$; the constant in (4.3.24) is well-defined. Arguing as in the case $1 \leq p < 2$, for

$$R(\varepsilon) := \max \left\{ R_0, R_2, 1, \sqrt{\frac{1}{\alpha - \alpha_1} \log_+ \left(\frac{16E_\alpha^2}{\Delta_2\eta^2\varepsilon^2} \right)}, \sqrt{\frac{2}{\alpha} \log_+ \left(\frac{16C_2^{(2)}}{\alpha e \Delta_2\eta^2\varepsilon^2} \right)} \right\},$$

since

$$\frac{\alpha_1}{2}(1 - \eta)^2\varepsilon^2 = \alpha_1\eta^2\varepsilon^2 = \frac{\lambda_3}{2}\eta^2\varepsilon^2$$

with $\eta = \sqrt{2} - 1$, then the right-hand side of (4.3.29) is bounded by

$$\exp \left(A_K(\varepsilon) - N \left(\frac{\lambda_2}{2}\eta^2\varepsilon^2 - \frac{\Delta_2}{8}\eta^2\varepsilon^2 \right) \right) + \exp \left(-N \left(\frac{\lambda_3}{2}\eta^2\varepsilon^2 - \frac{\Delta_2}{8}\eta^2\varepsilon^2 \right) \right). \quad (4.3.37)$$

If in addition

$$A_K(\varepsilon) \leq \frac{\Delta_2}{8}\eta^2 N \varepsilon^2, \quad (4.3.38)$$

then

$$\mathbb{P} [W_2(\hat{\nu}^N, \nu) > \varepsilon] \leq 2 \exp \left(-\frac{\lambda_3 + \lambda'}{4}\eta^2 N \varepsilon^2 \right).$$

Whenever

$$N\varepsilon^2 \geq \frac{4\log 2}{\Delta_2\eta^2}, \quad (4.3.39)$$

this yields

$$\mathbb{P} [W_2(\hat{\nu}^N, \nu) > \varepsilon] \leq \exp \left(-\frac{\lambda'}{2}\eta^2 N \varepsilon^2 \right). \quad (4.3.40)$$

Fix $n' > n_1 > n$ and set $\sigma := n' - n_1 > 0$. We choose $C_4^{(2)}$ such that

$$R(\varepsilon) \leq C_4^{(2)} \sqrt{\log(e/\varepsilon)} \quad \text{for all } 0 < \varepsilon < 1. \quad (4.3.41)$$

For instance, one may take

$$C_4^{(2)} := \max \left\{ R_0, R_2, 1, \sqrt{\frac{1}{\alpha - \alpha_1} \left(2 + \log_+ \frac{16E_\alpha^2}{\Delta_2\eta^2} \right)}, \sqrt{\frac{2}{\alpha} \left(2 + \log_+ \frac{16C_2^{(2)}}{\alpha\epsilon\Delta_2\eta^2} \right)} \right\}.$$

Arguing as in Case 1, from (4.3.41) we obtain

$$A_K(\varepsilon) \leq B_K^{(2)} \varepsilon^{-n'}, \quad (4.3.42)$$

with

$$B_K^{(2)} := C_3 \left(\frac{C_4^{(2)}}{\delta_1} \right)^{n_1} \exp \left(\frac{2|K|(n-1)(C_4^{(2)})^2}{\sigma} + \sigma \right) \left(\frac{n_1}{\epsilon\sigma} \right)^{n_1/2}.$$

Thus, by (4.3.42), the condition (4.3.38) is implied by

$$N \geq \frac{8B_K^{(2)}}{\Delta_2\eta^2} \varepsilon^{-(n'+2)}.$$

For the condition $N\varepsilon^2 \geq 4\log 2/(\Delta_2\eta^2)$, namely (4.3.39), to be verified as well, since $0 < \varepsilon < 1$, it is enough to take

$$N \geq N_0^{(2)} \varepsilon^{-(n'+2)}, \quad N_0^{(2)} := \max \left\{ \frac{8B_K^{(2)}}{\Delta_2\eta^2}, \frac{4\log 2}{\Delta_2\eta^2} \right\}.$$

Since $\eta^2 = 3 - 2\sqrt{2}$, (4.3.40) gives

$$\mathbb{P} [W_2(\hat{\nu}^N, \nu) > \varepsilon] \leq \exp \left(-(3 - 2\sqrt{2}) \frac{\lambda'}{2} N \varepsilon^2 \right),$$

which concludes the proof. $\square$

Proof of Theorem 4.16. Since the proof follows the same line as the one of Theorem 4.15, we only highlight the differences.

Step 1 (Truncation). Let $s \in [1, q/p]$ and set $\bar{s} := \max\{s/2, 1\}$. Since now only the polynomial moment $M_q := \int_E d^q(o, x) d\nu(x)$ is finite with $p < q$, and instead of (4.3.12) we have

$$W_p^p(\nu_R, \nu) \leq 2^p M_q R^{p-q}. \quad (4.3.43)$$

Recall that

$$W_p^p(\hat{\nu}^N, \hat{\nu}_R^N) \leq \frac{1}{N} \sum_{i=1}^N Z^i, \quad Z^i := 2^p d^p(o, X^i) 1_{\{d(o, X^i) > R\}}.$$

Using

$$\mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N (Z^i - \mathbb{E}[Z^i]) \right|^s \right] \leq C_s N^{\bar{s}-s} \mathbb{E}[(Z^1)^s], \quad C_s := \begin{cases} 2^{s+1}, & 1 \leq s \leq 2; \\ (2s)^s, & s > 2; \end{cases}$$

together with

$$\mathbb{E} Z_1 \leq 2^p M_q R^{p-q}, \quad \mathbb{E} Z_1^s \leq 2^{ps} M_q R^{ps-q},$$

whenever $t^p > 2^p M_q R^{p-q}$, Markov's inequality then yields

$$\begin{aligned} \mathbb{P} [W_p^p(\hat{\nu}^N, \hat{\nu}_R^N) > t^p] &\leq \mathbb{P} \left[\left| \frac{1}{N} \sum_{i=1}^N (Z_i - \mathbb{E}[Z^i]) \right|^s > |t^p - \mathbb{E}[Z^1]|^s \right] \\ &\leq C_s 2^{ps} M_q N^{\bar{s}-s} \frac{R^{ps-q}}{(t^p - 2^p M_q R^{p-q})^s}. \end{aligned} \quad (4.3.44)$$

For $R \geq 1$ is large enough so that $W_p(\nu_R, \nu) \leq \varepsilon/8$, using the triangle inequality and (4.3.44) we obtain

$$\begin{aligned} \mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] &\leq \mathbb{P} [W_p(\hat{\nu}^N, \hat{\nu}_R^N) > \frac{\varepsilon}{2}] + \mathbb{P} [W_p(\hat{\nu}_R^N, \nu_R) > \frac{3\varepsilon}{8}] \\ &\quad + \mathbb{P} [W_p(\nu_R, \nu) > \frac{\varepsilon}{8}] \\ &\leq \mathbb{P} [W_p(\hat{\nu}_R^N, \nu_R) > \frac{3\varepsilon}{8}] + C_s 2^{ps} M_q N^{\bar{s}-s} \frac{R^{ps-q}}{\left(\left(\frac{\varepsilon}{2}\right)^p - 2^p M_q R^{p-q}\right)^s}. \end{aligned} \quad (4.3.45)$$

Letting $0 < \varepsilon < 1$,

$$R \geq R_0 \varepsilon^{-\frac{p}{q-p}}, \quad R_0 := \max \left\{ (16M_q^{1/p})^{\frac{p}{q-p}}, (2^{2p+1} M_q)^{\frac{1}{q-p}}, 1 \right\}, \quad (4.3.46)$$

not only makes $W_p(\nu_R, \nu) \leq \varepsilon/8$ hold recalling (4.3.43), but also

$$\left(\frac{\varepsilon}{2}\right)^p - 2^p M_q R^{p-q} \geq 2^{-(p+1)} \varepsilon^p,$$

and (4.3.45) becomes

$$\mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq \mathbb{P} [W_p(\hat{\nu}_R^N, \nu_R) > \frac{3\varepsilon}{8}] + C_1 N^{\bar{s}-s} \varepsilon^{-ps}, \quad C_1 := C_s 2^{s(2p+1)} M_q. \quad (4.3.47)$$

Step 2 (Covering number estimates). We apply the quantitative Sanov estimate (4.3.20),

$$\mathbb{P}[\hat{\nu}_R^N \in \mathcal{A}] \leq \left(\frac{16eR}{\delta'} \right)^{pN(B_R(o), \frac{\delta'}{2})} \exp \left(-N \inf_{\nu' \in \mathcal{A}_{\delta'}} H(\nu' | \nu_R) \right) \quad (4.3.48)$$

to

$$\mathcal{A} := \left\{ \nu' \in \mathcal{P}(B_R(o)) : W_p(\nu', \nu_R) > \frac{3\varepsilon}{8} \right\}, \quad \delta' := \frac{\varepsilon}{8}.$$

If $\nu' \in \mathcal{A}_{\delta'}$, then $W_p(\nu', \nu_R) \geq \varepsilon/4$. Note that ν_R satisfies a penalized T_p inequality thanks to the Csiszár–Kullback–Pinsker inequality [12, Particular case 2.5];

$$W_p^{2p}(\nu', \nu_R) \leq 2^{2p-1} R^{2p} H(\nu' | \nu_R), \quad \forall \nu' \in \mathcal{P}(B_R(o)).$$

Thus for $\nu' \in \mathcal{A}_{\delta'}$,

$$H(\nu' | \nu_R) \geq 2^{1-6p} \left(\frac{\varepsilon}{R} \right)^{2p}. \quad (4.3.49)$$

Let $n_1 > n$ and recall the estimate (4.3.25),

$$pN \left(B_R(o), \frac{\delta'}{2} \right) \log \left(\frac{16eR}{\delta'} \right) \leq C_2 \exp \left(2\sqrt{|K|(n-1)} R \right) \left(\frac{R}{\delta'} \right)^{n_1} \quad (4.3.50)$$

with

$$C_2 := \frac{p 10^n (16e)^{n_1-n}}{e(n_1-n)}.$$

Combining (4.3.48), (4.3.49), and (4.3.50), we obtain

$$\mathbb{P} \left[W_p(\hat{\nu}_R^N, \nu_R) > \frac{3\varepsilon}{8} \right] \leq \exp \left(C_2 8^{n_1} e^{2\sqrt{|K|(n-1)} R} \left(\frac{R}{\varepsilon} \right)^{n_1} - 2^{1-6p} N \left(\frac{\varepsilon}{R} \right)^{2p} \right). \quad (4.3.51)$$

Step 3 (Choice of parameters). Fix $n' := n_1 > n$ and recall the condition (4.3.46) on R ; set

$$R := R_0 \varepsilon^{-\frac{p}{q-p}}.$$

We can estimate the exponential term in (4.3.51) as

$$\begin{aligned} \exp \left(C_2 8^{n'} e^{2\sqrt{|K|(n-1)} R} \left(\frac{R}{\varepsilon} \right)^{n'} - 2^{1-6p} N \left(\frac{\varepsilon}{R} \right)^{2p} \right) &\leq \exp \left(-2^{-6p} N \left(\frac{\varepsilon}{R} \right)^{2p} \right) \\ &\leq \exp \left(-C_3 N \varepsilon^{2p(1+\frac{p}{q-p})} \right), \end{aligned}$$

where $C_3 := (8R_0)^{-2p}$, provided

$$N \geq N'_0 \varepsilon^{-(n'+2p)(1+\frac{p}{q-p})} \exp\left(2\sqrt{|K|(n-1)R}\right), \quad N'_0 := 2^{6p} C_2 8^{n'} R_0^{n'+2p}. \quad (4.3.52)$$

Combining with (4.3.51), we obtain

$$\mathbb{P}\left[W_p(\hat{\nu}^N, \nu) > \varepsilon\right] \leq \exp\left(-C_3 N \varepsilon^{2p(1+\frac{p}{q-p})}\right) + C_1 N^{\bar{s}-s} \varepsilon^{-ps}. \quad (4.3.53)$$

Now set

$$s := \frac{q}{p} - \frac{\delta}{2}.$$

For the case (i), $1 \leq p < q/2$, since $s > 2$ and $\bar{s} = s/2$, then

$$C_1 N^{\bar{s}-s} \varepsilon^{-ps} \leq \frac{1}{2} \varepsilon^{-q} N^{-\frac{q}{2p} + \frac{\delta}{2}}. \quad (4.3.54)$$

if in addition $N_0 \geq (2C_1)^{4/\delta}$. For the case (ii), $q/2 \leq p < q$, since $1 < s < 2$ and $\bar{s} = 1$, then

$$C_1 N^{\bar{s}-s} \varepsilon^{-ps} \leq \frac{1}{2} \varepsilon^{-q} N^{1-\frac{q}{p}+\delta}. \quad (4.3.55)$$

provided the same additional condition $N_0 \geq (2C_1)^{4/\delta}$.

It remains to dominate the exponential term of (4.3.53). We claim that, under (4.3.52), after possibly enlarging N'_0 ,

$$\exp\left(-C_3 N \varepsilon^{2p(1+\frac{p}{q-p})}\right) \leq \frac{1}{2} \varepsilon^{-q} N^{-\beta}, \quad \beta := \begin{cases} \frac{q}{2p} - \frac{\delta}{2}, & \text{case (i);} \\ \frac{q}{p} - 1 - \delta, & \text{case (ii).} \end{cases} \quad (4.3.56)$$

Set

$$a := (n' + 2p) \left(1 + \frac{p}{q-p}\right) = \frac{q(n' + 2p)}{q-p}, \quad \gamma := 2p \left(1 + \frac{p}{q-p}\right) = \frac{2pq}{q-p}.$$

A possible choice is given by

$$N''_0 := \max \left\{ \frac{4\beta}{C_3}, \frac{4}{C_3} \left[\log 2 + \frac{c}{e} + \beta \log_+ \left(\frac{4\beta}{C_3} \right) \right] \right\}, \quad c := \frac{(a\beta - q)_+}{a - \gamma}. \quad (4.3.57)$$

Then $N''_0 \geq N'_0$ implies (4.3.56) for all $N \geq N''_0$. To see this, observe first that for $0 < \varepsilon < 1$, $N \geq N'_0 \varepsilon^{-a}$, and $N'_0 \geq \beta/C_3$, it holds

$$N \geq \frac{\beta}{C_3} \varepsilon^{-\gamma}.$$

Therefore, the function $k \mapsto e^{-C_3 k \varepsilon^\gamma} k^\beta$ is nonincreasing for $k \geq N$, and it is enough to prove (4.3.56) at $N = N_0'' \varepsilon^{-a}$. This reduces to

$$e^{-C_3 N_0'' \varepsilon^{-(a-\gamma)}} \leq \frac{1}{2} (N_0'')^{-\beta} \varepsilon^{-(q-a\beta)}.$$

Equivalently, taking logarithms, it is enough that

$$C_3 N_0'' \varepsilon^{-(a-\gamma)} \geq \log 2 + \beta \log N_0'' + c \log \left(\varepsilon^{-(a-\gamma)} \right),$$

but $\log(\varepsilon^{-(a-\gamma)}) \leq \varepsilon^{-(a-\gamma)}/e$ for $\varepsilon \leq 1$, and since (4.3.57) implies

$$C_3 N_0'' \geq \log 2 + \beta \log N_0'' + \frac{c}{e},$$

the claim follows.

We thus require

$$N \geq N_0 \varepsilon^{-\frac{q(n'+2p)}{q-p}} \exp \left(2\sqrt{|K|(n-1)R} \right), \quad N_0 := \max \left\{ N_0', N_0'', (1 \vee 2C_1)^{4/\delta} \right\}. \quad (4.3.58)$$

Combining (4.3.47) and (4.3.53)–(4.3.56) under (4.3.58), we obtain in case (i)

$$\mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq \varepsilon^{-q} N^{-\frac{q}{2p} + \frac{\delta}{2}},$$

while in case (ii)

$$\mathbb{P} [W_p(\hat{\nu}^N, \nu) > \varepsilon] \leq \varepsilon^{-q} N^{1-\frac{q}{p} + \delta}.$$

This concludes the proof. $\square$

4.4 An interacting particle system

4.4.1 Description of the interacting particle system

Let (M, g) be a complete Riemannian manifold of dimension $m \geq 2$. Let $V : M \rightarrow \mathbb{R}$ be a confinement potential and $W : M \times M \rightarrow \mathbb{R}$ an interaction potential. For $N \geq 1$, we consider the McKean–Vlasov interacting particle system

$$dX_t^{i,N} = \sqrt{2} dB_t^{i,M} - \nabla V(X_t^{i,N}) dt - \frac{1}{N} \sum_{j=1}^N \nabla_x W(X_t^{i,N}, X_t^{j,N}) dt, \quad i = 1, \dots, N, \quad (4.4.1)$$

where $B^{1,M}, \dots, B^{N,M}$ are independent Brownian motions on M . We denote by

$$\hat{\mu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}$$

the associated empirical measure.

Formally, as $N \rightarrow \infty$, the empirical measure $\hat{\mu}_t^N$ converges to a deterministic probability measure μ_t . The limiting equation is obtained by replacing the empirical interaction in (4.4.1) by its average against μ_t . Thus μ_t solves the nonlinear Fokker–Planck equation, also called the aggregation-diffusion equation,

$$\partial_t \mu_t = \Delta \mu_t + \nabla \cdot (\mu_t [\nabla V + \nabla (W * \mu_t)]), \quad (4.4.2)$$

where

$$W * \mu_t(x) := \int_M W(x, y) d\mu_t(y), \quad \nabla (W * \mu_t)(x) = \int_M \nabla_x W(x, y) d\mu_t(y).$$

This is the usual propagation-of-chaos picture for McKean–Vlasov diffusions. In this chapter, we do not address the general well-posedness theory for (4.4.2); instead, we assume that a sufficiently regular solution $(\mu_t)_{t \geq 0}$ is given and prove a quantitative concentration estimate around it.

As in Bolley–Guillin–Villani [11], we introduce a noninteracting proxy system $(Y_t^i)_{1 \leq i \leq N}$. Its particles are independent and have common law μ_t :

$$dY_t^i = \sqrt{2} d\bar{B}_t^{i,M} - \nabla V(Y_t^i) dt - \int_M \nabla_x W(Y_t^i, z) d\mu_t(z) dt. \quad (4.4.3)$$

We write

$$\hat{\nu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}.$$

Since $\hat{\nu}_t^N$ is the empirical measure of independent variables with common law μ_t , the concentration estimate of Theorem 4.15 can be applied to it once suitable moment estimates are propagated along (4.4.2).

The proof has two parts. First, we prove square-exponential moment bounds and time-increment estimates for the proxy process Y_t . These estimates yield concentration estimates for

$$\sup_{0 \leq t \leq T} W_1(\hat{\nu}_t^N, \mu_t).$$

Second, we couple the interacting system $X_t^{i,N}$ with the proxy system Y_t^i . The coupling is constructed by parallel displacement of the Brownian noises. For this reason, it is convenient to write the stochastic equations on the orthonormal frame bundle $O(M)$. The coupling estimate then reduces the control of $W_1(\hat{\mu}_t^N, \mu_t)$ to the concentration of the independent proxy empirical measure $\hat{\nu}_t^N$.

4.4.2 Assumptions on the potentials

Fix $o \in M$. We assume that there exist constants

$$\beta, \tilde{\beta} > 0, \quad A_V, A_W \geq 0, \quad L_V, L_{W,1}, L_{W,2} > 0,$$

such that, for all $x, y, z \in M$,

$$\text{Hess } V \geq \beta g, \quad |\nabla V(x)| \leq A_V + L_V r(x), \quad (4.4.4)$$

and

$$\text{Hess}_x W(\cdot, y) \geq \tilde{\beta} g, \quad |\nabla_x W(x, y)| \leq A_W + L_{W,1}(d(o, x) + d(o, y)). \quad (4.4.5)$$

We also assume that the dependence of $\nabla_x W$ on the second variable is Lipschitz, uniformly in the first variable; i.e., for every $x, y, z \in M$ and every unit vector $u \in T_x M$,

$$|(\nabla_x W(x, y) - \nabla_x W(x, z)) \cdot u| \leq L_{W,2} d(y, z). \quad (4.4.6)$$

The lower Hessian bound in the first variable gives dissipation, the linear growth assumption controls the moment estimates, and the Lipschitz assumption in the second variable is used in the coupling estimate through Kantorovich duality.

As an example, on a Hadamard manifold, the interaction potential

$$W(x, y) = \frac{\ell}{2} d^2(x, y), \quad \ell > 0,$$

satisfies the assumptions. Indeed, globally $\nabla_x W(x, y) = -\ell \log_x(y)$, and therefore

$$|\nabla_x W(x, y)| = \ell d(x, y) \leq \ell(d(o, x) + d(o, y)).$$

Moreover, by Hessian comparison [72, Theorem 11.7],

$$\text{Hess}_x W(\cdot, y) = \ell \text{Hess}_x \left(\frac{1}{2} d^2(\cdot, y) \right) \geq \ell g.$$

Finally, since M is Hadamard, the logarithm map $y \mapsto \log_x(y)$ is 1-Lipschitz. Hence

$$|\nabla_x W(x, y) - \nabla_x W(x, z)| \leq \ell d(y, z).$$

Thus, the assumptions (4.4.5) and (4.4.6) on W hold with $A_W = 0, L_{W,1} = L_{W,2} = \tilde{\beta} = \ell$. The potential

$$V(x) = \frac{\ell}{2} d^2(o, x)$$

thus also satisfies (4.4.4) with $A_V = 0$ and $L_V = \beta = \ell$.

4.4.3 Preliminary results

We establish propagation of concentration-cost inequality $T_1(\lambda(t))$ for (4.4.2), which we recall is equivalent to propagation of square-exponential moments.

Lemma 4.19 (Propagation of square-exponential moments). *Let (M, g) be a complete non-compact Riemannian manifold of dimension $m \geq 2$ satisfying $\text{Ric} \geq K$ for some $K \leq 0$. Let μ_t solve the aggregation-diffusion equation (4.4.2) with potentials V, W satisfying respectively (4.4.4) and (4.4.5). Assume that μ_0 has finite square-exponential moment; i.e., there is a constant $\alpha(0) > 0$ such that*

$$E_{\alpha(0)} := \int_M e^{\alpha(0)d^2(o,x)} d\mu_0(x) < +\infty.$$

Then for every $t > 0$, there exists a continuous function $\alpha(t) > 0$ such that

$$E_{\alpha(t)} := \int_M e^{\alpha(t)d^2(o,x)} d\mu_t(x) < +\infty.$$

If in addition $\beta + \tilde{\beta} > L_{W,1}$, then $\sup_{t \geq 0} E_{\alpha(t)} < +\infty$.

Proof. (Step 1) Propagation of second moment. Set $r(x) := d(o, x)$ and

$$M_2(t) := \int_M r^2(x) d\mu_t(x),$$

Differentiating and integrating by parts, formally,

$$\begin{aligned} \frac{d}{dt} M_2(t) &= \int_M r^2(x) \partial_t \mu_t(x) d\text{Vol}(x) \\ &= \int_M r^2(x) \left\{ \Delta \mu_t(x) + \nabla \cdot \left(\mu_t(x) [\nabla V + \nabla (W * \mu_t)](x) \right) \right\} d\text{Vol}(x) \\ &= \int_{M \setminus (\{o\} \cup \text{Cut}(o))} \Delta(r^2)(x) d\mu_t(x) \\ &\quad - \int_M \nabla(r^2)(x) \cdot \left([\nabla V + \nabla (W * \mu_t)](x) \right) d\mu_t(x). \end{aligned} \tag{4.4.7}$$

Note that r^2 is smooth only outside of $\text{Cut}(o)$, which is of measure zero (see e.g. [57, Theorem 3.4.1]). One can rigorously justify integration by parts outside of $\{o\} \cup \text{Cut}(o)$ by cutoffs and approximations; such justification is beyond the scope of this chapter and will be omitted.

By Laplacian comparison theory [72, Theorem 11.15], outside of $\{o\} \cup \text{Cut}(o)$,

$$\Delta r(x) \leq (m-1) \sqrt{\frac{|K|}{m-1}} \coth \left(\sqrt{\frac{|K|}{m-1}} r(x) \right) \leq \frac{m-1}{r(x)} + \sqrt{|K|(m-1)},$$

where in the last equality we used $\coth u \leq 1 + 1/u$. Then, since $|\nabla r(x)| = 1$,

$$\Delta r^2(x) = 2r(x)\Delta r(x) + 2|\nabla r(x)|^2 = 2r(x)\Delta r(x) + 2 \leq 2m + 2\sqrt{|K|(m-1)}r(x). \quad (4.4.8)$$

Recall we assumed that

$$\text{Hess } V \geq \beta, \quad \text{Hess } W(\cdot, y) \geq \tilde{\beta}, \quad |\nabla_x W(o, y)| \leq A_W + L_{W,1}r(y).$$

Let $\Lambda := \beta + \tilde{\beta}$. If $\gamma : [0, r(x)] \rightarrow M$ is a unit-speed minimizing geodesic from o to x , then outside of $\{o\} \cup \text{Cut}(o)$,

$$\begin{aligned} \nabla r(x) \cdot \nabla V(x) &= \frac{d}{ds} \Big|_{s=r(x)} V(\gamma(s)) \\ &= \frac{d}{ds} \Big|_{s=0} V(\gamma(s)) + \int_0^{r(x)} \frac{d^2}{ds^2} V(\gamma(s)) \, ds \\ &= \langle \nabla V(o), \dot{\gamma}(0) \rangle + \int_0^{r(x)} \text{Hess } V_{\gamma(s)}(\dot{\gamma}(s), \dot{\gamma}(s)) \, ds \\ &\geq -|\nabla V(o)| + \beta r(x). \end{aligned}$$

Therefore, for $x \in M \setminus (\{o\} \cup \text{Cut}\{o\})$, it holds

$$\nabla r^2(x) \cdot \nabla V(x) \geq 2\beta r^2(x) - 2|\nabla V(o)|r(x). \quad (4.4.9)$$

Similarly, for every $x \in M \setminus (\{o\} \cup \text{Cut}\{o\})$ and for every $y \in M$,

$$\nabla r^2(x) \cdot \nabla_x W(x, y) \geq 2\tilde{\beta} r^2(x) - 2|\nabla_x W(o, y)|r(x). \quad (4.4.10)$$

Combining (4.4.7)–(4.4.9), and after integrating (4.4.10) in $d\mu_t(y)$, we obtain

$$\frac{d}{dt} M_2(t) \leq 2m + 2C_1 \sqrt{M_2(t)} - 2(\beta + \tilde{\beta} - L_{W,1}) M_2(t) \quad (4.4.11)$$

with

$$C_1 := \sqrt{|K|(m-1)} + |\nabla V(o)| + A_W.$$

By Gronwall's inequality, given $T > 0$, from (4.4.11) we conclude

$$\sup_{t \in [0, T]} M_2(t) := M_{2,T} < +\infty,$$

and if in addition $\Lambda := \beta + \tilde{\beta} > L_{W,1}$, then $M_{2,T} < C_{M_2}$ for some constant $C_{M_2} > 0$, independently of T .

Step 2 (Propagation of square-exponential moments). Let

$$E_{\alpha(t)} := \int_M e^{\alpha(t)r^2(x)} d\mu_t(x),$$

where $\alpha(t) > 0$ will be chosen below. Since outside of $\{o\} \cup \text{Cut}(o)$,

$$\nabla e^{\alpha r^2} = \alpha e^{\alpha r^2} \nabla r^2, \quad \Delta e^{\alpha r^2} = e^{\alpha r^2} (\alpha \Delta r^2 + \alpha^2 |\nabla r^2|^2), \quad |\nabla r^2|^2 = 4r^2,$$

we obtain, using (4.4.8)–(4.4.10),

$$\begin{aligned} \frac{d}{dt} E_{\alpha(t)} &\leq \int_M e^{\alpha(t)r^2} \left[2m\alpha(t) + \left(\frac{d}{dt} \alpha(t) + 4\alpha(t)^2 - 2\Lambda\alpha(t) \right) r^2 + 2\alpha(t)C_2r \right] d\mu_t \\ &\leq (2m + C_2^2) \alpha(t) E_{\alpha(t)} + \left\{ \frac{d}{dt} \alpha(t) + 4\alpha(t)^2 + (1 - 2\Lambda) \alpha(t) \right\} \int_M r^2 e^{\alpha r^2} d\mu_t, \end{aligned} \quad (4.4.12)$$

$$(4.4.13)$$

where we used $2\alpha C_2 r \leq \alpha C_2^2 + \alpha r^2$, with

$$C_2 := \sqrt{|K|(m-1)} + |\nabla V(o)| + A_W + L_{W,1} \sqrt{M_{2,T}}.$$

We choose α as the solution of

$$\frac{d}{dt} \alpha(t) + 4\alpha(t)^2 + (1 - 2\Lambda) \alpha(t) = 0, \quad \alpha(0) = \alpha_0,$$

so that the term in brackets in (4.4.13) vanishes. Since $\alpha(t) > 0$ on $[0, T]$, (4.4.13) becomes

$$\frac{d}{dt} E_{\alpha(t)}(t) \leq \alpha(t)(2m + C_2^2) E_{\alpha(t)}(t).$$

Hence, for every $T > 0$, Gronwall's inequality yields

$$\sup_{0 \leq t \leq T} E_{\alpha(t)}(t) \leq E_{\alpha_0}(0) \exp \left((2m + C_2^2) \int_0^T \alpha(s) ds \right) < +\infty.$$

Assume now furthermore that $\Lambda > L_{W,1}$ and set

$$C_2^\infty := \sqrt{|K|(m-1)} + |\nabla V(o)| + A_W + L_{W,1} C_{M_2}^{1/2}.$$

In this case, we choose $\alpha(t)$ as the solution of

$$\frac{d}{dt} \alpha(t) = \Lambda \alpha(t) - 4\alpha(t)^2, \quad \alpha(0) = \alpha_0$$

so that

$$\alpha_- := \min \left\{ \alpha_0, \frac{\Lambda}{4} \right\} \leq \alpha(t) \leq \max \left\{ \alpha_0, \frac{\Lambda}{4} \right\} =: \alpha_+.$$

Using $2C_2^\infty r \leq \frac{\Lambda}{2}r^2 + \frac{2(C_2^\infty)^2}{\Lambda}$ in the differential inequality (4.4.12) gives

$$\frac{d}{dt} E_{\alpha(t)} \leq \alpha(t) \int_M \left(C_3 - \frac{\Lambda}{2} r^2 \right) e^{\alpha(t)r^2} d\mu_t, \quad C_3 := 2m + \frac{2(C_2^\infty)^2}{\Lambda}.$$

Splitting the integral over $\{r \leq R_*\}$ and $\{r > R_*\}$ for $R_*^2 := 4C_3/\Lambda$ yields

$$\frac{d}{dt} E_{\alpha(t)} \leq 2C_3\alpha_+ e^{\alpha_+ R_*^2} - C_3\alpha_- E_{\alpha(t)},$$

and thus $\sup_{t \geq 0} E_{\alpha(t)} < +\infty$. □

We now use the propagation of square-exponential moments for the solution of the aggregation-diffusion equation to derive uniform-in-time concentration estimates for a suitable noninteracting particle system. In this auxiliary system, the empirical mean-field interaction is replaced by the limiting law μ_t . Equivalently, the dynamics are written in Stratonovich form on the orthonormal frame bundle $\pi : O(M) \rightarrow M$ as

$$dU_t^i = \sqrt{2} \sum_{k=1}^m H_k(U_t^i) \circ dB_t^{i,k} + \tilde{b}_t(U_t^i) dt, \quad Y_t^i = \pi(U_t^i), \quad i = 1, \dots, N, \quad (4.4.14)$$

where $B_t = (B_t^1, \dots, B_t^m)$ is a standard Brownian motion on $\mathbb{R}^m$, $\tilde{b}_t$ denotes the horizontal lift of the time-dependent vector field

$$b_t(x) := -\nabla V(x) - \nabla(W * \mu_t)(x),$$

and H_k is the fundamental horizontal vector field associated with the canonical basis vector $e_k \in \mathbb{R}^m$; see Chapter 2 for the notion of horizontal lift to the orthonormal frame bundle and the monograph of Hsu [57]. We obtain that the empirical measure of $(Y_t^i)_{1 \leq i \leq N}$ given by

$$\hat{\nu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}$$

satisfies the same quantitative estimate as [11, Proposition 5.1], under some additional assumptions on the potentials V and W .

Proposition 4.20. *In the setting of Lemma 4.19, for every $T > 0$, there are constants $C_T, a_T > 0$ such that, for all $0 < \delta < T$ and for all $\varepsilon > 0$,*

$$\mathbb{P} \left[\sup_{s, t: 0 \leq t_0 \leq s \leq t \leq t_0 + \delta} W_1(\hat{\nu}_s^N, \hat{\nu}_t^N) > \varepsilon \right] \leq \exp(-N(a_T \varepsilon^2 - C_T \delta))$$

This is a consequence of the following result.

Lemma 4.21 (Estimates for the noninteracting proxy). *In the setting of Lemma 4.19, let $(Y_t)_{t \geq 0}$ solve (4.4.14). Then for every $T > 0$, there are constants $C_T, a_T > 0$ such that, for every $0 \leq t_0 \leq s \leq t \leq t_0 + \delta \leq T$, $T > 0$,*

$$\mathbb{E} [d^2(Y_t, Y_s)] \leq C_T(t - s) \quad (4.4.15)$$

and

$$\mathbb{E} \left[\sup_{s, t: t_0 \leq s \leq t \leq t_0 + \delta} \exp(a_T d^2(Y_t, Y_s)) \right] \leq 1 + C_T \delta. \quad (4.4.16)$$

Proof of Lemma 4.21. Fix $0 \leq s \leq t \leq T$ and denote $r := d(\cdot, Y_s)$. By Kendall's formula [66, Theorem 1.1] (see also [5] and [57, Theorem 3.5.1]),

$$dr^2(Y_t) = \sqrt{2}d\beta_t + [\Delta r(Y_t) + b_t(Y_t) \cdot \nabla r(Y_t)] dt - dL_t. \quad (4.4.17)$$

Here ∇r and Δr are understood in distributional sense on $\{Y_s\} \cup \text{Cut}(Y_s)$, β_t is a one-dimensional Brownian motion and L_t is a nondecreasing process which increases only when the Brownian motion is at the cut locus. Since $-dL_t \leq 0$, it can be dropped in upper bounds. Hence, applying Itô's formula to r^2 , noting that $d\langle r(Y) \rangle_t = 2dt$ and recalling (4.4.8),

$$dr^2(Y_t) \leq 2\sqrt{2}r(Y_t)d\beta_t + \left[2m + 2\sqrt{|K|(m-1)}r_t + 2r_t b_t(Y_t) \cdot \nabla r(Y_t) \right] dt. \quad (4.4.18)$$

We now estimate the drift term. From the Hessian lower bounds, $\text{Hess } V \geq \beta$ and $\text{Hess } W(\cdot, y) \geq \tilde{\beta}$, we obtain

$$b_t(Y_t) \cdot \nabla r(Y_t) \leq |\nabla V(Y_s)| + |\nabla(W * \mu_t)(Y_s)| - (\beta + \tilde{\beta}) r(Y_t).$$

Under the growth assumptions on V and W , and recalling that $M_{2,T} := \sup_{t \in [0, T]} M_2(t) < +\infty$ by Lemma 4.19, we have

$$|\nabla V(Y_s)| + |\nabla(W * \mu_t)(Y_s)| \leq A_V + L_V d(o, Y_s) + A_W + L_{W,1} \left(d(o, Y_s) + \sqrt{M_{2,T}} \right).$$

Inserting this into (4.4.18) and using Cauchy–Schwartz inequality yields

$$dr^2(Y_t) \leq 2\sqrt{2}r(Y_t)d\beta_t + C'_T (1 + d^2(o, Y_s) + r^2(Y_t)) dt \quad (4.4.19)$$

with

$$C'_T := \max \left\{ 3, (L_V + L_{W,1})^2, 2m + |K|(m-1) + \left(A_V + A_W + L_{W,1} \sqrt{M_{2,T}} \right)^2 \right\}.$$

After localizing and taking conditional expectation with respect to Y_s , we obtain

$$\frac{d}{dt} \mathbb{E} \left[r^2(Y_t) \mid Y_s \right] \leq C'_T (1 + d^2(o, Y_s)) + C'_T \mathbb{E} \left[r^2(Y_t) \mid Y_s \right].$$

Since $r(Y_s) = d(Y_s, Y_s) = 0$, by Gronwall's inequality,

$$\mathbb{E} \left[r^2(Y_t) \mid Y_s \right] \leq C'_T (1 + d^2(o, Y_s)) \int_s^t e^{C'_T(t-u)} du \leq C'_T e^{C'_T T} (1 + d^2(o, Y_s)) (t - s). \quad (4.4.20)$$

Taking the expectation and recalling that $\sup_{0 \leq s \leq T} \mathbb{E}[d^2(o, Y_s)] = M_{2,T} < +\infty$ by [Lemma 4.19](#) since $\text{Law}(Y_s) = \mu_s$, we obtain [\(4.4.15\)](#);

$$\mathbb{E} [d^2(Y_t, Y_s)] = \mathbb{E} [r^2(Y_t)] \leq C'_T e^{C'_T T} (1 + M_{2,T}) (t - s).$$

We now estimate $e^{a(t)r^2(Y_t)}$, where $a(t) > 0$ will be chosen below. From [\(4.4.19\)](#), we have

$$dr^2(Y_t) \leq 2\sqrt{2} r(Y_t) d\beta_t + C'_T (1 + d^2(o, Y_s) + r^2(Y_t)) dt, \quad d\langle r^2(Y) \rangle_t = 8r^2(Y_t) dt.$$

By Itô's formula,

$$\begin{aligned} & de^{a(t)r^2(Y_t)} \\ & \leq dM_t + e^{a(t)r^2(Y_t)} \left[a(t)C'_T (1 + d^2(o, Y_s)) + \left(\frac{d}{dt}a(t) + C'_T a(t) + 4a(t)^2 \right) r^2(Y_t) \right] dt \end{aligned} \quad (4.4.21)$$

for some local martingale M_t . We choose a as the solution of

$$a'(u) + C'_T a(u) + 4a(u)^2 = 0, \quad a(s) := a_0 > 0. \quad (4.4.22)$$

Then a is nonincreasing on $[s, T]$. Hence [\(4.4.21\)](#) becomes

$$de^{a(u)r^2(Y_u)} \leq dM_u + C'_T a(u) (1 + d^2(o, Y_s)) e^{a(u)r^2(Y_u)} du. \quad (4.4.23)$$

After localization and conditioning on Y_s , combining Gronwall's inequality with $a(u) \leq a_0$ gives

$$\mathbb{E} \left[e^{a(t)d^2(Y_t, Y_s)} \mid Y_s \right] \leq \exp \left(C'_T a_0 (t - s) (1 + d^2(o, Y_s)) \right). \quad (4.4.24)$$

Taking expectations yields

$$\mathbb{E} \left[e^{a(t)d^2(Y_t, Y_s)} \right] \leq e^{C'_T a_0 (t-s)} \mathbb{E} \left[e^{C'_T a_0 (t-s)d^2(o, Y_s)} \right].$$

In particular, by choosing a_o small enough, say

$$0 < a_0 < \frac{\inf_{0 \leq \tau \leq T} \alpha(\tau)}{TC'_T}, \quad (4.4.25)$$

where α is the continuous function of [Lemma 4.19](#), we obtain

$$\mathbb{E} \left[e^{a(t)d^2(Y_t, Y_s)} \right] \leq 1 + C_T(t - s) \quad (4.4.26)$$

for some constant $C_T > 0$; one may take

$$C_T := e^{C'_T a_0 T} C'_T a_0 \left(1 + \frac{E_T - 1}{\underline{\alpha}_T} \right)$$

with

$$\underline{\alpha}_T := \inf_{0 \leq \tau \leq T} \alpha(\tau), \quad E_T := \sup_{0 \leq \tau \leq T} \int_M e^{\alpha_T d^2(o, x)} d\mu_\tau(x).$$

We now aim to obtain (4.4.26) with a supremum over $0 \leq t_0 \leq s \leq t \leq t_0 + \delta \leq T$ inside the expectation. Fix $t_0 \in [0, T - \delta]$, set

$$r := d(\cdot, Y_{t_0}), \quad S_u := \sup_{t_0 \leq v \leq u} e^{a(v)r^2(Y_v)}.$$

Applying (4.4.23) with $s = t_0$, integrating, and taking the supremum, we get, after localization,

$$S_u \leq 1 + \sup_{t_0 \leq \tau \leq u} |M_\tau - M_{t_0}| + C'_T a_0 (1 + d^2(o, Y_{t_0})) \int_{t_0}^u S_\tau d\tau.$$

Here the martingale part is given by

$$dM_u = 2\sqrt{2} a(u) r(Y_u) e^{a(u)r^2(Y_u)} d\beta_u,$$

and therefore

$$d\langle M \rangle_u \leq 8a_0^2 r^2(Y_u) e^{2a(u)r^2(Y_u)} du.$$

Hence, by the Burkholder–Davis–Gundy inequality,

$$\mathbb{E} \left[\sup_{t_0 \leq \tau \leq u} |M_\tau - M_{t_0}| \middle| Y_{t_0} \right] \leq C_{\text{BDG}} \sqrt{8a_0} \left(\int_{t_0}^u \mathbb{E} \left[r^2(Y_\tau) e^{2a(\tau)r^2(Y_\tau)} \middle| Y_{t_0} \right] d\tau \right)^{1/2}.$$

Choosing $a_0 > 0$ smaller if necessary, by applying (4.4.24) with a larger exponential parameter, we obtain

$$\mathbb{E} \left[r^2(Y_\tau) e^{2a(\tau)r^2(Y_\tau)} \middle| Y_{t_0} \right] \leq C_T (\tau - t_0) (1 + d^2(o, Y_{t_0})) e^{C_T \delta (1 + d^2(o, Y_{t_0}))}.$$

Consequently, after increasing C_T that we do not relabel,

$$\mathbb{E} \left[\sup_{t_0 \leq \tau \leq u} |M_\tau - M_{t_0}| \middle| Y_{t_0} \right] \leq C_T \delta (1 + d^2(o, Y_{t_0})) e^{C_T \delta (1 + d^2(o, Y_{t_0}))}.$$

Hence,

$$\begin{aligned} & \mathbb{E} [S_u | Y_{t_0}] \\ & \leq 1 + C_T \delta (1 + d^2(o, Y_{t_0})) e^{C_T \delta (1 + d^2(o, Y_{t_0}))} + C_T (1 + d^2(o, Y_{t_0})) \int_{t_0}^u \mathbb{E} [S_\tau | Y_{t_0}] d\tau. \end{aligned}$$

By Gronwall's inequality, after increasing C_T ,

$$\mathbb{E} \left[\sup_{t_0 \leq u \leq t_0 + \delta} e^{a(u)d^2(Y_u, Y_{t_0})} \middle| Y_{t_0} \right] \leq 1 + C_T \delta (1 + d^2(o, Y_{t_0})) e^{C_T \delta (1 + d^2(o, Y_{t_0}))}.$$

For any $s, t \in [t_0, t_0 + \delta]$, the triangle inequality gives

$$d^2(Y_t, Y_s) \leq 2d^2(Y_t, Y_{t_0}) + 2d^2(Y_s, Y_{t_0}) \leq 4 \sup_{t_0 \leq u \leq t_0 + \delta} d^2(Y_u, Y_{t_0}).$$

Moreover, the solution of (4.4.22) admits a strictly positive lower bound $a_* = a_*(T) > 0$, uniform in $t_0 \in [0, T]$. Hence, setting $a_T := \frac{a_*}{4}$, we obtain

$$\sup_{t_0 \leq s \leq t \leq t_0 + \delta} e^{a_T d^2(Y_t, Y_s)} \leq \sup_{t_0 \leq u \leq t_0 + \delta} e^{a(u)d^2(Y_u, Y_{t_0})}.$$

Choosing $a_0 > 0$ small enough, the propagated square-exponential moment, together with the fact that $a(u)$ is uniformly bounded from below on intervals of length at most T yield, after possibly decreasing $a_T > 0$,

$$\mathbb{E} \left[\sup_{t_0 \leq s \leq t \leq t_0 + \delta} e^{a_T d^2(Y_t, Y_s)} \right] \leq 1 + C_T \delta.$$

for large enough $C_T > 0$.

□

4.4.4 Coupling the interacting and non interacting proxy

We first derive an estimate for Kendall's formula of two diffusions with different drifts using parallel coupling of the noises under the condition of Ricci lower bound. We now explain this parallel coupling.

Let X_t and Y_t solve

$$dX_t = \sqrt{2} dB_t^M + b_t^X(X_t) dt, \quad dY_t = \sqrt{2} d\bar{B}_t^M + b_t^Y(Y_t) dt,$$

where B_t^M and $\overline{B}_t^M$ are Brownian motions on M . Equivalently, on the orthonormal frame bundle $\pi : O(M) \rightarrow M$ in Stratonovich form, denoting by tildes horizontal lifts,

$$d\Xi_t = \sqrt{2} \sum_{k=1}^m H_k(\Xi_t) \circ dB_t^k + \widetilde{b}_t^X(\Xi_t) dt, \quad X_t = \pi(\Xi_t),$$

and

$$dU_t = \sqrt{2} \sum_{k=1}^m H_k(U_t) \circ d\overline{B}_t^k + \widetilde{b}_t^Y(U_t) dt, \quad Y_t = \pi(U_t).$$

Away from the joint cut locus, let $\gamma_t : [0, d(X_t, Y_t)] \rightarrow M$ be the unit-speed minimizing geodesic from X_t to Y_t , and denote by

$$P_{X_t \rightarrow Y_t} : T_{X_t} M \rightarrow T_{Y_t} M$$

parallel transport along γ_t . We couple the Brownian noises by parallel displacement. More precisely, for each pair (X_t^i, Y_t^i) we set

$$d\overline{B}_t^i = O_t^i dB_t^i, \quad O_t^i := (U_t^i)^{-1} P_{X_t^i \rightarrow Y_t^i} \Xi_t^i.$$

This parallel coupling is extended through the cut locus via the standard approximation procedure; see [107, Section 2.1]. Since the block-diagonal matrix $\text{diag}(O_t^1, \dots, O_t^N)$ is orthogonal, $(\overline{B}^1, \dots, \overline{B}^N)$ is again a standard $\mathbb{R}^{Nm}$ -valued Brownian motion. In particular, $\overline{B}^1, \dots, \overline{B}^N$ are independent, and consequently so are $Y^1, \dots, Y^N$, with common law μ_t . The Brownian part of the coupled process (X_t, Y_t) is therefore driven along the vector fields

$$\sqrt{2} (E_k(t), \overline{E}_k(t)), \quad k = 1, \dots, m,$$

where

$$E_k(t) := \Xi_t e_k, \quad \overline{E}_k(t) := P_{X_t \rightarrow Y_t} E_k(t).$$

Applying Kendall's formula [107, Theorem 2.3.2] to the distance function $\rho := d(\cdot, \cdot)$ on the product manifold $M \times M$ (see also [5], [57], and [66, Theorem 3.5.1]), one obtains

$$d\rho(X_t, Y_t) \leq \mathcal{I}_t dt + b_t^X(X_t) \cdot \nabla_x \rho(X_t, Y_t) dt + b_t^Y(Y_t) \cdot \nabla_y \rho(X_t, Y_t) dt, \quad (4.4.27)$$

where

$$\mathcal{I}_t := \sum_{k=1}^m \text{Hess } \rho_{(X_t, Y_t)} \left((E_k, \overline{E}_k)(t), (E_k, \overline{E}_k)(t) \right). \quad (4.4.28)$$

On the cut locus and diagonal, that is on the complement of $\text{cut} := \{(x, y) \in M^2 : x \notin \text{Cut}(y)\}$ and of the diagonal $\text{diag} := \{(x, x) \in M\}$, Kendall's formula (4.4.27) is understood in the distributional sense. Outside of the cut locus and the diagonal, (4.4.27) is

an equality. The fact that (4.4.27) is an inequality comes from the distributional Laplacian of ρ which creates a nonincreasing process only supported on the cut locus and the diagonal, as in (4.4.17)

Note that there is no martingale term due to the parallel coupling. Indeed,

$$dZ_t = \sqrt{2} \sum_{k=1}^m (E_k(t), \bar{E}_k(t)) \circ dB_t^k + (b_t^X(X_t), b_t^Y(Y_t)) dt,$$

and applying Itô's formula to ρ yields then

$$d\rho(Z_t) = \sqrt{2} \sum_{k=1}^m d\rho_{Z_t}(E_k(t), \bar{E}_k(t)) dB_t^k + \mathcal{I}_t dt + b_t^X(X_t) \cdot \nabla_x \rho(Z_t) dt + b_t^Y(Y_t) \cdot \nabla_y \rho(Z_t) dt$$

But

$$\nabla_x \rho(Z_t) = -\dot{\gamma}_t(0), \quad \nabla_y \rho(Z_t) = \dot{\gamma}_t(\rho(Z_t)),$$

and thus, since $\bar{E}_k(t) = P_{X_t \rightarrow Y_t} E_k(t)$,

$$\begin{aligned} d_x \rho(Z_t) E_k(t) + d_y \rho(Z_t) \bar{E}_k(t) &= -E_k(t) \cdot \dot{\gamma}_t(0) + \bar{E}_k(t) \cdot \dot{\gamma}_t(\rho(Z_t)) \\ &= -E_k(t) \cdot \dot{\gamma}_t(0) + E_k(t) \cdot \dot{\gamma}_t(0) \\ &= 0. \end{aligned}$$

We now estimate the index term (4.4.28) under the Ricci lower bound $\text{Ric} \geq K$ with $K \leq 0$. We choose $E_1(t) := \dot{\gamma}_t(0)$ so the radial direction does not contribute to the Hessian of ρ . Hence

$$\mathcal{I}_t = \sum_{k=2}^m \text{Hess}_{(X_t, Y_t)} \rho \left((E_k(t), \bar{E}_k(t)), (E_k(t), \bar{E}_k(t)) \right).$$

For $k = 2, \dots, m$, let J_k be the parallel vector field along γ_t such that

$$J_k(0) = E_k(t), \quad J_k(\rho(X_t, Y_t)) = \bar{E}_k(t).$$

By the index lemma [57, Lemma 6.7.1],

$$\text{Hess}_{(X_t, Y_t)} \rho \left((E_k(t), \bar{E}_k(t)), (E_k(t), \bar{E}_k(t)) \right) \leq I_{\gamma_t}(J_k, J_k),$$

where

$$I_{\gamma_t}(J, J) := \int_0^{\rho(X_t, Y_t)} (|\nabla_{\dot{\gamma}_t} J|^2 - \langle R(J, \dot{\gamma}_t) \dot{\gamma}_t, J \rangle) ds.$$

But J_k is parallel, and therefore,

$$I_{\gamma_t}(J_k, J_k) = - \int_0^{\rho(X_t, Y_t)} \langle R(J_k, \dot{\gamma}_t) \dot{\gamma}_t, J_k \rangle ds.$$

Summing over $k = 2, \dots, m$, and using that $\{J_k\}_{2 \leq k \leq m}$ form a parallel orthonormal frame orthogonal to $\dot{\gamma}_t$, we obtain

$$\mathcal{I}_t \leq - \int_0^{\rho(X_t, Y_t)} \sum_{k=2}^m \langle R(J_k, \dot{\gamma}_t) \dot{\gamma}_t, J_k \rangle ds = - \int_0^{\rho(X_t, Y_t)} \text{Ric}(\dot{\gamma}_t, \dot{\gamma}_t) ds \leq -K\rho(X_t, Y_t).$$

Consequently,

$$d\rho(X_t, Y_t) \leq [K|\rho(X_t, Y_t) + b_t^X(X_t) \cdot \nabla_x \rho(X_t, Y_t) + b_t^Y(Y_t) \cdot \nabla_y \rho(X_t, Y_t)] dt. \quad (4.4.29)$$

Proposition 4.22 (Coupling estimate). *Let*

$$\hat{\mu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}, \quad \hat{\nu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i},$$

where

$$\begin{aligned} dX_t^i &= \sqrt{2} dB_t^{i,M} - \nabla V(X_t^i) dt - \frac{1}{N} \sum_{j=1}^N \nabla_x W(X_t^i, X_t^{j,N}) dt, \\ dY_t^i &= \sqrt{2} d\bar{B}_t^{i,M} - \nabla V(Y_t^i) dt - \int_M \nabla_x W(Y_t^i, z) d\mu_t(z) dt. \end{aligned}$$

Assume that $X_0^i = Y_0^i$, and couple each pair (X_t^i, Y_t^i) by parallel coupling. Then in the setting of [Lemma 4.19](#), under the additional assumption [\(4.4.6\)](#) for every $t \geq 0$, it holds

$$W_1(\hat{\mu}_t^N, \mu_t) \leq W_1(\hat{\nu}_t^N, \mu_t) + L_{W,2} \int_0^t e^{-\alpha_M(t-s)} W_1(\hat{\nu}_s^N, \mu_s) ds, \quad (4.4.30)$$

where

$$\alpha_M := \beta + \tilde{\beta} - |K| - L_{W,2}.$$

Proof. We apply the parallel-coupling Kendall formula [\(4.4.29\)](#) to the pair (X_t^i, Y_t^i) ;

$$d\rho(X_t^i, Y_t^i) \leq [K|\rho(X_t^i, Y_t^i) + b_t^i(X_t^i) \cdot \nabla_x \rho(X_t^i, Y_t^i) + \bar{b}_t(Y_t^i) \cdot \nabla_y \rho(X_t^i, Y_t^i)] dt, \quad (4.4.31)$$

where

$$b_t^i(x) := -\nabla V(x) - \frac{1}{N} \sum_{j=1}^N \nabla_x W(x, X_t^j), \quad \bar{b}_t(x) := -\nabla V(x) - \nabla_x (W * \mu_t).$$

We first estimate the contribution of V . Since $\text{Hess } V \geq \beta$, integrating along the geodesic γ_i from X_t^i to Y_t^i yields

$$\nabla V(Y_t^i) \cdot \dot{\gamma}_i(\rho_i(t)) - P_{X_t^i \rightarrow Y_t^i} \nabla V(X_t^i) \cdot \dot{\gamma}_i(\rho_i(t)) \geq \beta \rho_i(t).$$

Equivalently,

$$-\nabla V(X_t^i) \cdot \nabla_x \rho(X_t^i, Y_t^i) - \nabla V(Y_t^i) \cdot \nabla_y \rho(X_t^i, Y_t^i) \leq -\beta \rho_i(t). \quad (4.4.32)$$

We now treat the interaction. Its contribution in (4.4.31) is given by

$$\begin{aligned} \mathcal{W}_i(t) &:= -\frac{1}{N} \sum_{j=1}^N \nabla_x W(X_t^i, X_t^j) \cdot \nabla_x \rho(X_t^i, Y_t^i) \\ &\quad - \int_M \nabla_x W(Y_t^i, z) d\mu_t(z) \cdot \nabla_y \rho(X_t^i, Y_t^i). \end{aligned}$$

After parallel transport, adding and subtracting vector fields in $T_{Y_t^i} M$, we deduce that

$$\begin{aligned} \mathcal{W}_i(t) &= \left(\frac{1}{N} \sum_{j=1}^N \left[P_{X_t^i \rightarrow Y_t^i} \nabla_x W(X_t^i, X_t^j) - \nabla_x W(Y_t^i, X_t^j) \right] \right) \cdot \nabla_y \rho(X_t^i, Y_t^i) \\ &\quad + \left(\frac{1}{N} \sum_{j=1}^N \left[\nabla_x W(Y_t^i, X_t^j) - \nabla_x W(Y_t^i, Y_t^j) \right] \right) \cdot \nabla_y \rho(X_t^i, Y_t^i) \\ &\quad + \left(\frac{1}{N} \sum_{j=1}^N \nabla_x W(Y_t^i, Y_t^j) - \int_M \nabla_x W(Y_t^i, z) d\mu_t(z) \right) \cdot \nabla_y \rho(X_t^i, Y_t^i). \end{aligned}$$

The lower Hessian bound $\text{Hess}_x W(\cdot, z) \geq \tilde{\beta}$, for every $z \in M$, yields

$$\left(P_{X_t^i \rightarrow Y_t^i} \nabla_x W(X_t^i, z) - \nabla_x W(Y_t^i, z) \right) \cdot \nabla_y \rho(X_t^i, Y_t^i) \leq -\tilde{\beta} \rho(X_t^i, Y_t^i),$$

the assumption (4.4.6) on $\nabla_x W$ gives

$$\left| \left(\frac{1}{N} \sum_{j=1}^N \left[\nabla_x W(Y_t^i, X_t^j) - \nabla_x W(Y_t^i, Y_t^j) \right] \right) \cdot \nabla_y \rho(X_t^i, Y_t^i) \right| \leq \frac{L_{W,2}}{N} \sum_{i=1}^N \rho(X_t^i, Y_t^i),$$

and furthermore using Kantorovich duality,

$$\begin{aligned} \left| \left(\frac{1}{N} \sum_{j=1}^N \nabla_x W(Y_t^i, Y_t^j) - \int_M \nabla_x W(Y_t^i, z) d\mu_t(z) \right) \cdot \nabla_y \rho(X_t^i, Y_t^i) \right| \\ \leq L_{W,2} W_1(\hat{\nu}_t^N, \mu_t). \end{aligned}$$

Hence,

$$\mathcal{W}_i(t) \leq -\tilde{\beta}\rho(X_t^i, Y_t^i) + L_{W,2} R_N(t) + L_{W,2} W_1(\hat{\nu}_t^N, \mu_t), \quad R_N(t) := \frac{1}{N} \sum_{i=j}^N \rho(X_t^j, Y_t^j). \quad (4.4.33)$$

Combining (4.4.31)–(4.4.33) and averaging over $i = 1, \dots, N$,

$$\frac{d}{dt} R_N(t) \leq -\left(\beta + \tilde{\beta} - |K| - L_{W,2}\right) R_N(t) + L_{W,2} W_1(\hat{\nu}_t^N, \mu_t).$$

In other words,

$$\frac{d}{dt} R_N(t) \leq -\alpha_M R_N(t) + L_{W,2} W_1(\hat{\nu}_t^N, \mu_t).$$

Since $X_0^i = Y_0^i$, we have $R_N(0) = 0$. Gronwall's inequality yields

$$R_N(t) \leq L_{W,2} \int_0^t e^{-\alpha_M(t-s)} W_1(\hat{\nu}_s^N, \mu_s) ds.$$

Finally, pairing each X_t^i with Y_t^i gives

$$W_1(\hat{\mu}_t^N, \hat{\nu}_t^N) \leq \frac{1}{N} \sum_{i=1}^N d(X_t^i, Y_t^i) = R_N(t).$$

Therefore, by the triangle inequality,

$$\begin{aligned} W_1(\hat{\mu}_t^N, \mu_t) &\leq W_1(\hat{\mu}_t^N, \hat{\nu}_t^N) + W_1(\hat{\nu}_t^N, \mu_t) \\ &\leq L_{W,2} \int_0^t e^{-\alpha_M(t-s)} W_1(\hat{\nu}_s^N, \mu_s) ds + W_1(\hat{\nu}_t^N, \mu_t), \end{aligned}$$

which proves (4.4.30). $\square$

From this point on, the proof of [Theorem 4.18](#) is the same as in Bolley–Guillin–Villani [11] by combining [Theorem 4.15](#), [Lemma 4.21](#) and [Proposition 4.22](#).

4.5 Appendix

Proof of [Lemma 4.8](#). Since the covering number $N(B(x, R), \delta)$ is bounded by the packing number $P(B(x, R), \delta)$, it is sufficient to estimate the latter. Let $\{x_i\}_{i=1}^{P(B(x, R), \delta)}$ be a maximal δ -packing of $B(x, R)$; in particular, all balls $B(x_i, \delta/2)$ are disjoint and contained in $B(x, R + \delta/2)$ and therefore

$$P(B(x, R), \delta) \leq \max_{x_i} \frac{\mathfrak{m}[B(x, R + \delta/2)]}{\mathfrak{m}[B(x_i, \delta/2)]}. \quad (4.5.1)$$

On the other hand, since $x_i \in B(x, R)$, we have $B(x, R) \subseteq B(x_i, 2R)$. Hence, by the Bishop–Gromov [Theorem 4.7](#),

$$\frac{\mathfrak{m}[B(x_i, \delta/2)]}{V_{K,n}(\delta/2)} \geq \frac{\mathfrak{m}[B(x_i, 2R)]}{V_{K,n}(2R)} \geq \frac{\mathfrak{m}[B(x, R)]}{V_{K,n}(2R)}, \quad (4.5.2)$$

Inserting (4.5.2) into (4.5.1), we obtain

$$\begin{aligned} P(B(x, R), \delta) &\leq \frac{\mathfrak{m}[B(x, R + \delta/2)]}{\mathfrak{m}[B(x, R)]} \frac{V_{K,n}(2R)}{V_{K,n}(\delta/2)} \\ &\leq \frac{V_{K,n}(R + \delta/2)}{V_{K,n}(R)} \frac{V_{K,n}(2R)}{V_{K,n}(\delta/2)}. \end{aligned} \quad (4.5.3)$$

after another application of Bishop–Gromov.

If $K = 0$, then $V_{0,n}(r)$ is proportional to r^n , and therefore

$$P(B(x, R), \delta) \leq \left(1 + \frac{\delta}{2R}\right)^n \left(\frac{4R}{\delta}\right)^n \leq \left(\frac{5R}{\delta}\right)^n,$$

since $\delta \leq R/2$. Otherwise, for $K < 0$, using the inequality

$$\frac{V_{K,n}(r_2)}{V_{K,n}(r_1)} = \frac{\int_0^{r_2} \sinh^{n-1} \left(\sqrt{\frac{|K|}{n-1}} s \right) ds}{\int_0^{r_1} \sinh^{n-1} \left(\sqrt{\frac{|K|}{n-1}} s \right) ds} \leq \left(\frac{r_2}{r_1}\right)^n \exp\left(\sqrt{|K|(n-1)}(r_2 - r_1)\right)$$

valid for all $0 < r_1 \leq r_2$ in (4.5.3), we obtain

$$P(B(x, R), \delta) \leq \left(\frac{5R}{\delta}\right)^n \exp\left(2\sqrt{|K|(n-1)}R\right).$$

□

Proof of [Proposition 4.9](#). Given $o \in E$, we truncate ν on the ball $B(o, R)$ as $\nu_R := \nu[R]^{-1} \nu|_{B(o, R)}$. Then we sample independent random variables $\{X^i\}_{i=1}^N$ and $\{Y^i\}_{i=1}^N$ according to ν and ν_R respectively. Notice that

$$X_R^i := \begin{cases} X^i & \text{if } d(o, X^i) < R; \\ Y^i & \text{otherwise,} \end{cases}$$

are independent and distributed according to ν_R . Denote by

$$\widehat{\nu}_R^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_R^i}, \quad \widehat{\pi}^N := \frac{1}{N} \sum_{i=1}^N \delta_{X^i} \otimes \delta_{X_R^i},$$

respectively the empirical measure associated to ν_R and the trivial coupling between $\widehat{\nu}^N$ and $\widehat{\nu}_R^N$. Then

$$W_p^p\left(\widehat{\nu}^N, \widehat{\nu}_R^N\right) \leq \int_{E^2} d^p(x, y) d\widehat{\pi}^N(x, y) \leq \frac{1}{N} \sum_{i=1}^N 2^p d^p(o, X^i) 1_{d(o, X^i) \geq R},$$

and by independence of the X^i 's,

$$\mathbb{E}\left[W_p^p\left(\widehat{\nu}^N, \widehat{\nu}_R^N\right)\right] \leq 2^p \mathbb{E}\left[d^p(o, X^1) 1_{d(o, X^1) \geq R}\right] = 2^p \int_{E \setminus B(o, R)} d^p(o, x) d\nu(x). \quad (4.5.4)$$

The expected Wasserstein distance between the truncated empirical measure and the truncated measure is bounded by the estimate (4.2.3);

$$\mathbb{E}\left[W_p\left(\widehat{\nu}_R^N, \nu_R\right)\right] \leq C \left(\varepsilon + N^{-\frac{1}{2p}} \int_{\varepsilon}^{R/2} N(B(o, R), \delta)^{\frac{1}{2p}} d\delta \right). \quad (4.5.5)$$

Moreover, since $X^1 \sim \nu$ and $X_R^1 \sim \nu_R$,

$$W_p^p(\nu, \nu_R) \leq \mathbb{E}\left[d^p(X^1, X_R^1)\right] \leq 2^p \int_{E \setminus B(o, R)} d^p(o, x) d\nu(x). \quad (4.5.6)$$

Thus, combining the triangle inequality with estimates (4.5.4), (4.5.5), and (4.5.6), we obtain

$$\begin{aligned} \mathbb{E}\left[W_p\left(\widehat{\nu}^N, \nu\right)\right] &\leq \mathbb{E}\left[W_p\left(\widehat{\nu}^N, \widehat{\nu}_R^N\right)\right] + \mathbb{E}\left[W_p\left(\widehat{\nu}_R^N, \nu_R\right)\right] + \mathbb{E}\left[W_p\left(\nu_R, \nu\right)\right] \\ &\leq 4 \left(\int_{E \setminus B(o, R)} d^p(o, x) d\nu(x) \right)^{\frac{1}{p}} + C \left(\varepsilon + N^{-\frac{1}{2p}} \int_{\varepsilon}^{R/2} N(B(o, R), \delta)^{\frac{1}{2p}} d\delta \right). \end{aligned} \quad (4.5.7)$$

Recall the covering estimate (4.2.4),

$$N(B(x, R), \delta) \leq \left(\frac{5R}{\delta} \right)^n \exp\left(2\sqrt{|K|(n-1)R}\right).$$

Then, denoting $\kappa := \sqrt{|K|(n-1)}$, we get

$$\int_{\varepsilon}^{R/2} N(B(o, R), \delta)^{\frac{1}{2p}} d\delta \lesssim_{K, n, p} \begin{cases} R & \text{if } K = 0, n < 2p; \\ R \log(R/\varepsilon) & \text{if } K = 0, n = 2p; \\ R^{\frac{n}{2p}} \varepsilon^{1-\frac{n}{2p}} & \text{if } K = 0, n > 2p; \\ e^{\kappa R/p} R & \text{if } K < 0, n < 2p; \\ e^{\kappa R/p} R \log(R/\varepsilon) & \text{if } K < 0, n = 2p; \\ e^{\kappa R/p} R^{\frac{n}{2p}} \varepsilon^{1-\frac{n}{2p}} & \text{if } K < 0, n > 2p. \end{cases} \quad \begin{aligned} &(4.5.8) \\ &(4.5.9) \\ &(4.5.10) \\ &(4.5.11) \\ &(4.5.12) \\ &(4.5.13) \end{aligned}$$

Here, the subscripts of $\lesssim$ denote the dependence of the constant. Now, we optimize ε and R in terms of N to balance the first term of (4.5.7) depending on the decay of the probability measure with the second term of (4.5.7) estimated by (4.5.8)–(4.5.13) depending on K, n , and p . Note that if $M_q := \int_E d^q(o, x) d\nu(x) < +\infty$ for some $q > p$, then

$$\left(\int_{E \setminus B(o, R)} d^p(o, x) d\nu(x) \right)^{\frac{1}{p}} \leq M_q^{1/p} R^{1-\frac{q}{p}}. \quad (4.5.14)$$

Case 1: Finite polynomial moment and $K = 0$.

i) If $n < 2p$, then inserting (4.5.8) and (4.5.14) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4M_q^{1/p} R^{1-\frac{q}{p}} + C_{n,p} \left(\varepsilon + N^{-\frac{1}{2p}} R \right).$$

Setting $R := C_{n,p,M_q} N^{\frac{1}{2q}}$ and $\varepsilon := c(N) \ll R^{1-\frac{q}{p}}$, it holds

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{n,p,M_q} N^{-\frac{q-p}{2pq}}.$$

ii) If $n = 2p$, then inserting (4.5.9) and (4.5.14) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4M_q^{1/p} R^{1-\frac{q}{p}} + C \left(\varepsilon + N^{-\frac{1}{2p}} R \log \left(\frac{R}{\varepsilon} \right) \right).$$

Setting $\varepsilon := CN^{-\frac{1}{2p}} R$ and $R := C_{n,p,q,M_q} N^{\frac{1}{2q}} (\log N)^{-\frac{p}{q}}$, it holds

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{n,p,q,M_q} N^{-\frac{q-p}{2pq}} (\log N)^{\frac{q-p}{q}}.$$

iii) If $n > 2p$, then inserting (4.5.10) and (4.5.14) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4M_q^{1/p} R^{1-\frac{q}{p}} + C \left(\varepsilon + N^{-\frac{1}{2p}} R^{\frac{n}{2p}} \varepsilon^{1-\frac{n}{2p}} \right).$$

Setting $\varepsilon := C_{n,p} N^{-\frac{1}{n}} R$ and $R := C_{n,p,q} N^{\frac{p}{nq}}$, it holds

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{n,p,q} N^{-\frac{q-p}{nq}}.$$

Case 2: Finite polynomial moment and $K < 0$.

iv) If $n < 2p$, then inserting (4.5.11) and (4.5.14) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4M_q^{1/p} R^{1-\frac{q}{p}} + C \left(\varepsilon + N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)}R/p} R \right). \quad (4.5.15)$$

Balancing the first and third term of (4.5.15) yields

$$R^{-q/p} \asymp N^{-1/(2p)} e^{\sqrt{|K|(n-1)} R/p}.$$

Taking logarithms,

$$\frac{\sqrt{|K|(n-1)}}{p} R + \frac{q}{p} \log R \approx \frac{1}{2p} \log N + \log C_{K,n,p,q}.$$

This suggests the ansatz $R := A \log N + B \log \log N$. We thus set

$$R := \frac{1}{2\sqrt{|K|(n-1)}} (\log N - 2q \log \log N).$$

For N large enough, we have both

$$\begin{aligned} N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)} R/p} R &= (\log N)^{-q/p} R \\ &\lesssim_{K,n,p,q} (\log N)^{-\frac{q-p}{p}}. \end{aligned}$$

and

$$R^{1-\frac{q}{p}} \lesssim_{K,n,p,q} (\log N)^{-\frac{q-p}{p}}.$$

Therefore, choosing $\varepsilon := c(N) \ll R^{1-\frac{q}{p}}$, for N large enough, it holds

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,q} (\log N)^{-\frac{q-p}{p}}.$$

v) If $n = 2p$, then inserting (4.5.12) and (4.5.14) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4M_q^{1/p} R^{1-\frac{q}{p}} + C \left(\varepsilon + N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)} R/p} R \log \left(\frac{R}{\varepsilon} \right) \right).$$

Setting $\varepsilon := C_{K,n,p,q} R^{1-\frac{q}{p}} < R/4$ for N large enough, the choice

$$R := \frac{1}{2\sqrt{|K|(n-1)}} (\log N - (2q + 2p) \log \log N)$$

gives the ansatz approximation to

$$\frac{\sqrt{|K|(n-1)}}{p} R + \frac{q}{p} \log R + \log \log R \approx \frac{1}{2p} \log N + \log C_{K,n,p,q},$$

and yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,q} (\log N)^{-\frac{q-p}{p}}.$$

vi) If $n > 2p$, then inserting (4.5.13) and (4.5.14) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4M_q^{1/p} R^{1-\frac{q}{p}} + C \left(\varepsilon + N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)}R/p} R^{\frac{n}{2p}} \varepsilon^{1-\frac{n}{2p}} \right).$$

Setting $\varepsilon := C_{K,n,p,q} R^{1-\frac{q}{p}} < R/4$ for N large enough, the choice

$$R := \frac{1}{2\sqrt{|K|(n-1)}} \left(\log N - \frac{nq}{p} \log \log N \right),$$

gives the ansatz approximation to

$$\frac{\sqrt{|K|(n-1)}}{p} R + \frac{nq}{2p^2} \log R \approx \frac{1}{2p} \log N + \log C_{K,n,p,q}$$

and also yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,q} (\log N)^{-\frac{q-p}{p}}.$$

Now, we treat probability measures with finite exponential moments. If

$$\mathcal{E}_\alpha := \int_E e^{\alpha d(o,x)} d\nu(x) < +\infty$$

for some $\alpha > 0$, then for $R \geq p/\alpha$,

$$\left(\int_{E \setminus B(o,R)} d^p(o,x) d\nu(x) \right)^{\frac{1}{p}} \leq \mathcal{E}_\alpha^{1/p} R e^{-\frac{\alpha}{p} R}. \quad (4.5.16)$$

Case 3: Finite exponential moment and $K = 0$.

vii) If $n < 2p$, then inserting (4.5.8) and (4.5.16) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4\mathcal{E}_\alpha^{1/p} R e^{-\frac{\alpha}{p} R} + C_{n,p} \left(\varepsilon + N^{-\frac{1}{2p}} R \right).$$

Setting $R := \frac{1}{2\alpha} \log N$ and $\varepsilon := c(N) \ll R e^{-\frac{\alpha}{p} R}$, it holds

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{n,p,\alpha,\mathcal{E}_\alpha} N^{-\frac{1}{2p}} \log N.$$

viii) If $n = 2p$, then inserting (4.5.9) and (4.5.16) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4\mathcal{E}_\alpha^{1/p} R e^{-\frac{\alpha}{p} R} + C \left(\varepsilon + N^{-\frac{1}{2p}} R \log \left(\frac{R}{\varepsilon} \right) \right).$$

Setting $\varepsilon := C_{n,p,\alpha,\varepsilon_\alpha} R e^{-\frac{\alpha}{p}R} < R/4$ for N large enough, the choice

$$R := \frac{1}{2\alpha} (\log N - 2p \log \log N)$$

gives the ansatz approximation to

$$\frac{\alpha}{p}R + \log R \approx \frac{1}{2p} \log N + \log C_{n,p,\alpha,\varepsilon_\alpha}$$

so that

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{n,p,\alpha,\varepsilon_\alpha} N^{-\frac{1}{2p}} (\log N)^2.$$

ix) If $n > 2p$, then inserting (4.5.10) and (4.5.16) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4\varepsilon_\alpha^{1/p} R e^{-\frac{\alpha}{p}R} + C \left(\varepsilon + N^{-\frac{1}{2p}} R^{\frac{n}{2p}} \varepsilon^{1-\frac{n}{2p}} \right).$$

Setting $\varepsilon := C_{n,p} N^{-\frac{1}{n}} R$ and $R := \frac{p}{\alpha n} \log N$, it holds

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{n,p,\alpha,\varepsilon_\alpha} N^{-\frac{1}{n}} \log N.$$

Case 4: Finite exponential moment and $K < 0$.

x) If $n < 2p$, then inserting (4.5.11) and (4.5.16) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4\varepsilon_\alpha^{1/p} R e^{-\frac{\alpha}{p}R} + C \left(\varepsilon + N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)}R/p} R \right).$$

Setting $\varepsilon := c(N) \ll R e^{-\frac{\alpha}{p}R}$ and

$$R := \frac{1}{2(\alpha + \sqrt{|K|(n-1)})} \log N,$$

we deduce that for N large enough,

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,\alpha,\varepsilon_\alpha} N^{-\frac{\alpha}{2p(\alpha + \sqrt{|K|(n-1)})}} \log N.$$

xi) If $n = 2p$, then inserting (4.5.12) and (4.5.16) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4\varepsilon_\alpha^{1/p} R e^{-\frac{\alpha}{p}R} + C \left(\varepsilon + N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)}R/p} R \log \left(\frac{R}{\varepsilon} \right) \right).$$

Setting $\varepsilon := C_{K,n,p,\alpha,\varepsilon_\alpha} R e^{-\frac{\alpha}{p}R}$ and

$$R := \frac{1}{2\left(\alpha + \sqrt{|K|(n-1)}\right)} (\log N - 2p \log \log N),$$

we deduce that for N large enough,

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,\alpha,\varepsilon_\alpha} N^{-\frac{\alpha}{2p(\alpha + \sqrt{|K|(n-1)})}} (\log N)^{1 + \frac{\alpha}{\alpha + \sqrt{|K|(n-1)}}}.$$

xii) If $n > 2p$, then inserting (4.5.13) and (4.5.16) into (4.5.7) yields

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \leq 4\varepsilon_\alpha^{1/p} R e^{-\frac{\alpha}{p}R} + C \left(\varepsilon + N^{-\frac{1}{2p}} e^{\sqrt{|K|(n-1)}R/p} R^{\frac{n}{2p}} \varepsilon^{1-\frac{n}{2p}} \right).$$

Optimizing in ε for fixed R gives

$$\varepsilon := C_{K,n,p} R N^{-\frac{1}{n}} \exp \left(\frac{2\sqrt{|K|(n-1)}}{n} R \right).$$

Hence

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,\alpha,\varepsilon_\alpha} \varepsilon_\alpha^{1/p} R e^{-\frac{\alpha}{p}R} + R N^{-\frac{1}{n}} \exp \left(\frac{2\sqrt{|K|(n-1)}}{n} R \right).$$

Setting

$$R := \frac{p}{n\alpha + 2p\sqrt{|K|(n-1)}} \log N,$$

we deduce that for N large enough,

$$\mathbb{E} [W_p(\hat{\nu}^N, \nu)] \lesssim_{K,n,p,\alpha,\varepsilon_\alpha} N^{-\frac{\alpha}{n\alpha + 2p\sqrt{|K|(n-1)}}} \log N.$$

□

CHAPTER 5

Well-Posedness of the Multi-Species Landau Equation

This chapter is the content of the joint work [65] with Raphael Winter and Havva Yoldaş. "On the existence of solutions to the multi-species Landau equation." arXiv preprint arXiv:2512.17082 (2025).

5.1 Introduction

We study the spatially homogeneous multi-species Landau equation, a kinetic model for the time evolution of a plasma consisting of several distinct particle species, such as ions and electrons. The homogeneous multi-species Landau system, satisfied by the probability densities $f_i = f_i(t, v)$ with $(t, v) \in (0, \infty) \times \mathbb{R}^3$ for N species, is given by

$$\partial_t f_i = \sum_{j=1}^N Q_{ij}(f_i, f_j), \quad 1 \leq i \leq N, \quad (5.1.1)$$

where $Q_{ij}(f_i, f_j)$ is the Landau collision operator, derived by Landau in 1936 [69],

$$Q_{ij}(f_i, f_j)(v_i) = \frac{1}{m_i} \nabla_{v_i} \cdot \left(\int_{\mathbb{R}^3} c_{ij} A(v_i - v_j) \left(\frac{\nabla_{v_i}}{m_i} - \frac{\nabla_{v_j}}{m_j} \right) (f_i \otimes f_j)(v_i, v_j) dv_j \right).$$

Here, the constants m_i denote the particle masses for each species, and

$$A(z) := |z|^{2+\gamma} \Pi_z^\perp, \quad \Pi_z^\perp := \text{Id} - \frac{z \otimes z}{|z|^2}, \quad c_{ij} := \frac{|\log \Lambda_c| q_i^2 q_j^2 n_i n_j}{8\pi \varepsilon_0^2}, \quad (5.1.2)$$

where $|\log \Lambda_c|$ is the Coulomb logarithm, ε_0 is the vacuum permittivity, and q_i are the particle charges. The constants n_i are the number densities of each species per unit volume,

so we can assume f_i to be probability densities. The constant $\gamma \in [-3, 1]$ determines the singularity of the interaction kernel, the physically most relevant case describing the interaction of charged particles via Coulomb potentials corresponds to $\gamma = -3$. The Landau equation can be obtained in the *grazing collisions limit* from the Boltzmann equation, see e.g., [3, 27, 104, 71].

The breakthrough result of Guillen–Silvestre [52] establishes the global existence of solutions for the Landau–Coulomb equation, when f solves (5.1.1) for only one species. They achieve this by showing that the Fisher information

$$I(f) = \int_{\mathbb{R}^3} |\nabla \log f|^2 f \, dv \quad (5.1.3)$$

is non-increasing along the solution of the Landau equation. On the other hand, for the multi-species Landau equation only a couple of mathematically rigorous results are available.

As the multi-species Landau equation arises from (grazing) elastic collisions, it conserves, at least formally, the number of particles for all species, as well as the total momentum $\mathcal{P}$ and total energy $\mathcal{E}$. Therefore, we have the following conserved quantities,

$$\mathcal{P} := \sum_{i=1}^N \int_{\mathbb{R}^3} m_i v f_i(v) \, dv, \quad \mathcal{E} := \frac{1}{2} \sum_{i=1}^N \int_{\mathbb{R}^3} m_i |v|^2 f_i(v) \, dv. \quad (5.1.4)$$

Moreover, the Boltzmann entropy $\mathcal{H}$, given by

$$\mathcal{H}(F_N) := \sum_{i=1}^N \int_{\mathbb{R}^3} f_i \log f_i \, dv,$$

where $F_N := (f_1, \dots, f_N)$ is a solution to (5.1.1), is formally non-increasing in time.

In this chapter, we tackle the problem of global well-posedness for the spatially homogeneous multi-species Landau equation. To this end, we first derive a natural generalization of the Fisher information in the multi-species setting. Since the global equilibrium of a centered multi-species system is given by the Gibbs state

$$F_{\text{eq}}(\mathbf{v}_N) := \prod_{i=1}^N \mathcal{M}_{T, m_i}(v_i), \quad (5.1.5)$$

where $\mathbf{v}_N := (v_1, \dots, v_N)$, $T > 0$ is the temperature and $\mathcal{M}_{T, m_i}$ is the Maxwellian velocity distribution for particles of mass m_i ; i.e.,

$$\mathcal{M}_{T, m_i}(v) := \left(\frac{m_i}{2\pi T} \right)^{\frac{3}{2}} \exp \left(-\frac{m_i}{2T} |v|^2 \right), \quad (5.1.6)$$

it is natural to seek a multi-species Fisher information $\mathcal{I}$ of the form

$$\mathcal{I}(F_N; \mathbf{a}) = \sum_{i=1}^N a_i \int_{\mathbb{R}^3} |\nabla \log f_i|^2 f_i \, dv, \quad (5.1.7)$$

where $\mathbf{a} = (a_1, \dots, a_N)$ positive constants $a_i > 0$ to be chosen. We can only expect monotonicity of this functional (5.1.7) if the equilibrium (5.1.5) is a minimizer among all centered configurations with fixed total kinetic energy $\mathcal{E}$ defined in (5.1.4). Testing this condition on factorized functions with the same kinetic energy shows that, up to a multiplicative constant, the unique choice is $a_i = m_i^{-1}$. We therefore define

$$\mathcal{I}(F_N) := \sum_{i=1}^N \frac{1}{m_i} \int_{\mathbb{R}^3} |\nabla \log f_i|^2 f_i \, dv. \quad (5.1.8)$$

This functional is indeed a Lyapunov functional for the homogeneous multi-species Landau equation provided that $\gamma \in [-\sqrt{8}, -2]$ which is our first main result, see [Theorem 5.1](#). We recast the cross-interaction operators in a new set of variables (5.2.6) and differential operators (5.2.5). After this transformation, the operators can be treated similar to [52], and the constant in the log-Sobolev inequality without additional symmetry is responsible for the threshold value $\gamma_* := -\sqrt{8}$.

We can prove that the threshold γ_* is sharp and the monotonicity of the Fisher information no longer holds in general for $\gamma \in [-3, \gamma_*)$. More precisely, we construct explicit counterexamples for the two-species Landau system when the mass ratio between the two species is sufficiently large. We do so by investigating first the limiting equation where one species has infinite mass. The finite mass case can then be recovered by a continuity argument. The counterexamples are given in [Theorem 5.2](#).

The first two theorems pose the question whether the multi-species Landau system is indeed globally well-posed below the threshold value γ_* or if the symmetry under exchange of particles is necessary to obtain global well-posedness results using the information-theoretical approach. This is the main objective of our analysis. We are able to answer this question positively and provide a new method to construct Lyapunov functionals that do not require this symmetry assumption.

We consider the two-species Landau system where one species has infinite mass, i.e., the collisional dynamics of electrons and ions. While the counterexample in [Theorem 5.2](#) shows that the time derivative of the Fisher information along this dynamics is no longer under control, we construct a new Lyapunov functional $\Lambda(f)$, defined in (5.1.12), by combining a weighted spherical Fisher information, the classical Fisher information and the Boltzmann entropy. With this new functional, we establish global well-posedness for $\gamma \in [-3, \gamma_*)$ in [Theorem 5.3](#). This result reveals a new method that can show global well-posedness for equations which do not satisfy the symmetry property required by the existing methods.

5.1.1 Main results

Theorem 5.1 (Global well-posedness). *Let $N \in \mathbb{N}$ be the number of species and $m_i > 0$, $1 \leq i \leq N$ their respective particle masses. Assume that the interaction is soft with $-2 \geq \gamma \geq -\sqrt{8}$. Further assume that the initial distributions satisfy $f_i^\circ \in L_{2\ell}^1 \cap L_\ell^2 \cap L \log L$ for some $\ell \geq 7$ and are probability densities. Then there exists a global-in-time strong solution $F_N = (f_i)_{i=1}^N$ which solves (5.1.1), and*

$$f_i \in C^\infty((0, \infty) \times \mathbb{R}^3), \quad \text{for } 1 \leq i \leq N.$$

Moreover, for any $t > 0$, the generalized Fisher information $\mathcal{I}(F_N)$ defined in (5.1.8) is finite and non-increasing in time on $(0, \infty)$.

The proof of this theorem is the content of [Section 5.2.4](#).

The theorem covers a large range of soft potentials, down to the threshold $\gamma_* = -\sqrt{8}$. This naturally raises the question whether the monotonicity of the Fisher information persists for $\gamma < \gamma_*$. We answer this question negatively in the cases where the mass ratio between the two species is sufficiently large or infinite. This includes physically important cases such as the collisional dynamics of electrons interacting with ions, where the mass ratio is about 1836 between protons and electrons. In the limiting case of infinite mass ratio, the resulting equation can be found in [88, Equation 6.4.11] and reads

$$\partial_t f = Q_\gamma(f, f) + c_{ei} Lf, \quad (5.1.9)$$

with the choice $\gamma = -3$ and $c_{ei} > 0$ in the general formula

$$\begin{aligned} Q_\gamma(f, f)(v) &:= \nabla_v \cdot \int_{\mathbb{R}^3} A(v - v^*) (\nabla_v - \nabla_{v^*}) (f \otimes f)(v, v^*) dv^*, \\ Lf(v) &:= \nabla_v \cdot (|v|^{2+\gamma} \Pi_v^\perp \nabla_v f(v)). \end{aligned}$$

In other words, the interaction with ions induces a diffusion on spheres in velocity space. We prove that the Fisher information is a Lyapunov function for this equation if the initial datum f° is even, but is *not* monotone decreasing in general.

Theorem 5.2 (Non-monotonicity of the Fisher information below γ_*). *Let $-3 \leq \gamma < \gamma_* = -\sqrt{8}$.*

(i). Infinite mass ratio: *There exists a probability density $f^\circ \in \mathcal{S}(\mathbb{R}^3)$ such that the local-in-time solution f with initial data f° to (5.1.9) satisfies*

$$\frac{d}{dt} \mathcal{I}(f) \big|_{t=0} > 0,$$

where $\mathcal{I}(f)$ is defined in (5.1.3). On the other hand, if f° is even, i.e. $f^\circ(v) = f^\circ(-v)$, then the Fisher information $\mathcal{I}(f)$ is a Lyapunov functional for (5.1.9) for any $-3 \leq \gamma \leq -2$.

(ii). **Large but finite mass ratio:** For $N = 2$, $c_{ij} > 0$, $i, j \in \{1, 2\}$, and $\frac{m_2}{m_1} > 0$ large enough, there exist probability densities $f_1^\circ, f_2^\circ \in \mathcal{S}(\mathbb{R}^3)$ such that the local-in-time solution $F_2 = (f_1, f_2)$ with initial data f_1°, f_2° to (5.1.1) has increasing Fisher information (5.1.8)

$$\frac{d}{dt} \mathcal{I}(F_2) \big|_{t=0} > 0.$$

The proof of this theorem is the content of [Section 5.3.1](#).

This theorem shows that the Fisher information is no longer a Lyapunov functional. Moreover, the proof reveals that the time derivative of the Fisher information cannot be controlled in terms of $I(f)$ alone, which makes any direct growth estimate for $I(f)$ inaccessible. This prompts the question whether (5.1.9) is in fact globally well-posed. We answer this positively even for Coulomb interactions between electrons and ions. We consider the equation

$$\partial_t f = Q_\gamma(f, f) + c_{ei} \nabla \cdot \left(\frac{\Pi_v^\perp}{|v|} \nabla f \right), \quad (5.1.10)$$

where $Q_\gamma(f, f)$ is given by the single-species Landau operator with coefficient $\gamma \in [-3, 2]$ in (5.1.2).

To obtain a new Lyapunov functional, we introduce, in addition to the standard Fisher information, a weighted spherical Fisher information $J(f)$ defined by

$$J(f) := \int_{\mathbb{R}^3} \frac{|\Pi_v^\perp \nabla \log f(v)|^2}{|v|} f(v) dv.$$

Furthermore, we introduce the relative Boltzmann entropy $H(f|\mathcal{M}_{T^\circ,1})$ with respect to the Maxwellian $\mathcal{M}_{T^\circ,1}$ (5.1.6) where the temperature T° is chosen such that

$$\int_{\mathbb{R}^3} \frac{1}{2} |v|^2 \mathcal{M}_{T^\circ,1}(v) dv = \int_{\mathbb{R}^3} \frac{1}{2} |v|^2 f^\circ(v) dv = E.$$

The relative entropy $H(f|\mathcal{M}_{T^\circ,1})$ can be written explicitly as

$$H(f|\mathcal{M}_{T^\circ,1}) := H(f) - \int_{\mathbb{R}^3} f \log \mathcal{M}_{T^\circ,1} dv := \int_{\mathbb{R}^3} f \log f dv - \int_{\mathbb{R}^3} f \log \mathcal{M}_{T^\circ,1} dv.$$

We are now in the position to state the central result of the chapter.

Theorem 5.3 (New Lyapunov functional for the infinite-mass case). *Let $-3 \leq \gamma < -\sqrt{8}$, $c_{ei} > 0$ and f be the (local) strong solution to (5.1.10) with initial datum given by a*

probability density f° satisfying $f^\circ \in L_{2\ell}^1 \cap L_\ell^2 \cap L \log L$ for some $\ell \geq 7$ and $J(f^\circ) < \infty$. In the case of $\gamma = -3$, we need to further assume that c_{ei} and the initial datum f° satisfy

$$1 + \frac{J(f^\circ)}{I^{\frac{3}{2}}(f^\circ)} + H(f^\circ | \mathcal{M}_{T^\circ,1}) \leq \frac{c_{ei}}{K}, \quad (5.1.11)$$

where $K > 0$ is a constant (computable in terms of the optimal constants in several functional inequalities). Then there exist constants $a = a(c_{ei}, \gamma, I(f^\circ), J(f^\circ), H(f^\circ | \mathcal{M}_{T^\circ,1})) > 0$ and $R = R(c_{ei}, \gamma, I(f^\circ), J(f^\circ), H(f^\circ | \mathcal{M}_{T^\circ,1})) > 0$ such that the non-negative functional

$$\Lambda(f) = I(f) + a J(f) + R H(f | \mathcal{M}_{T^\circ,1}) \quad (5.1.12)$$

is non-increasing in time. In particular, the Fisher information $I(f)$ remains globally bounded by a constant depending on the initial datum f° , and the solution to (5.1.10) exists globally-in-time.

Remark 5.4. For the condition (5.1.11) to hold, it is necessary that $c_{ei} > K$. The constant $K > 0$ constructed in our proof is rather large see [Remark 5.22](#).

The proof of this theorem is the content of [Section 5.3.3](#).

Let us emphasize that this result shows that the global-in-time boundedness of the Fisher information persists, even though the Fisher information itself is no longer monotone for $\gamma < \gamma_*$. For general finite mass-ratio of two species, the question of well-posedness of the two-species system remains open when $\gamma < \gamma_*$. In the linearized setting and large mass ratio, the problem is related to the two-scale problem considered in [50].

5.1.2 Summary of previous results

In this section, we summarize the theory of the Landau equation with a particular focus on the results after [52]. We refer the reader to [52] for a detailed review and historical remarks. The global well-posedness theory for the spatially homogeneous Landau equation is well-established for hard potentials and Maxwellian molecules, i.e., $\gamma \geq 0$, see e.g., [104] for $\gamma = 0$ and [29, 30] for $\gamma > 0$. For moderately soft potentials, i.e., $\gamma \in (-2, 0]$, regularity estimates developed in [48, 90, 112] imply global smooth solutions. For the very soft potentials, i.e., $\gamma \in [-3, -2)$ including the case of Coulomb interactions, this problem was open until the recent breakthrough result by Guillen–Silvestre [52]. They prove that the Fisher information is monotone decreasing for the spatially homogeneous Landau equation with Coulomb interactions, which gives the existence of global smooth solutions. Since then, this new information-theoretical approach has been used to address other open problems in kinetic theory. We refer the reader to the recent notes by Villani [103] on this topic. In [59], Imbert–Silvestre–Villani generalized this method to show the existence of global smooth solutions to the space-homogeneous Boltzmann equation in the regime of very soft potentials. The Fisher information approach has been generalized to initial data without

finite Fisher information, see [16, 28, 44, 61]. Let us briefly discuss the Lenard–Balescu equation which is considered as a more accurate model than the Landau equation for weakly coupled systems with long-range forces (especially Coulomb plasmas). The presence of collective effects leads to an additional non-local nonlinearity in the collision operator and currently no Fisher information based theory is known. The state of the art regarding the spatially homogeneous Lenard–Balescu equation is due to [32, 33, 95].

The global well-posedness of the spatially in-homogeneous Landau equation with Coulomb interactions still remains open for general initial data. However, it is possible to obtain rigorous results when the collision operator is mollified in space. In [49], the authors study a fuzzy version of the spatially in-homogeneous Landau equation where the collisions are delocalized via a spatially dependent kernel $\kappa(x - x_*)$. They prove the global well-posedness of the fuzzy Landau equation for moderately soft interactions where $\gamma \in (-2, 0]$ and a positive spatial kernel $\kappa(x) \sim \langle x \rangle^{-\lambda}$, $\lambda > 0$ when the initial data is sufficiently smooth. For $\kappa \equiv 1$, and $\gamma \in [-3, 1]$, they also prove that the spatial Fisher information decays monotonically and the full Fisher information remains uniformly bounded in finite time intervals. We also refer to the recent article series [35, 36, 34] for more details on the fuzzy Landau equation.

When it comes to the multi-species Landau equation, mathematical results are rather scarce. Classical references for the derivation of (5.1.1) are [73, Chapter 4] and [88, Chapter 6]. This problem is closely related to the study of the multi-species Boltzmann equation; thus, the existing results use similar approaches. Let us start by summarizing some landmark results in this case. In [92], the authors study the Boltzmann equation for a mixture of two gases in one spatial dimension, and prove sharp pointwise-in- (t, x) asymptotics where one species is near vacuum and the other one is near Maxwellian equilibrium. One of the first spectral gap results for the spatially in-homogeneous multi-species Boltzmann system on $(x, v) \in \mathbb{T}^3 \times \mathbb{R}^3$ covering hard potentials and Maxwellian molecules, $\gamma \in [0, 1]$, with Grad's cutoff assumption is [23]. This work uses a multi-species adaptation of the quantitative hypocoercivity techniques developed in [81]. However, [23] does not allow species to have different masses in the mixture. In [15], the authors develop a perturbative Cauchy theory for the multi-species Boltzmann equation posed on $(x, v) \in \mathbb{T}^3 \times \mathbb{R}^3$ in the weighted space $L_v^1 L_x^\infty$ with a polynomial weight. They prove the existence of a spectral gap for the linear multi-species Boltzmann operator allowing different masses which induce a loss of symmetry and the standard methods developed for the mono-species case are not immediately available. They study the perturbed linear equation by means of the $L^1 - L^\infty$ theory due to [53].

To the best of our knowledge, there are only two works on the Cauchy problem of the multi-species Landau system. In [51], the authors study the system of the spatially in-homogeneous Landau system defined in the phase space $(x, p) \in \mathbb{T}^3 \times \mathbb{R}^3$ with moderately soft potentials, $\gamma \in [-2, 1]$. Using a similar strategy as in [23], they provide explicit spectral gap and hypocoercivity estimates for a linearized multi-species Landau system where all species are assumed to be close to equilibrium. In [75], the authors study the Landau

system on $(x, p) \in \mathbb{R}^3 \times \mathbb{R}^3$ for $\gamma \in [-2, 1]$ for two species where one species starts near vacuum and the other one starts near a Maxwellian equilibrium state. Using a similar approach as in the case of the multi-species Boltzmann system [92], they show that the solutions to the linearized system become instantaneously smooth both in x and p without requiring any smoothness assumption on the initial datum. On the numerics side, [19] develops a deterministic, structure-preserving particle method for the spatially homogeneous, multi-species Landau equation by regularizing the collision operator so that each species distribution can be approximated as a sum of Dirac masses whose locations evolve according to an ODE system.

5.1.3 Preliminaries and structure of the chapter

Throughout the chapter, we follow the notational convention introduced in [52] and write $\langle I'(f), g \rangle$ for the Gateaux derivative of a functional I in direction g . Moreover, we recall the Japanese bracket $\langle v \rangle = \sqrt{1 + |v|^2}$ and the following standard notation for weighted spaces.

Definition 5.5 (Weighted spaces). *We work in weighted spaces L_t^2 and H_t^k , $l \geq 0$, $k \in \mathbb{N}$, defined by*

$$\begin{aligned} \|f\|_{L_t^2}^2 &= \int_{\mathbb{R}^3} f^2(v) \langle v \rangle^l dv, \\ \|f\|_{H_t^k}^2 &= \sum_{j=0}^k \|\nabla^j f\|_{L_t^2}^2. \end{aligned}$$

To control the tail of f , we also use the moments

$$\|f\|_{L_t^1} = \int_{\mathbb{R}^3} f(v) \langle v \rangle^l dv.$$

We introduce the number densities n_i in (5.1.2) so that we can assume throughout the chapter that f, f_i are probability densities.

We use the standard notation for the Frobenius product of real matrices and the induced Hilbert-Schmidt norm

$$A : B := \sum_{i,j} A_{ij} B_{ij} = \text{trace}(A^t B) \quad \text{and} \quad \|A\|_{\text{HS}}^2 := A : A.$$

For convenience, we will use the tensor product notation for functions f and g ,

$$(f \otimes g)(v, w) := f(v)g(w),$$

and suppress the arguments wherever there is no risk of confusion.

Let us briefly summarize the structure of the chapter. In [Section 5.2](#) we treat the finite-mass case and prove [Theorem 5.1](#). In [Section 5.3](#) we consider the case of two species with large or infinite mass ratio $\frac{m_2}{m_1}$. [Section 5.3.2](#) contains preliminary results on the dissipation functionals for the infinite mass case, while the proof of [Theorem 5.2](#) is carried out in [Section 5.3.1](#). The novel dissipation functional and the proof of [Theorem 5.3](#) are the content of [Section 5.3.3](#). Some frequently used identities in spherical calculus are gathered in [Appendix 5.4.1](#).

5.2 The multi-species Landau system with finite masses

5.2.1 Lifting argument for the multi-species system

We perform a lifting procedure by doubling the number of variables as in [52]. To this end, we define

$$m_{N+i} := m_i, \quad c_{i,N+j} := c_{ij}, \quad 1 \leq i, j \leq N,$$

so that particles i and $N+i$ belong to the same species. We introduce the lifted linear operators acting on functions

$$\tilde{F}(\mathbf{v}_{2N}) : (\mathbb{R}^3)^{2N} \rightarrow \mathbb{R}, \quad \mathbf{v}_{2N} := (v_1, \dots, v_{2N}),$$

by

$$\tilde{Q}_{ij}(\tilde{F}) := \left(\frac{\nabla v_i}{m_i} - \frac{\nabla v_j}{m_j} \right) \cdot \left(c_{ij} A(v_i - v_j) \left(\frac{\nabla v_i}{m_i} - \frac{\nabla v_j}{m_j} \right) \tilde{F} \right)$$

with

$$1 \leq i \leq N, \quad N+1 \leq j \leq 2N,$$

and consider the linear evolution of the Cauchy problem

$$\begin{cases} \partial_t \tilde{F} = \tilde{Q}(\tilde{F}) = \sum_{\substack{1 \leq i \leq N \\ N+1 \leq j \leq 2N}} \tilde{Q}_{ij}(\tilde{F}), \\ \tilde{F}^\circ(v_1, \dots, v_{2N}) = \prod_{i=1}^N f_i^\circ(v_i) f_i^\circ(v_{N+i}). \end{cases} \quad (5.2.1)$$

Since the lifted equation and its initial datum are invariant under the exchange of the two copies, uniqueness of the lifted linear evolution implies

$$\Pi_i \tilde{F}(t) = \Pi_{N+i} \tilde{F}(t), \quad i = 1, \dots, N. \quad (5.2.2)$$

The evolution (5.2.1) is useful because the equations for the marginals can be related back to the multi-species system (5.1.1). To see this, we consider the marginals

$$\Pi_k \tilde{F} := \int_{(\mathbb{R}^3)^{2N-1}} \tilde{F} \, d\hat{\mathbf{v}}_k, \quad 1 \leq k \leq 2N,$$

where $\hat{\mathbf{v}}_k$ denotes the vector $\mathbf{v}_{2N} := (v_1, \dots, v_{2N})$ with v_k omitted. For $1 \leq i \leq N$, we have

$$\partial_t (\Pi_i \tilde{F}) = \sum_{j=N+1}^{2N} \frac{\nabla_{v_i}}{m_i} \cdot \left(\int_{(\mathbb{R}^3)^{2N-1}} c_{ij} A(v_i - v_j) \left(\frac{\nabla_{v_i}}{m_i} - \frac{\nabla_{v_j}}{m_j} \right) \tilde{F}(v_1, \dots, v_{2N}) \, d\hat{\mathbf{v}}_i \right),$$

and thus, at time $t = 0$, we recover the multi-species equation (5.1.1),

$$\partial_t (\Pi_i \tilde{F}) \Big|_{t=0} = \sum_{j=1}^N Q_{ij}(f_i^\circ, f_j^\circ) = \partial_t f_i \Big|_{t=0}$$

For the lifted multi-species system we introduce the weighted Fisher information

$$\tilde{\mathcal{I}}(\tilde{F}) := \int_{(\mathbb{R}^3)^{2N}} \sum_{i=1}^{2N} \frac{1}{m_i} \left| \nabla_{v_i} \log \tilde{F} \right|^2 \tilde{F} \, d\mathbf{v}_{2N}.$$

Although this definition is asymmetric, the weighted Fisher information retains the key property that we need: it controls a suitably scaled sum of the marginals' Fisher information.

Lemma 5.6 (Dissipation inequality for the weighted Fisher information). *Let $\tilde{F} : (\mathbb{R}^3)^{2N} \rightarrow \mathbb{R}$ be a probability density with finite weighted Fisher information $\tilde{\mathcal{I}}(\tilde{F})$ and marginals*

$$\Pi_i \tilde{F} = \Pi_{N+i} \tilde{F} = f_i, \quad 1 \leq i \leq N.$$

Then

$$\sum_{i=1}^N \frac{1}{m_i} I(f_i) \leq \frac{1}{2} \tilde{\mathcal{I}}(\tilde{F}), \quad (5.2.3)$$

with equality if and only if $\tilde{F} = (f_1 \otimes \dots \otimes f_N)^{\otimes 2}$. As a consequence, we have the inequality

$$\frac{d}{dt} \mathcal{I}(F_N) \Big|_{t=0} \leq \frac{1}{2} \frac{d}{dt} \tilde{\mathcal{I}}(\tilde{F}) \Big|_{t=0}, \quad (5.2.4)$$

for F_N and $\tilde{F}$ solutions to (5.1.1) and (5.2.1) respectively.

Proof. Consider the change of variables $\Phi : (\mathbb{R}^3)^{2N} \rightarrow (\mathbb{R}^3)^{2N}$ defined by

$$\Phi : (v_1, \dots, v_{2N}) \mapsto (\sqrt{m_1} v_1, \dots, \sqrt{m_{2N}} v_{2N}).$$

Then the push-forward density $G := |\det D\Phi|^{-1} \tilde{F} \circ \Phi^{-1}$ satisfies

$$\int_{(\mathbb{R}^3)^{2N}} |\nabla \log G|^2 G \, d\mathbf{v}_{2N} = \tilde{\mathcal{I}}(\tilde{F}).$$

The claim is then a direct consequence of Carlen's inequality [17, Theorem 3]. The factor $\frac{1}{2}$ stems from the fact that $\Pi_i \tilde{F} = f_i = f_{N+i} = \Pi_{N+i} \tilde{F}$, so each probability distribution appears twice as a marginal.

Finally, the inequality (5.2.4) follows from the same argument as in the single-species case:

$$\begin{aligned} \left. \frac{d}{dt} \mathcal{I}(F_N) \right|_{t=0} &= \sum_{i,j=1}^N \left\langle \frac{1}{m_i} I'(f_i^\circ), Q_{ij}(f_i^\circ, f_j^\circ) \right\rangle \\ &= \sum_{i=1}^N \left\langle \frac{1}{m_i} I'(f_i^\circ), \partial_t (\Pi_i \tilde{F}) \Big|_{t=0} \right\rangle \\ &= \frac{d}{dt} \left(\sum_{i=1}^N \frac{1}{m_i} I(\Pi_i \tilde{F}) - \frac{1}{2} \tilde{\mathcal{I}}(\tilde{F}) \right) \Big|_{t=0} + \frac{1}{2} \frac{d}{dt} \tilde{\mathcal{I}}(\tilde{F}) \Big|_{t=0} \\ &\leq \frac{1}{2} \frac{d}{dt} \tilde{\mathcal{I}}(\tilde{F}) \Big|_{t=0}, \end{aligned}$$

where the last inequality follows from (5.2.3) and (5.2.2) using that the initial datum is factorized. $\square$

5.2.2 Dissipation of the weighted Fisher information

This section is devoted to the proof of the dissipation of the weighted Fisher information for strong solutions. This is the fundamental tool to prove Theorem 5.1.

We introduce the following differential operators,

$$\mathbb{V}_{i,j} := \left(\frac{\nabla v_i}{\sqrt{m_i}}, \frac{\nabla v_j}{\sqrt{m_j}} \right), \quad \mathbb{W}_{i,j} := \frac{\nabla v_i}{m_i} - \frac{\nabla v_j}{m_j}. \quad (5.2.5)$$

Lemma 5.7. *Let $\tilde{F}(v_1, \dots, v_{2N}) = \otimes_{k=1}^{2N} f_k$ be a probability density with marginals f_k and finite weighted Fisher information $\tilde{\mathcal{I}}(\tilde{F})$. Then for any pair $(i, j) : 1 \leq i \leq N, N+1 \leq j \leq 2N$,*

$$\left\langle \tilde{\mathcal{I}}'(\tilde{F}), \tilde{Q}_{ij}(\tilde{F}) \right\rangle = \left\langle \tilde{\mathcal{I}}'_{ij}(f_i \otimes f_j), \tilde{Q}_{ij}(f_i \otimes f_j) \right\rangle,$$

where

$$\tilde{Q}_{ij}(f_i \otimes f_j) := \mathbb{W}_{i,j} \cdot (c_{ij} A(v_i - v_j) \mathbb{W}_{i,j}(f_i \otimes f_j))$$

and

$$\tilde{\mathcal{I}}_{ij}(f_i \otimes f_j) := \int_{(\mathbb{R}^3)^2} |\nabla_{i,j} \log(f_i \otimes f_j)|^2 (f_i \otimes f_j).$$

Proof. By linearity, it suffices to show that for any index $k \neq i, j$ and

$$\tilde{\mathcal{I}}_k(\tilde{F}) := \int_{(\mathbb{R}^3)^{2N}} \frac{1}{m_k} |\nabla_{v_k} \log \tilde{F}|^2 \tilde{F},$$

we have

$$\langle \tilde{\mathcal{I}}'_k(\tilde{F}), \tilde{Q}_{ij}(\tilde{F}) \rangle = 0.$$

A direct computation gives

$$\begin{aligned} & \langle \tilde{\mathcal{I}}'_k(\tilde{F}), \tilde{Q}_{ij}(\tilde{F}) \rangle \\ &= \int_{(\mathbb{R}^3)^{2N}} \frac{2}{m_k} \left(\nabla_{v_k} \left(\frac{\tilde{Q}_{ij}(\tilde{F})}{\tilde{F}} \right) \cdot \nabla_{v_k} \log \tilde{F} \right) \tilde{F} + \int_{(\mathbb{R}^3)^{2N}} \frac{1}{m_k} |\nabla_{v_k} \log \tilde{F}|^2 \tilde{Q}_{ij}(\tilde{F}) \\ &= \frac{1}{m_k} \int_{(\mathbb{R}^3)^{2N}} 2 \nabla_{v_k} \left(\frac{\tilde{Q}_{ij}(\tilde{F})}{\tilde{F}} \right) \cdot \nabla_{v_k} \tilde{F} + |\nabla_{v_k} \log \tilde{F}|^2 \tilde{Q}_{ij}(\tilde{F}). \end{aligned}$$

Since $\tilde{F}$ is factorized, $\tilde{Q}_{ij}(\tilde{F})/\tilde{F}$ does not depend on v_k while $\nabla_{v_k} \log \tilde{F}$ only depends on v_k . As a consequence, the above integrals vanish, which proves the claim. $\square$

Let $\tilde{F}$ be a solution to (5.2.1). At time $t = 0$, the initial data is factorized as $\tilde{F}^\circ = (f_1^\circ \otimes \cdots \otimes f_N^\circ)^{\otimes 2}$, and Lemma 5.7 allows us to rewrite each contribution to the dissipation as

$$\begin{aligned} \partial_t \tilde{\mathcal{I}}(\tilde{F}) \Big|_{t=0} &= \sum_{\substack{1 \leq i \leq N \\ N+1 \leq j \leq 2N}} \langle \tilde{\mathcal{I}}'(\tilde{F}^\circ), \tilde{Q}_{ij}(\tilde{F}^\circ) \rangle \\ &= \sum_{\substack{1 \leq i \leq N \\ N+1 \leq j \leq 2N}} \langle \tilde{\mathcal{I}}'_{ij}(f_i^\circ \otimes f_j^\circ), \tilde{Q}_{ij}(f_i^\circ \otimes f_j^\circ) \rangle. \end{aligned}$$

We introduce two new physical variables, namely the relative velocity and the center-of-momentum frame: for a fixed pair (i, j) we set

$$z := v_i - v_j, \quad z^* := \frac{m_i v_i + m_j v_j}{M_{ij}}, \quad \alpha_{ij} := \frac{1}{m_i} + \frac{1}{m_j} = \frac{M_{ij}}{m_i m_j}, \quad (5.2.6)$$

where $M_{ij} := m_i + m_j$. Then the transformation $(v_i, v_j) \mapsto (z, z^*)$ is volume-preserving and we obtain

$$\mathbb{V}_{i,j} \mapsto \left(\frac{\nabla_z}{\sqrt{m_i}} + \frac{\sqrt{m_i}}{M_{ij}} \nabla_{z^*}, -\frac{\nabla_z}{\sqrt{m_j}} + \frac{\sqrt{m_j}}{M_{ij}} \nabla_{z^*} \right), \quad \mathbb{V}_{i,j} \mapsto \alpha_{ij} \nabla_z.$$

Using these new variables and denoting $\underline{f}_{ij}(z, z^*) := (f_i \otimes f_j)(v_i, v_j)$ we have

$$|\mathbb{V}_{i,j} \log(f_i \otimes f_j)|^2 = \alpha_{ij} |\nabla_z \log \underline{f}_{ij}|^2 + \frac{1}{M_{ij}} |\nabla_{z^*} \log \underline{f}_{ij}|^2.$$

Therefore, we can decompose $\tilde{\mathcal{I}}_{ij}$ into three functionals

$$\begin{aligned} \tilde{\mathcal{I}}_{ij}(f_i \otimes f_j) &= \int_{\mathbb{R}^6} |\mathbb{V}_{i,j} \log(f_i \otimes f_j)|^2 (f_i \otimes f_j) \\ &= \int_{\mathbb{R}^6} \alpha_{ij} |\nabla_z \log \underline{f}_{ij}|^2 \underline{f}_{ij} + \int_{\mathbb{R}^6} \frac{1}{M_{ij}} |\nabla_{z^*} \log \underline{f}_{ij}|^2 \underline{f}_{ij} \\ &= \alpha_{ij} \int_{\mathbb{R}^6} |(\text{Id} - \Pi_z^\perp) \nabla_z \log \underline{f}_{ij}|^2 \underline{f}_{ij} + \alpha_{ij} \int_{\mathbb{R}^6} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \underline{f}_{ij} \\ &\quad + \int_{\mathbb{R}^6} \frac{1}{M_{ij}} |\nabla_{z^*} \log \underline{f}_{ij}|^2 \underline{f}_{ij} \\ &=: \tilde{\mathcal{R}}_{ij;\text{rel}}(\underline{f}_{ij}) + \tilde{\mathcal{J}}_{ij;\text{rel}}(\underline{f}_{ij}) + \tilde{\mathcal{I}}_{ij;\text{cmf}}(\underline{f}_{ij}). \end{aligned} \tag{5.2.7}$$

One should interpret functionals $\tilde{\mathcal{R}}_{ij;\text{rel}}(\underline{f}_{ij})$ and $\tilde{\mathcal{J}}_{ij;\text{rel}}(\underline{f}_{ij})$ as the radial and spherical parts of the weighted Fisher information in relative velocity respectively, and the third functional $\tilde{\mathcal{I}}_{ij;\text{cmf}}(\underline{f}_{ij})$ as the remaining weighted Fisher information in the center-of-momentum frame.

In the variables (5.2.6), the lifted multi-species Landau operator takes a simpler form, namely,

$$\begin{aligned} &\frac{\tilde{Q}_{ij}(f_i \otimes f_j)}{f_i \otimes f_j} \\ &:= \frac{\mathbb{V}_{i,j} \cdot (c_{ij} A(v_i - v_j) \mathbb{V}_{i,j}(f_i \otimes f_j))}{f_i \otimes f_j} \\ &= c_{ij} (\mathbb{V}_{i,j} \cdot (A \mathbb{V}_{i,j} \log(f_i \otimes f_j)) + \mathbb{V}_{i,j} \log(f_i \otimes f_j) \cdot A \mathbb{V}_{i,j} \log(f_i \otimes f_j)) \\ &= c_{ij} \alpha_{ij}^2 \left(\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log \underline{f}_{ij}) + |z|^{2+\gamma} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right), \end{aligned} \tag{5.2.8}$$

This expression takes an even simpler form in geometric notation using the spherical coordinates in the relative velocity z :

$$r := |z|, \quad \omega := \frac{z}{|z|} =: \hat{z}. \tag{5.2.9}$$

Defining $\bar{f}_{ij}(r, \omega, z^*) := \underline{f}_{ij}(z, z^*)$, in view of (5.4.1) and (5.4.5),

$$\frac{\tilde{Q}_{ij}(\underline{f}_i \otimes \underline{f}_j)}{\underline{f}_i \otimes \underline{f}_j} = c_{ij} \alpha_{ij}^2 r^\gamma \left(\Delta_{\mathbb{S}^2} \log \bar{f}_{ij} + |\nabla_{\mathbb{S}^2} \log \bar{f}_{ij}|^2 \right). \quad (5.2.10)$$

We are now set to compute the Gateaux derivative of each functional in the direction $\tilde{Q}_{ij}$.

Lemma 5.8. *Let $1 \leq i \leq N$ and $N+1 \leq j \leq 2N$, and let $f_i, f_{j-N} (:= f_j)$ be solutions of (5.1.1) given by Lemma 5.13. The Gateaux derivatives of the functionals $\tilde{\mathcal{R}}_{ij;rel}(\underline{f}_{ij})$, $\tilde{\mathcal{J}}_{ij;rel}(\underline{f}_{ij})$ and $\tilde{\mathcal{I}}_{ij;cmf}(\underline{f}_{ij})$ in the direction $\tilde{Q}_{ij}(\underline{f}_i \otimes \underline{f}_j)$ are given by*

$$\left\langle \tilde{\mathcal{I}}'_{ij;cmf}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle = -2c_{ij} \frac{\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_{z^*} \nabla_z \log \underline{f}_{ij} : (\nabla_{z^*} \nabla_z \log \underline{f}_{ij} A(z)) \underline{f}_{ij}. \quad (5.2.11)$$

$$\left\langle \tilde{\mathcal{J}}'_{ij;rel}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle = -2c_{ij} \alpha_{ij}^3 \int_{\mathbb{R}^6} |z|^{2+\gamma} \left\| (\Pi_z^\perp \nabla_z)^2 \log \underline{f}_{ij} \right\|_{\text{HS}}^2 \underline{f}_{ij}. \quad (5.2.12)$$

$$\begin{aligned} \left\langle \tilde{\mathcal{R}}'_{ij;rel}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle &= -2c_{ij} \alpha_{ij}^3 \times \\ &\int_{\mathbb{R}^6} |z|^{2+\gamma} \left(\left| \Pi_z^\perp \nabla_z (\hat{z} \cdot \nabla_z \log \underline{f}_{ij}) + \frac{\gamma}{2|z|} \Pi_z^\perp \nabla_z \log \underline{f}_{ij} \right|^2 - \frac{\gamma^2}{4|z|^2} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right) \underline{f}_{ij}. \end{aligned} \quad (5.2.13)$$

Proof. We start with the proof of (5.2.11). The Gateaux derivative is given by

$$\begin{aligned} \left\langle \tilde{\mathcal{I}}'_{ij;cmf}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle &= \frac{c_{ij} \alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} 2 \nabla_{z^*} \log \underline{f}_{ij} \cdot \nabla_{z^*} \left(\frac{\nabla_z \cdot (A(z) \nabla_z \underline{f}_{ij})}{\underline{f}_{ij}} \right) \underline{f}_{ij} \\ &\quad + |\nabla_{z^*} \log \underline{f}_{ij}|^2 \nabla_z \cdot (A(z) \nabla_z \underline{f}_{ij}) \\ &=: \mathsf{T}_1 + \mathsf{T}_2. \end{aligned}$$

Integrating by parts yields readily

$$\mathsf{T}_2 = -\frac{2c_{ij} \alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_z \nabla_{z^*} \log \underline{f}_{ij} : (\nabla_{z^*} \log \underline{f}_{ij} \otimes A(z) \nabla_z \log \underline{f}_{ij}) \underline{f}_{ij}.$$

For T_1 , we use the symmetry of A to rewrite it as

$$\begin{aligned}
 T_1 &= \frac{2c_{ij}\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_{z^*} \log \underline{f}_{ij} \cdot \nabla_{z^*} \left(\frac{\nabla_z \cdot (A(z) \nabla_z \underline{f}_{ij})}{\underline{f}_{ij}} \right) \underline{f}_{ij} \\
 &= \frac{2c_{ij}\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \left(\nabla_{z^*} \log \underline{f}_{ij} \cdot \nabla_{z^*} (\nabla_z \cdot (A(z) \nabla_z \log \underline{f}_{ij})) \right. \\
 &\quad \left. + \nabla_z \log \underline{f}_{ij} \cdot A(z) \nabla_z \log \underline{f}_{ij} \right) \underline{f}_{ij} \\
 &= -\frac{2c_{ij}\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_{z^*} \nabla_z \log \underline{f}_{ij} : (\nabla_{z^*} \nabla_z \log \underline{f}_{ij} A(z)) \underline{f}_{ij} \\
 &\quad + \frac{2c_{ij}\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_z \nabla_{z^*} \log \underline{f}_{ij} : (\nabla_{z^*} \log \underline{f}_{ij} \otimes A(z) \nabla_z \log \underline{f}_{ij}) \underline{f}_{ij},
 \end{aligned}$$

where, in the third line, we used integration by parts. Combining T_1 and T_2 gives (5.2.11).

Next, we prove the second identity (5.2.12). Using (5.2.8), we have

$$\begin{aligned}
 &\langle \tilde{\mathcal{J}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \rangle \\
 &= \alpha_{ij} \int_{\mathbb{R}^6} \left(2\Pi_z^\perp \nabla_z \log \underline{f}_{ij} \right. \\
 &\quad \cdot \nabla_z \left(c_{ij}\alpha_{ij}^2 \left(\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log \underline{f}_{ij}) + |z|^{2+\gamma} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right) \right) \underline{f}_{ij} \\
 &\quad \left. + \alpha_{ij} \int_{\mathbb{R}^6} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right. \\
 &\quad \times \left(c_{ij}\alpha_{ij}^2 \left(\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log \underline{f}_{ij}) + |z|^{2+\gamma} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right) \right) \underline{f}_{ij}.
 \end{aligned}$$

Denoting $\bar{g}(r, \omega, z^*) := \underline{f}_{ij}(z, z^*)$ and recalling (5.2.10) and (5.4.1), this expression becomes

$$\begin{aligned}
 &\langle \tilde{\mathcal{J}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \rangle \\
 &= c_{ij}\alpha_{ij}^3 \int_0^\infty \int_{\mathbb{S}^2} 2r^\gamma \nabla_{\mathbb{S}^2} \log \bar{g} \cdot \nabla_{\mathbb{S}^2} \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} \, d\omega \, dr \\
 &\quad + c_{ij}\alpha_{ij}^3 \int_0^\infty \int_{\mathbb{S}^2} r^\gamma |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} \, d\omega \, dr.
 \end{aligned}$$

Integrating by parts yields

$$\begin{aligned} \langle \tilde{\mathcal{I}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \rangle &= c_{ij} \alpha_{ij}^3 \int_0^\infty \int_{\mathbb{S}^2} r^\gamma \left(2 |\nabla_{\mathbb{S}^2} \log \bar{g}|^4 + 3 \nabla_{\mathbb{S}^2} (|\nabla_{\mathbb{S}^2} \log \bar{g}|^2) \right. \\ &\quad \left. \cdot \nabla_{\mathbb{S}^2} \log \bar{g} - 2 (\Delta_{\mathbb{S}^2} \log \bar{g})^2 \right) \bar{g} \, d\omega \, dr. \end{aligned} \quad (5.2.14)$$

In order to simplify this expression, we first recall Bochner's formula on $\mathbb{S}^2$:

$$\frac{1}{2} \Delta_{\mathbb{S}^2} (|\nabla_{\mathbb{S}^2} \log \bar{g}|^2) = \Gamma_2(\log \bar{g}, \log \bar{g}) + \nabla_{\mathbb{S}^2} (\Delta_{\mathbb{S}^2} \log \bar{g}) \cdot \nabla_{\mathbb{S}^2} \log \bar{g}, \quad (5.2.15)$$

where Γ_2 is the carré du champs operator of the spherical Laplacian $\Delta_{\mathbb{S}^2}$:

$$\Gamma_2(\log \bar{g}, \log \bar{g}) = \|\nabla_{\mathbb{S}^2}^2 \log \bar{g}\|_{\text{HS}}^2 + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2.$$

Multiplying (5.2.15) by $\bar{g}$ and integrating by parts, we obtain

$$\begin{aligned} -\frac{1}{2} \int_{\mathbb{S}^2} \left(\nabla_{\mathbb{S}^2} (|\nabla_{\mathbb{S}^2} \log \bar{g}|^2) \cdot \nabla_{\mathbb{S}^2} \log \bar{g} \right) \bar{g} \\ = \int_{\mathbb{S}^2} \left(\|\nabla_{\mathbb{S}^2}^2 \log \bar{g}\|_{\text{HS}}^2 + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 + |\nabla_{\mathbb{S}^2} \log \bar{g}|^4 \right. \\ \left. + \nabla_{\mathbb{S}^2} (|\nabla_{\mathbb{S}^2} \log \bar{g}|^2) \cdot \nabla_{\mathbb{S}^2} \log \bar{g} - (\Delta_{\mathbb{S}^2} \log \bar{g})^2 \right) \bar{g}. \end{aligned}$$

This can be recast in the more convenient form,

$$\begin{aligned} \frac{3}{2} \int_{\mathbb{S}^2} \left(\nabla_{\mathbb{S}^2} (|\nabla_{\mathbb{S}^2} \log \bar{g}|^2) \cdot \nabla_{\mathbb{S}^2} \log \bar{g} \right) \bar{g} \\ = \int_{\mathbb{S}^2} \left((\Delta_{\mathbb{S}^2} \log \bar{g})^2 - \Gamma_2(\log \bar{g}, \log \bar{g}) - |\nabla_{\mathbb{S}^2} \log \bar{g}|^4 \right) \bar{g}. \end{aligned} \quad (5.2.16)$$

Finally, inserting (5.2.16) into (5.2.14) yields

$$\langle \tilde{\mathcal{I}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \rangle = -2c_{ij} \alpha_{ij}^3 \int_{\mathbb{R}^3} \int_0^\infty \int_{\mathbb{S}^2} r^\gamma \Gamma_2(\log \bar{g}, \log \bar{g}) \bar{g} \, d\omega \, dr \, dz^*. \quad (5.2.17)$$

The result follows recalling (5.4.8).

To prove (5.2.13), we compute the Gateaux derivative

$$\begin{aligned}
 & \left\langle \tilde{\mathcal{R}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle \\
 &= \alpha_{ij} \int_{\mathbb{R}^6} \left(2 (\text{Id} - \Pi_z^\perp) \nabla_z \log \underline{f}_{ij} \right. \\
 & \quad \cdot \nabla_z \left(c_{ij} \alpha_{ij}^2 \left(\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log \underline{f}_{ij}) + |z|^{2+\gamma} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right) \right) \underline{f}_{ij} \Bigg) \\
 & \quad + \alpha_{ij} \int_{\mathbb{R}^6} \left(|\text{Id} - \Pi_z^\perp| \nabla_z \log \underline{f}_{ij}|^2 \right. \\
 & \quad \times \left. c_{ij} \alpha_{ij}^2 \left(\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log \underline{f}_{ij}) + |z|^{2+\gamma} |\Pi_z^\perp \nabla_z \log \underline{f}_{ij}|^2 \right) \right) \underline{f}_{ij} \Bigg).
 \end{aligned}$$

Recalling (5.2.10) and (5.4.1), this becomes

$$\begin{aligned}
 & \left\langle \tilde{\mathcal{R}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle \\
 &= c_{ij} \alpha_{ij}^3 \int_0^\infty \int_{\mathbb{S}^2} 2 \partial_r \log \bar{g} \partial_r \left(r^\gamma \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \right) \bar{g} r^2 \, d\omega \, dr \\
 & \quad + c_{ij} \alpha_{ij}^3 \int_0^\infty \int_{\mathbb{S}^2} r^\gamma (\partial_r \log \bar{g})^2 \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} r^2 \, d\omega \, dr \\
 &=: c_{ij} \alpha_{ij}^3 \int_0^\infty (K_1 + K_2) r^2 \, dr
 \end{aligned}$$

Integrating by parts, we obtain

$$\begin{aligned}
 K_1 &= 2 \int_{\mathbb{S}^2} \partial_r \log \bar{g} \partial_r \left(r^\gamma \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \right) \bar{g} \, d\omega \\
 &= 2 \int_{\mathbb{S}^2} r^\gamma \partial_r \log \bar{g} \left(\frac{\gamma}{r} + \partial_r \right) \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} \, d\omega \\
 &= 2 \int_{\mathbb{S}^2} r^\gamma \left(\frac{1}{2} \partial_r \log \bar{g} \partial_r \left(|\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) - |\nabla_{\mathbb{S}^2} \partial_r \log \bar{g}|^2 - \frac{\gamma}{2r} \partial_r \left(|\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \right) \bar{g} \, d\omega
 \end{aligned}$$

and

$$\begin{aligned}
 K_2 &= \int_{\mathbb{S}^2} r^\gamma (\partial_r \log \bar{g})^2 \left(\Delta_{\mathbb{S}^2} \log \bar{g} + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} \, d\omega \\
 &= - \int_{\mathbb{S}^2} r^\gamma \partial_r \log \bar{g} \partial_r \left(|\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} \, d\omega.
 \end{aligned}$$

Thus, we have

$$\begin{aligned}
 & \langle \tilde{\mathcal{R}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \rangle \\
 &= -2c_{ij}\alpha_{ij}^3 \int_{\mathbb{S}^2} \int_0^\infty \left(|\nabla_{\mathbb{S}^2} \partial_r \log \bar{g}|^2 + \frac{\gamma}{2r} \partial_r \left(|\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \right) \bar{g} r^{2+\gamma} dr d\omega \\
 &= -2c_{ij}\alpha_{ij}^3 \int_{\mathbb{S}^2} \int_0^\infty \left(|\nabla_{\mathbb{S}^2} \partial_r \log \bar{g} + \frac{\gamma}{2r} \nabla_{\mathbb{S}^2} \log \bar{g}|^2 - \frac{\gamma^2}{4r^2} |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} r^{2+\gamma} dr d\omega
 \end{aligned}$$

where in the second line we completed the square,

$$|\nabla_{\mathbb{S}^2} \partial_r \log \bar{g}|^2 + \frac{\gamma}{2r} \partial_r \left(|\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) = \left| \nabla_{\mathbb{S}^2} \partial_r \log \bar{g} + \frac{\gamma}{2r} \nabla_{\mathbb{S}^2} \log \bar{g} \right|^2 - \frac{\gamma^2}{4r^2} |\nabla_{\mathbb{S}^2} \log \bar{g}|^2.$$

Changing back to variables (z, z^*) completes the proof. $\square$

Definition 5.9. We define the optimal constant Λ_3 in the log-Sobolev inequality, for any function ψ ,

$$\int_{\mathbb{S}^2} \Gamma_2(\log \psi, \log \psi) \psi d\omega \geq \Lambda_3 \int_{\mathbb{S}^2} |\nabla_{\mathbb{S}^2} \log \psi|^2 \psi d\omega, \quad (5.2.18)$$

and Λ_3^{sym} as the optimal constant under the additional symmetry constraint $\psi(\omega) = \psi(-\omega)$. Note that $\Lambda_3 = 2$ and $\Lambda_3^{\text{sym}} \geq 5.5$, see [60, Remark 1.4 and Theorem 1.1]).

Proposition 5.10 (Coercivity of the two-species weighted Fisher dissipation). *The dissipation of the two-species weighted Fisher information $\tilde{\mathcal{I}}_{ij}(f_i \otimes f_j)$ in the direction $\tilde{Q}_{ij}(f_i \otimes f_j)$ is non-negative whenever $-2 \geq \gamma \geq -\sqrt{8}$,*

$$D_{\tilde{\mathcal{I}}_{ij}, \tilde{Q}_{ij}}(f_i \otimes f_j) := -\langle \tilde{\mathcal{I}}'_{ij}(f_i \otimes f_j), \tilde{Q}_{ij}(f_i \otimes f_j) \rangle \geq 0.$$

Moreover, setting $C_{ij} := \Lambda_3 \mathbb{1}_{j \neq N+i} + \Lambda_3^{\text{sym}} \mathbb{1}_{j=N+i}$, we have the lower bound

$$\begin{aligned}
 & D_{\tilde{\mathcal{I}}_{ij}, \tilde{Q}_{ij}}(f_i \otimes f_j) \\
 & \geq 2c_{ij}\alpha_{ij}^3 \int_{\mathbb{R}^6} |z|^{2+\gamma} \left| \Pi_z^\perp \nabla_z (\hat{z} \cdot \nabla_z \log \underline{f}_{ij}) + \frac{\gamma}{2|z|} \Pi_z^\perp \nabla_z \log \underline{f}_{ij} \right|^2 \underline{f}_{ij} dz dz^* \\
 & \quad + 2c_{ij}\alpha_{ij}^3 \left(C_{ij} - \frac{\gamma^2}{4} \right) \int_{\mathbb{R}^6} |z|^\gamma \left| \Pi_z^\perp \nabla_z \log \underline{f}_{ij}(z, z^*) \right|^2 \underline{f}_{ij}(z, z^*) dz dz^* \\
 & \quad + 2c_{ij} \frac{\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_{z^*} \nabla_z \log \underline{f}_{ij}(z, z^*) : (\nabla_{z^*} \nabla_z \log \underline{f}_{ij}(z, z^*) A(z)) \underline{f}_{ij}(z, z^*) dz dz^*.
 \end{aligned}$$

Proof. Thanks to the decomposition defined earlier

$$\tilde{\mathcal{I}}_{ij}(f_i \otimes f_j) = \tilde{\mathcal{R}}_{ij;\text{rel}}(\underline{f}_{ij}) + \tilde{\mathcal{J}}_{ij;\text{rel}}(\underline{f}_{ij}) + \tilde{\mathcal{I}}_{ij;\text{cmf}}(\underline{f}_{ij})$$

, the dissipation $D_{\tilde{\mathcal{I}}_{ij}, \tilde{Q}_{ij}}(f_i \otimes f_j)$ is given by

$$\begin{aligned}
 & D_{\tilde{\mathcal{I}}_{ij}, \tilde{Q}_{ij}}(f_i \otimes f_j) \\
 &= -\left\langle \tilde{\mathcal{R}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle - \left\langle \tilde{\mathcal{J}}'_{ij;\text{rel}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle - \left\langle \tilde{\mathcal{I}}'_{ij;\text{cmf}}(\underline{f}_{ij}), \tilde{Q}_{ij}(\underline{f}_{ij}) \right\rangle \\
 &= 2c_{ij} \frac{\alpha_{ij}^2}{M_{ij}} \int_{\mathbb{R}^6} \nabla_{z^*} \nabla_z \log \underline{f}_{ij} : (\nabla_{z^*} \nabla_z \log \underline{f}_{ij} A(z)) \underline{f}_{ij} \\
 &\quad + 2c_{ij} \alpha_{ij}^3 \int_{\mathbb{R}^6} |z|^{2+\gamma} \left(\left\| (\Pi_z^\perp \nabla_z)^2 \log \underline{f}_{ij} \right\|_{\text{HS}}^2 - \frac{\gamma^2}{4|z|^2} \left| \Pi_z^\perp \nabla_z \log \underline{f}_{ij} \right|^2 \right) \underline{f}_{ij} \quad (5.2.19) \\
 &\quad + 2c_{ij} \alpha_{ij}^3 \int_{\mathbb{R}^6} |z|^{2+\gamma} \left| \Pi_z^\perp \nabla_z (\hat{z} \cdot \nabla_z \log \underline{f}_{ij}) + \frac{\gamma}{2|z|} \Pi_z^\perp \nabla_z \log \underline{f}_{ij} \right|^2 \underline{f}_{ij}.
 \end{aligned}$$

Notice that except (5.2.19), all the terms above are non-negative. To obtain the lower bound, we use the log-Sobolev inequality on $\mathbb{S}^2$, i.e., (5.2.18). Recalling $\bar{g}(r, \omega, z^*) := \underline{f}_{ij}(z, z^*)$, we have

$$\int_{\mathbb{S}^2} \left(\Gamma_2(\log \bar{g}, \log \bar{g}) - \frac{\gamma^2}{4} |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \right) \bar{g} \geq \int_{\mathbb{S}^2} \left(C_{ij} - \frac{\gamma^2}{4} \right) |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \bar{g}. \quad (5.2.20)$$

Here, we use that for $j = N + i$, $m_j = m_i$, the function

$$\bar{g}(r, \omega, z^*) = \underline{f}_{ij}(z, z^*) = f_i \left(\frac{1}{2}z + z^* \right) f_i \left(-\frac{1}{2}z + z^* \right) = f_i(v_i) f_i(v_{N+i}),$$

is symmetric in z ; thus, the constant of the log-Sobolev inequality becomes Λ_3^{sym} . Inserting back (5.2.20) into (5.2.19), and recalling (5.4.8) conclude the proof. $\square$

Remark 5.11. Without invoking the log-Sobolev inequality, the sign of the dissipation would remain unclear due to (5.2.19).

Remark 5.12. It is possible to prove Proposition 5.10 in an extrinsic fashion.

5.2.3 Local well-posedness

To derive global-in-time well-posedness of the multi-species Landau equation from the monotonicity of the Fisher information, we first establish local well-posedness. The proof is similar to the single-species case, and therefore we are brief on some of the details.

Lemma 5.13 (Local well-posedness). *Let $l \geq 5$ and $f_i^\circ \in L_l^2 \cap L \log L$ be probability densities for $1 \leq i \leq N$. Then for some $T > 0$ depending only on $\|f_i^\circ\|_{L_l^2}$, $\mathcal{H}(F_N^\circ)$, there exists a unique strong solution $f_i \in L^\infty([0, T], L_l^2) \cap C^\infty((0, T) \times \mathbb{R}^3)$ to the system*

$$\partial_t f_i = \sum_{j=1}^N Q_{ij}(f_i, f_j), \quad f_i(0, v) = f_i^\circ(v), \quad (5.2.21)$$

which satisfies

$$\sup_{t \in [0, T]} \sum_{i=1}^N \|f_i(t)\|_{L_t^2}^2 + c \sum_{i=1}^N \int_0^T \|f_i\|_{H_{t+\gamma}^1}^2 \leq C_T \sum_{i=1}^N \|f_i^\circ\|_{L_t^2}^2. \quad (5.2.22)$$

Moreover, the entropy is non-increasing,

$$\frac{d}{dt} \mathcal{H}(F_N) = -\mathcal{D}_{\mathcal{H}}(F_N) = - \sum_{i,j=1}^N D_{ij}(f_i, f_j), \quad (5.2.23)$$

where $D_{ij}(f_i, f_j)$ is defined as

$$\frac{c_{ij}}{2} \int_{\mathbb{R}^6} \left(\left(\Pi_{v_i - v_j}^\perp |v_i - v_j|^{2+\gamma} \mathbb{V}_{i,j} \log(f_i \otimes f_j) \right) \cdot \mathbb{V}_{i,j} \log(f_i \otimes f_j) \right) (f_i \otimes f_j) dv_i dv_j.$$

Proof. The existence of the solution follows by a standard regularization argument, along with the conservation of mass, total momentum and energy. For the regularity estimate, we consider the time derivative

$$\frac{d}{dt} \frac{1}{2} \int_{\mathbb{R}^3} f_i^2(v) \langle v \rangle^l dv = \sum_{j=1}^N \int_{\mathbb{R}^3} f_i(v) Q_{ij}(f_i, f_j) \langle v \rangle^l dv.$$

We then estimate each term in the sum separately. By definition of Q_{ij} , we have

$$\begin{aligned} & \int_{\mathbb{R}^3} f_i(v) Q_{ij}(f_i, f_j) \langle v \rangle^l dv \\ &= \int_{\mathbb{R}^3} f_i(v_i) \frac{\nabla v_i}{m_i} \cdot \left(\int_{\mathbb{R}^3} c_{ij} A(v_i - v_j) \left(\frac{\nabla v_i}{m_i} - \frac{\nabla v_j}{m_j} \right) (f_i \otimes f_j)(v_i, v_j) dv_j \right) \langle v_i \rangle^l dv_i \\ &= - \int_{\mathbb{R}^3} \mathcal{A}_{ij}[f_j] : \left(\frac{\nabla f_i}{m_i} \right)^{\otimes 2} \langle v \rangle^l dv - \int_{\mathbb{R}^3} \mathcal{A}_{ij}[f_j] : \left(\frac{\nabla f_i}{m_i} \otimes \frac{\nabla \langle v \rangle^l}{m_i} \right) f_i dv \\ &\quad + \int_{\mathbb{R}^3} \left(\frac{\nabla}{m_j} \cdot \mathcal{A}_{ij}[f_j] \right) \cdot \frac{\nabla}{m_i} (f_i \langle v \rangle^l) f_i dv \\ &=: -\text{I}_{ij} + \text{II}_{ij} + \text{III}_{ij}. \end{aligned}$$

where the operator $\mathcal{A}_{ij}$ is defined by

$$\mathcal{A}_{ij}[g](v) := \int_{\mathbb{R}^3} c_{ij} \Pi_{v-v^*}^\perp |v - v^*|^{2+\gamma} g(v^*) dv^*.$$

Since f_j are probability densities whose energy and entropy are bounded by the total energy and entropy of the system, we have the classical estimate, see [29],

$$\text{I}_{ij} \geq c \int_{\mathbb{R}^3} |\nabla f_i(v)|^2 \langle v \rangle^{l+\gamma} dv,$$

where $c > 0$ depends on the total energy and entropy of the initial data. Using standard interpolation estimates, making use of $l \geq 5$, we bound

$$\sum_{j=1}^N (|\mathbb{I}_{ij}| + |\mathbb{I}_{ji}|) \leq \frac{1}{4} \sum_{j=1}^N l_{ij} + C \left(1 + \sum_{j=1}^N \|f_j\|_{L_t^2} \right)^\beta \|f_i\|_{L_t^2}^2,$$

where $\beta > 0$ is a constant depending on γ . For the quantity

$$y(t) = \sum_{i=1}^N \int_{\mathbb{R}^3} f_i^2(v) \langle v \rangle^l dv,$$

this yields the differential inequality

$$\frac{d}{dt} y(t) + \frac{3}{2} \sum_{i,j=1}^N l_{ij} \leq C \langle y \rangle^{1+\beta}(t).$$

Hence, for some $T > 0$ we obtain the claimed estimate.

Finally, the entropy dissipation (5.2.23) is classical, see [51, Theorem 4]. $\square$

Corollary 5.14 (Production of Fisher information). *Let $l \geq 7$, the initial data $f_i^\circ \in L_{2l}^1 \cap L_l^2 \cap L \log L$ be probability densities for $1 \leq i \leq N$, and $f_i \in L^\infty([0, T], L_l^2) \cap C^\infty((0, T) \times \mathbb{R}^3)$ be the solution from Lemma 5.13 up to time $T > 0$. Then, after possibly decreasing $T > 0$, we have for all $1 \leq i \leq N$, $t \in (0, T)$ a.e.,*

$$\|f_i\|_{H_{l+\gamma}^2} \leq C_T < \infty, \quad \text{and} \quad \mathcal{I}(f_i) \leq C_T < \infty.$$

In particular, the Fisher information $\mathcal{I}(F)$ becomes finite.

Proof. Lemma 5.13 shows that for almost every $t_0 \in (0, T)$,

$$\|\nabla f_i(t_0)\|_{L_{l+\gamma}^2} < \infty, \quad 1 \leq i \leq N.$$

We first repeat the energy estimate below with weight $l + \gamma$. Using the uniform L_l^2 bound to control the lower-order terms, we obtain, for almost every $t_1 > t_0$,

$$\|\nabla^2 f_i(t_1)\|_{L_{l+2\gamma}^2} < \infty.$$

Since $l + 2\gamma > 0$, this implies $f_i(t_1) \in H^2(\mathbb{R}^3)$ and hence, by Sobolev embedding, $f_i(t_1) \in L^\infty(\mathbb{R}^3)$.

Moreover, by Lemma 5.15,

$$\|f_i(t_1)\|_{L_{2l}^1} < \infty.$$

Since $l - 2\gamma \leq 2l$, it follows that

$$\|f_i(t_1)\|_{L_{l-2\gamma}^2}^2 \leq \|f_i(t_1)\|_{L^\infty} \|f_i(t_1)\|_{L_{l-2\gamma}^1} < \infty.$$

An integration by parts therefore gives

$$\|\nabla f_i(t_1)\|_{L_l^2}^2 \leq C \|f_i(t_1)\|_{L_{l-2\gamma}^2} \|\nabla^2 f_i(t_1)\|_{L_{l+2\gamma}^2} + \frac{1}{2} \|\nabla f_i(t_1)\|_{L_l^2}^2 + C \|f_i(t_1)\|_{L_{l-2}^2}^2,$$

and consequently

$$\|\nabla f_i(t_1)\|_{L_l^2} < \infty.$$

We now derive energy estimates for $\partial_k f_i$ for any $k \in \{1, 2, 3\}$. The equation for the partial derivative reads

$$\partial_t \partial_k f_i = \sum_{j=1}^N Q_{ij}(\partial_k f_i, f_j) + Q_{ij}(f_i, \partial_k f_j).$$

We apply the same argument as in [Lemma 5.13](#), which yields

$$\frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^3} (\partial_k f_i)^2(v) \langle v \rangle^l dv = \sum_{j=1}^N \int_{\mathbb{R}^3} \partial_k f_i(v) (Q_{ij}(\partial_k f_i, f_j) + Q_{ij}(f_i, \partial_k f_j)) \langle v \rangle^l dv.$$

As before, we can separate the contributions from the individual species $1 \leq j \leq N$. Using the notation from above, we obtain

$$\begin{aligned} \int_{\mathbb{R}^3} \partial_k f_i(v) (Q_{ij}(\partial_k f_i, f_j) + Q_{ij}(f_i, \partial_k f_j)) \langle v \rangle^l dv &= - \int_{\mathbb{R}^3} \frac{\nabla \partial_k f_i}{m_i} \mathcal{A}[f_j] \frac{\nabla \partial_k f_i}{m_i} \langle v_i \rangle^l dv_i \\ &\quad - \int_{\mathbb{R}^3} \frac{\nabla \partial_k f_i}{m_i} \mathcal{A}[f_j] \partial_k f_i \frac{\nabla}{m_i} \langle v_i \rangle^l dv_i + \int_{\mathbb{R}^3} \partial_k f_i \frac{\nabla v_j}{m_j} \cdot \mathcal{A}[f_j] \frac{\nabla}{m_i} (\partial_k f_i \langle v_i \rangle^l) dv_i \\ &\quad - \int_{\mathbb{R}^3} \frac{\nabla f_i}{m_i} \mathcal{A}[\partial_k f_j] \frac{\nabla}{m_i} (\partial_k f_i \langle v_i \rangle^l) dv_i + \int_{\mathbb{R}^3} f_i \frac{\nabla v_j}{m_j} \cdot \mathcal{A}[\partial_k f_j] \frac{\nabla}{m_i} (\partial_k f_i \langle v_i \rangle^l) dv_i \\ &=: -\text{I}_{ij} + \text{II}_{ij} + \text{III}_{ij} + \text{IV}_{ij} + \text{V}_{ij}. \end{aligned}$$

We again use the lower bound for the matrix $\mathcal{A}$ in terms of total energy and entropy, and obtain the bound

$$\text{I}_{ij} \geq c \|\partial_k \nabla f_i\|_{L_{l+\gamma}^2}^2.$$

The remaining terms can be bounded by straightforward interpolation arguments to get the total bound

$$\partial_t \sum_{i=1}^N \|\nabla f_i\|_{L_l^2}^2 + c \sum_{i=1}^N \|\nabla^2 f_i\|_{L_{l+\gamma}^2}^2 \leq C \left(1 + \sum_{i=1}^N \|\nabla f_i\|_{L_l^2}^2\right)^{1+\beta},$$

for some constant $\beta > 0$. The upper bound for the solution of the corresponding ODE, initialized at t_1 , yields, after possibly choosing a smaller $T > t_1$, the bound

$$\int_{t_1}^T \|\nabla^2 f_i(t)\|_{L_{i+\gamma}^2}^2 dt \leq C_{t_1, T} < \infty.$$

To conclude the finiteness of the Fisher information, we recall that for any $k > 3$ we can estimate

$$I(f) = 4\|\sqrt{f}\|_{H^1}^2 \leq C_k \|f\|_{H_k^2},$$

see [98, Lemma 1] (observe the slightly different convention for $\|\cdot\|_{L_t^2}$ therein). This concludes the proof. $\square$

This local result is complemented by a control of the growth of the moments of the solution.

Lemma 5.15 (Growth bound for the moments). *Let $l, q \in \mathbb{N}$ and the initial data $f_i^\circ \in L_l^2 \cap L_q^1 \cap L \log L$ be probability densities and let $F_N = (f_i)_{i=1}^N$ be the unique local strong solution from Lemma 5.13. Then $f_i(t)$ are probability densities and the total momentum and total energy for the multi-species system are conserved,*

$$\begin{aligned} \mathcal{P}(t) &:= \sum_{i=1}^N \int_{\mathbb{R}^3} m_i v f_i(t, v) dv = \mathcal{P}(0), \\ \mathcal{E}(t) &:= \sum_{i=1}^N \int_{\mathbb{R}^3} \frac{1}{2} m_i |v|^2 f_i(t, v) dv = \mathcal{E}(0). \end{aligned}$$

The higher-order moments satisfy the growth estimate

$$\|f_i\|_{L_q^1} \leq C_q \sum_{j=1}^N \|f_j^\circ\|_{L_q^1} + C_q t^{1+\lceil \frac{(q-4)+}{2} \rceil}. \quad (5.2.24)$$

Remark 5.16. For $l \leq 4$, the estimate (5.2.24) coincides with the bound for the single species case, see [18]. On the other hand, for $l > 4$ the estimates are likely not sharp.

Proof of Lemma 5.15. Conservation of total mass and momentum follows from the fact that these quantities are preserved by each collision operator Q_{ij} . The growth bound for higher-order moments requires some care due to the asymmetry of the multi-species equation. Nevertheless, we can largely follow the strategy of [18]. Let $q \geq 0$, and consider the total q -th moment

$$M_q(t) = \sum_{i=1}^N \int_{\mathbb{R}^3} m_i \langle v \rangle^q f_i(v) dv, \quad (5.2.25)$$

and its time derivative

$$\begin{aligned}
 \frac{d}{dt} M_q(t) &= \sum_{i,j=1}^N \int_{\mathbb{R}^3} m_i \langle v \rangle^q Q_{ij}(f_i, f_j) dv \\
 &= \sum_{i=1}^N \int_{\mathbb{R}^3} m_i \langle v \rangle^q Q_{ii}(f_i, f_i) dv \\
 &\quad + \sum_{i < j} \int_{\mathbb{R}^3} m_i \langle v \rangle^q Q_{ij}(f_i, f_j) + m_j \langle v \rangle^q Q_{ji}(f_j, f_i) dv \\
 &=: \sum_{i=1}^N K_i + \sum_{i < j} K_{ij}.
 \end{aligned}$$

The terms K_i can be treated as in [18, Lemma 8], yielding for some $C > 0$

$$\sum_{i=1}^N K_i \leq \mathcal{D}_{\mathcal{H}}(F_N) M_q + C M_{q-4} + C,$$

where $\mathcal{D}_{\mathcal{H}}$ is the non-negative entropy dissipation functional introduced in (5.2.23). It remains to estimate the cross-interaction terms K_{ij} . Due to the m_i -weights in (5.2.25), the terms K_{ij} , $i \neq j$, take the form

$$-c_{ij} \int_{\mathbb{R}^6} (f_i \otimes f_j) |v_i - v_j|^{2+\gamma} \left(\Pi_{v_i - v_j}^\perp \mathbb{V}_{i,j} \log(f_i \otimes f_j) \right) \cdot (\nabla \langle v_i \rangle^q - \nabla \langle v_j \rangle^q) dv_i dv_j.$$

As in the single-species case, we introduce a smooth, radially symmetric cutoff function $\chi \in C_c^\infty$ satisfying $\mathbb{1}_{B_{\frac{1}{2}}} \leq \chi \leq \mathbb{1}_{B_1}$.

We then decompose K_{ij} into a contribution near the singularity and a far-field part:

$$\begin{aligned}
 K_{ij} &= -c_{ij} \int_{\mathbb{R}^6} (f_i \otimes f_j) \chi(v_i - v_j) |v_i - v_j|^{2+\gamma} \Pi_{v_i - v_j}^\perp \mathbb{V}_{i,j} \log(f_i \otimes f_j) \\
 &\quad \cdot (\nabla \langle v_i \rangle^q - \nabla \langle v_j \rangle^q) dv_i dv_j \\
 &\quad - c_{ij} \int_{\mathbb{R}^6} (f_i \otimes f_j) (1 - \chi)(v_i - v_j) |v_i - v_j|^{2+\gamma} \Pi_{v_i - v_j}^\perp \mathbb{V}_{i,j} \log(f_i \otimes f_j) \\
 &\quad \cdot (\nabla \langle v_i \rangle^q - \nabla \langle v_j \rangle^q) dv_i dv_j \\
 &=: K_{ij,1} + K_{ij,2}.
 \end{aligned}$$

The first term can be estimated exactly as in the single-species case, giving

$$|K_{ij,1}| \leq 2\mathcal{D}_{ij} M_q + C q^{-\gamma} M_{q-4}.$$

To treat the term $K_{ij,2}$, we integrate by parts and obtain

$$\begin{aligned} K_{ij,2} &= c_{ij} \int_{\mathbb{R}^6} (f_i \otimes f_j)(1-\chi)|v_i - v_j|^{2+\gamma} \mathbb{V}_{i,j}(\Pi_{v_i-v_j}^\perp(\nabla\langle v_i \rangle^q - \nabla\langle v_j \rangle^q)) dv_i dv_j \\ &= -2c_{ij}\alpha_{ij} \int_{\mathbb{R}^6} (f_i \otimes f_j)(1-\chi)|v_i - v_j|^\gamma (v_i - v_j) \cdot (\nabla\langle v_i \rangle^q - \nabla\langle v_j \rangle^q) dv_i dv_j \\ &\quad + c_{ij} \int_{\mathbb{R}^6} (f_i \otimes f_j)(1-\chi)|v_i - v_j|^{2+\gamma} \Pi_{v_i-v_j}^\perp : \mathbb{V}_{i,j}(\nabla\langle v_i \rangle^q - \nabla\langle v_j \rangle^q) dv_i dv_j. \end{aligned}$$

Now the first term in $K_{ij,2}$ is non-positive since

$$(v_i - v_j) \cdot (\nabla\langle v_i \rangle^q - \nabla\langle v_j \rangle^q) \geq 0,$$

while the second term can be bounded by

$$\int_{\mathbb{R}^6} (f_i \otimes f_j)(1-\chi)|v_i - v_j|^{2+\gamma} \Pi_{v_i-v_j}^\perp : \mathbb{V}_{i,j}(\nabla\langle v_i \rangle^q - \nabla\langle v_j \rangle^q) dv_i dv_j \leq CM_{q-2}.$$

Collecting terms, we obtain

$$\frac{d}{dt} M_q(t) \leq \mathcal{D}_{\mathcal{H}}(F_N) M_q + CM_{q-4} + CM_{q-2} \leq \mathcal{D}_{\mathcal{H}}(F_N) M_q + CM_{q-2}. \quad (5.2.26)$$

For $q \leq 4$, the moment M_{q-2} is uniformly bounded, and therefore

$$M_q(t) \leq CM_q(F_N^\circ) + Ct. \quad (5.2.27)$$

For the case $q > 4$, it suffices to iteratively insert the estimates obtained from (5.2.27) back into (5.2.26). $\square$

5.2.4 Global well-posedness

Proof of Theorem 5.1. The global well-posedness follows by combining the results of the previous sections. Let $\ell \geq 7$ and F_N° be the initial datum. The local well-posedness result in Lemma 5.13 yields a strong solution that can be continued as long as $\|f_i\|_{L_t^2}$ and $H(f_i)$ remain bounded. Since $\mathcal{H}(F_N)$ is monotone decreasing in time, it remains to provide a bound for the weighted L^2 norm.

To this end, we observe that Corollary 5.14 yields finiteness of the Fisher information for small $t > 0$, which is non-increasing in time due to Lemma 5.7 and Proposition 5.10. Furthermore, Lemma 5.15 gives polynomial bounds for the moments $\|f\|_{L_{2t}^1}$. By interpolation, we obtain

$$\|f\|_{L_t^2} \leq \|f\|_{L_{2t}^1}^{\frac{1}{4}} \|f\|_{L^3}^{\frac{3}{4}} \leq C \|f\|_{L_{2t}^1}^{\frac{1}{4}} I^{\frac{3}{4}}(f),$$

establishing the uniform bound and finishing the proof. $\square$

5.3 The two-species Landau system with large mass ratio

5.3.1 Non-monotonicity of the Fisher information: a counterexample

In this section, we prove [Theorem 5.2](#), which shows that, in general, the Fisher information is not a Lyapunov functional for the multi-species Landau equation, even for less singular interactions than the Coulomb interaction.

Before the proof, we derive the Gateaux derivatives of the Fisher information in the direction of the intra- and inter-species collisions.

Proposition 5.17 (Gateaux derivatives of the Fisher information). *Let f be a local-in-time strong solution to either (5.1.9) or (5.1.10) with initial datum $f^\circ \in L^1_{2l} \cap L^2_l \cap L \log L$ for some $l \geq 7$. Then, the dissipation of the Fisher information $I(f)$ in the direction $Q_\gamma(f, f)$ is given by*

$$\begin{aligned} \langle I'(f), Q_\gamma(f, f) \rangle &= -8 \int_{\mathbb{R}^6} |z|^{2+\gamma} \left| \Pi_z^\perp \nabla_z (\hat{z} \cdot \nabla_z \log \underline{f}_\otimes) + \frac{\gamma}{2|z|} \Pi_z^\perp \nabla_z \log \underline{f}_\otimes \right|^2 \underline{f}_\otimes \, dz \, dz^* \\ &\quad - 8 \int_{\mathbb{R}^6} |z|^\gamma \left(|z|^2 \left\| (\Pi_z^\perp \nabla_z)^2 \log \underline{f}_\otimes \right\|_{\text{HS}}^2 - \frac{\gamma^2}{4} \left| \Pi_z^\perp \nabla_z \log \underline{f}_\otimes \right|^2 \right) \underline{f}_\otimes \, dz \, dz^* \\ &\quad - 2 \int_{\mathbb{R}^6} \nabla_{z^*} \nabla_z \log \underline{f}_\otimes : (\nabla_{z^*} \nabla_z \log \underline{f}_\otimes A(z)) \underline{f}_\otimes \, dz \, dz^*, \end{aligned} \quad (5.3.1)$$

where $\underline{f}_\otimes(z, z^*) := (f \otimes f)(v, v^*)$ in the variables $(z, z^*) := (v - v^*, \frac{v+v^*}{2})$, is non-negative. More precisely, recalling [Definition 5.9](#) for Λ_3^{sym} , we have the lower bound,

$$\begin{aligned} D_{I, Q_\gamma}(f) &:= -\langle I'(f), Q_\gamma(f, f) \rangle \\ &\geq 8 \left(1 - \frac{\gamma^2}{4\Lambda_3^{\text{sym}}} \right) \int_{\mathbb{R}^6} |z|^{2+\gamma} \left\| (\Pi_z^\perp \nabla_z)^2 \log \underline{f}_\otimes \right\|_{\text{HS}}^2 \underline{f}_\otimes \, dz \, dz^*. \end{aligned} \quad (5.3.2)$$

Moreover, in the direction of the linear collision operator,

$$L_\beta f(v) := \nabla \cdot (|v|^{2+\beta} \Pi_v^\perp \nabla f), \quad \beta \in [-3, -2],$$

we have

$$\begin{aligned} \langle I'(f), L_\beta f \rangle &= -2 \int_{\mathbb{R}^3} |v|^{2+\beta} \left| \Pi_v^\perp \nabla (\hat{v} \cdot \nabla \log f) + \frac{\beta}{2|v|} \Pi_v^\perp \nabla \log f \right|^2 f \, dv \\ &\quad - 2 \int_{\mathbb{R}^3} |v|^{2+\beta} \left(\left\| (\Pi_v^\perp \nabla)^2 \log f \right\|_{\text{HS}}^2 - \frac{\beta^2}{4|v|^2} \left| \Pi_v^\perp \nabla \log f \right|^2 \right) f \, dv. \end{aligned} \quad (5.3.3)$$

Proof. We have

$$\langle I'(f), Q_\gamma(f, f) \rangle = \int_{\mathbb{R}^3} 2\nabla \left(\frac{Q_\gamma(f, f)}{f} \right) \cdot \nabla f + \left| \nabla \log f \right|^2 Q_\gamma(f, f) \, dv =: T_1 + T_2.$$

Then, introducing the notation $\mathbb{V} := (\nabla_v, \nabla_{v^*})$, $\mathbb{V} := \nabla_v - \nabla_{v^*}$,

$$\begin{aligned} T_1 &= 2 \int_{\mathbb{R}^6} \nabla_v \left(\frac{(\nabla_v - \nabla_{v^*}) \cdot (A(v - v^*) (\nabla_v - \nabla_{v^*}) (f \otimes f))}{f \otimes f} \right) \\ &\quad \cdot \nabla_v \log f(v) (f \otimes f) \, dv \, dv^* \\ &= 2 \int_{\mathbb{R}^6} \nabla_{v^*} \left(\frac{(\nabla_v - \nabla_{v^*}) \cdot (A(v - v^*) (\nabla_v - \nabla_{v^*}) (f \otimes f))}{f \otimes f} \right) \\ &\quad \cdot \nabla_{v^*} \log f(v^*) (f \otimes f) \, dv \, dv^* \\ &= \int_{\mathbb{R}^6} \left(\mathbb{V} \left(\frac{\mathbb{V} \cdot (A(v - v^*) \mathbb{V} (f \otimes f))}{f \otimes f} \right) \cdot \mathbb{V} \log (f \otimes f) \right) (f \otimes f) \, dv \, dv^*. \end{aligned}$$

Similarly, we can express T_2 as

$$\begin{aligned} T_2 &= \int_{\mathbb{R}^6} |\nabla_v \log f(v)|^2 (\nabla_v - \nabla_{v^*}) \cdot (A(v - v^*) (\nabla_v - \nabla_{v^*}) (f \otimes f)) \, dv \, dv^* \\ &= \int_{\mathbb{R}^6} |\nabla_{v^*} \log f(v^*)|^2 (\nabla_v - \nabla_{v^*}) \cdot (A(v - v^*) (\nabla_v - \nabla_{v^*}) (f \otimes f)) \, dv \, dv^* \\ &= \frac{1}{2} \int_{\mathbb{R}^6} |\mathbb{V} \log (f \otimes f)|^2 \mathbb{V} \cdot (A(v - v^*) \mathbb{V} (f \otimes f)) \, dv \, dv^*. \end{aligned}$$

Therefore, the Gateaux derivative can be written as

$$\begin{aligned} &\langle I'(f), Q_\gamma(f, f) \rangle \\ &= \int_{\mathbb{R}^6} \left(\mathbb{V} \left(\frac{\tilde{Q}_\gamma(f \otimes f)}{f \otimes f} \right) \cdot \mathbb{V} \log (f \otimes f) \right) (f \otimes f) + \frac{1}{2} \tilde{Q}_\gamma(f \otimes f) |\mathbb{V} \log (f \otimes f)|^2 \end{aligned}$$

with

$$\tilde{Q}_\gamma(f \otimes f) := \mathbb{V} \cdot (A(v - v^*) \mathbb{V} (f \otimes f)).$$

Notice that $\langle I'(f), Q_\gamma(f, f) \rangle = \frac{1}{2} \langle \tilde{\mathcal{I}}'_{ij}(f_i \otimes f_j), \tilde{Q}_{ij}(f_i \otimes f_j) \rangle$ for $f_i = f_j = f$, $m_i = m_j = 1$ (thus $\alpha_{ij} = 2$), and $c_{ij} = 1$, so (5.3.1) follows by repeating the argument in the proof of Proposition 5.10. The lower bound follows by the log-Sobolev inequality (5.2.18) as $\bar{g}(r, \omega, z^*) := \underline{f}_\otimes(z, z^*)$ is symmetric in $\omega \in \mathbb{S}^2$.

The linear operator L_β is a spherical diffusion in the variables $(r, \omega) := (|v|, \hat{v})$ with $\hat{v} := \frac{v}{|v|}$, i.e.,

$$L_\beta f(v) = r^\beta \Delta_{\mathbb{S}^2} \bar{f}(r, \omega), \quad \bar{f}(r, \omega) := f(v).$$

Similarly to (5.2.7), we decompose the Fisher information $I(f)$ into its radial and spherical parts,

$$I(f) = \int_{(\mathbb{R}^3)} \left(\left| (\text{Id} - \Pi_v^\perp) \nabla \log f \right|^2 + \left| \Pi_v^\perp \nabla \log f \right|^2 \right) f \, dv.$$

Then the claim (5.3.3) follows by repeating the arguments done for proving (5.2.12) and (5.2.13). $\square$

Proof of Theorem 5.2. We first prove the infinite mass case, and obtain the second part as a corollary.

(i). Infinite mass ratio: We recall the two-species system with one species with infinite mass. The equation for the finite-mass species then reads

$$\partial_t f = Q(f, f) + Lf, \quad Lf := \nabla \cdot (|v|^{2+\gamma} \Pi_v^\perp \nabla f). \quad (5.3.4)$$

The local well-posedness of this equation for regular decaying initial data f° follows as in Section 5.2.3. We refer to [55] for an extensive analysis in a more general setting.

We will construct a Schwartz function f_R° for the initial datum, such that the Fisher information increases near $t = 0$. To this end, we compute the time derivative of the Fisher information

$$\frac{d}{dt} I(f) \Big|_{t=0} = \langle I'(f^\circ), Lf^\circ \rangle + \langle I'(f^\circ), Q(f^\circ, f^\circ) \rangle. \quad (5.3.5)$$

From Proposition 5.17, we have, along the linear operator L ,

$$\begin{aligned} \langle I'(f^\circ), Lf^\circ \rangle &= -2 \int_{\mathbb{R}^3} |v|^{2+\gamma} \left| \Pi_v^\perp \nabla (\hat{v} \cdot \nabla \log f^\circ) + \frac{\gamma}{2|v|} \Pi_v^\perp \nabla \log f^\circ \right|^2 f^\circ \, dv \\ &\quad - 2 \int_{\mathbb{R}^3} |v|^{2+\gamma} \left(\left\| (\Pi_v^\perp \nabla)^2 \log f^\circ \right\|_{\text{HS}}^2 - \frac{\gamma^2}{4|v|^2} \left| \Pi_v^\perp \nabla \log f^\circ \right|^2 \right) f^\circ \, dv, \end{aligned}$$

and for the one along the nonlinear collision operator Q we have

$$\begin{aligned} &\langle I'(f^\circ), Q(f^\circ, f^\circ) \rangle \\ &= -8 \int_{\mathbb{R}^6} |z|^{2+\gamma} \left| \Pi_z^\perp \nabla_z (\hat{z} \cdot \nabla_z \log \underline{f}_\otimes^\circ) + \frac{\gamma}{2|z|} \Pi_z^\perp \nabla_z \log \underline{f}_\otimes^\circ \right|^2 \underline{f}_\otimes^\circ \, dz \, dz^* \\ &\quad - 8 \int_{\mathbb{R}^6} |z|^\gamma \left(|z|^2 \left\| (\Pi_z^\perp \nabla_z)^2 \log \underline{f}_\otimes^\circ \right\|_{\text{HS}}^2 - \frac{\gamma^2}{4} \left| \Pi_z^\perp \nabla_z \log \underline{f}_\otimes^\circ \right|^2 \right) \underline{f}_\otimes^\circ \, dz \, dz^* \\ &\quad - 2 \int_{\mathbb{R}^6} \nabla_{z^*} \nabla_z \log \underline{f}_\otimes^\circ : (\nabla_{z^*} \nabla_z \log \underline{f}_\otimes^\circ A(z)) \underline{f}_\otimes^\circ \, dz \, dz^*. \end{aligned}$$

We define initial data f_R° , $R > 0$ using the ansatz

$$f_R^\circ(v) = f_{R,1}^\circ(v) + f_{R,2}^\circ(v), \quad (5.3.6)$$

with

$$f_{R,1}^\circ(v) = \exp\left(R|v|^{\frac{|\gamma|}{2}} e_1 \cdot \hat{v}\right) \chi\left(\frac{v}{r}\right) \quad f_{R,2}^\circ(v) = \kappa \chi\left(\frac{v}{\tau}\right),$$

where χ is a radial standard cutoff function supported on the annulus defined by $\frac{1}{2} \leq |x| \leq 2$ and

$$r = \frac{R^{2/\gamma}}{|\log(2+R)|}. \quad (5.3.7)$$

For $R \geq 1$ we choose $\kappa(R), \tau(R) > 0$ such that the function f_R° has unit mass and kinetic energy.

Step 1. We first observe that $\kappa(R) \rightarrow \kappa_\infty > 0$, $\tau(R) \rightarrow \tau_\infty > 0$ converge to finite positive values as $R \rightarrow \infty$. Hence, for R sufficiently large, $f_{R,1}^\circ$ and $f_{R,2}^\circ$ have disjoint support and we can compute the dissipations along L separately;

$$\langle I'(f_R^\circ), Lf_R^\circ \rangle = \langle I'(f_{R,1}^\circ), Lf_{R,1}^\circ \rangle + \langle I'(f_{R,2}^\circ), Lf_{R,2}^\circ \rangle.$$

Using again the convergence of $\kappa(R)$ and $\tau(R)$, we conclude

$$\limsup_{R \rightarrow \infty} \left| \langle I'(f_{R,2}^\circ), Lf_{R,2}^\circ \rangle \right| < \infty.$$

Step 2. We have the following identities:

$$\begin{aligned} \Pi_v^\perp \nabla \log f_{R,1}^\circ(v) &= R|v|^{-(1+\gamma/2)} \Pi_v^\perp e_1, \\ \Pi_v^\perp \nabla (\hat{v} \cdot \nabla \log f_{R,1}^\circ(v)) &= R \frac{|\gamma|}{2} |v|^{-(2+\gamma/2)} \Pi_v^\perp e_1, \\ (\Pi_v^\perp \nabla)^2 \log f_{R,1}^\circ(v) &= -R|v|^{-(2+\gamma/2)} ((e_1 \cdot \hat{v}) \Pi_v^\perp + (\Pi_v^\perp e_1) \otimes \hat{v}). \end{aligned}$$

The functions $f_{R,1}^\circ$ are constructed such that the first term in (5.3.3) vanishes, and therefore

$$\begin{aligned} &\langle I'(f_{R,1}^\circ), Lf_{R,1}^\circ \rangle \\ &= -2 \int_{\mathbb{R}^3} |v|^{2+\gamma} \left(\left\| (\Pi_v^\perp \nabla)^2 \log f_{R,1}^\circ \right\|_{\text{HS}}^2 - \frac{\gamma^2}{4|v|^2} |\Pi_v^\perp \nabla \log f_{R,1}^\circ|^2 \right) f_{R,1}^\circ dv. \end{aligned}$$

Computing the derivatives, we find that

$$\begin{aligned}
 & \langle I'(f_{R,1}^\circ), Lf_{R,1}^\circ \rangle \\
 &= -2R^2 \int_{\mathbb{R}^3} |v|^{-2} \left(\|\Pi_v^\perp(\hat{v} \cdot e_1) + \Pi_v^\perp e_1 \otimes \hat{v}\|_{\text{HS}}^2 - \frac{\gamma^2}{4} |\Pi_v^\perp e_1|^2 \right) f_{R,1}^\circ \, dv \\
 &= -2R^2 \int_0^\infty \chi \left(\frac{l}{r} \right) \int_{\mathbb{S}^2} \left(2(\omega \cdot e_1)^2 - \left(\frac{\gamma^2 - 4}{4} \right) |\Pi_\omega^\perp e_1|^2 \right) e^{Rl^{|\gamma|/2} e_1 \cdot \omega} \, d\omega \, dl.
 \end{aligned}$$

Using

$$\int_{\mathbb{S}^2} (\omega \cdot e_1)^2 \, d\omega = \frac{4\pi}{3}, \quad \int_{\mathbb{S}^2} |\Pi_\omega^\perp e_1|^2 \, d\omega = \frac{8\pi}{3},$$

we obtain

$$\int_{\mathbb{S}^2} \left(2(\omega \cdot e_1)^2 - \left(\frac{\gamma^2 - 4}{4} \right) |\Pi_\omega^\perp e_1|^2 \right) \, d\omega = \frac{2\pi}{3} (8 - \gamma^2) < 0,$$

where we used that $\gamma^2 > 8$.

Moreover, on the support of $\chi(l/r)$ we have $l \leq 2r$, and hence

$$Rl^{|\gamma|/2} \leq 2^{|\gamma|/2} Rr^{|\gamma|/2} = 2^{|\gamma|/2} |\log(2 + R)|^{-|\gamma|/2} \longrightarrow 0.$$

Therefore,

$$\int_{\mathbb{S}^2} \left(2(\omega \cdot e_1)^2 - \left(\frac{\gamma^2 - 4}{4} \right) |\Pi_\omega^\perp e_1|^2 \right) e^{Rl^{|\gamma|/2} e_1 \cdot \omega} \, d\omega \longrightarrow \frac{2\pi}{3} (8 - \gamma^2) < 0$$

uniformly for l in the support of $\chi(l/r)$. Hence, for $R > 0$ large enough, we obtain for some $c_\gamma > 0$ that

$$\langle I'(f_{R,1}^\circ), Lf_{R,1}^\circ \rangle \geq c_\gamma R^2 r.$$

Step 3. We claim that

$$\left| \langle I'(f_R^\circ), Q(f_R^\circ, f_R^\circ) \rangle \right| \leq C \left(1 + r^{-1} + R^2 r^{|\gamma|-1} \right).$$

where r depends on R as in (5.3.7).

To this end, we choose the cutoff χ so that, in addition,

$$\chi^{\frac{1}{4}} |\nabla \log \chi| + \chi^{\frac{1}{4}} \|\nabla^2 \log \chi\|_{\text{HS}} \leq C$$

on the interior of its support. We then observe that

$$\begin{aligned} \chi^{\frac{1}{4}} \left(\frac{v}{r} \right) |\nabla_v \log f_{R,1}^\circ| &\leq \mathbb{1}_{|v| \leq 2r} C \left(R|v|^{\frac{|\gamma|}{2}-1} + r^{-1} \right), \\ \chi^{\frac{1}{4}} \left(\frac{v}{r} \right) \|\nabla_v^2 \log f_{R,1}^\circ\|_{\text{HS}} &\leq \mathbb{1}_{|v| \leq 2r} C \left(R|v|^{\frac{|\gamma|}{2}-2} + r^{-2} \right). \end{aligned}$$

Inserting these estimates into the formula for $\langle I'(f_R^\circ), Q(f_R^\circ, f_R^\circ) \rangle$, the claim follows by straightforward integration.

Step 4. We collect the estimates from Steps 1–3 and insert them into (5.3.5). Courtesy of the scaling relation (5.3.7), we obtain for $R > 1$ large enough

$$\left. \frac{d}{dt} I(f) \right|_{t=0} > c_\gamma R^2 r - C \left(1 + r^{-1} + R^2 r^{|\gamma|-1} \right) > 0,$$

establishing non-monotonicity.

The monotonicity of the Fisher information $I(f)$ for strong solutions with even initial data follows from Proposition 5.17. Since $f(t, v) = f(t, -v)$ holds for local-in-time solution, the Fisher information is dissipated by L . To see this, we recall that by (5.3.3),

$$\begin{aligned} \langle I'(f), Lf \rangle &= -2 \int_{\mathbb{R}^3} |v|^{2+\gamma} \left| \Pi_v^\perp \nabla (\hat{v} \cdot \nabla \log f) + \frac{\gamma}{2|v|} \Pi_v^\perp \nabla \log f \right|^2 f \, dv \\ &\quad - 2 \int_{\mathbb{R}^3} |v|^{2+\gamma} \left(\left\| (\Pi_v^\perp \nabla)^2 \log f \right\|_{\text{HS}}^2 - \frac{\gamma^2}{4|v|^2} |\Pi_v^\perp \nabla \log f|^2 \right) f \, dv \leq 0, \end{aligned}$$

where we use that the log-Sobolev inequality (5.2.18) holds with $\Lambda_3^{\text{sym}} \geq 5.5$ for even functions.

(ii). Large but finite mass ratio: To prove the second part of the theorem, we make use of the infinite mass case. More precisely, we set without loss of generality $m_1 = 1$ and $m_2 > 0$ to be chosen later. As initial data, we pick $f_1^\circ = f_R^\circ$ the function constructed in the previous part (5.3.6). For the second species, we take $f_2^\circ = f_{2,m_2}^\circ = \mathcal{M}_{1,m_2}$.

Due to the choice of the initial data, $Q_{22}(f_2^\circ, f_2^\circ) = 0$, so we have

$$\begin{aligned} \left. \frac{d}{dt} \mathcal{I}(F_2) \right|_{t=0} &= \frac{1}{m_1} [\langle I'(f_1^\circ), Q_{11}(f_1^\circ, f_1^\circ) \rangle + \langle I'(f_1^\circ), Q_{12}(f_1^\circ, f_2^\circ) \rangle] \\ &\quad + \frac{1}{m_2} \langle I'(f_2^\circ), Q_{21}(f_2^\circ, f_1^\circ) \rangle \\ &=: -D_{I,1} - \frac{1}{m_2} D_{I,2}. \end{aligned}$$

By explicit computation, we have the identities

$$\begin{aligned}\lim_{m_2 \rightarrow \infty} D_{I,1} &= -\langle I'(f_1^\circ), Lf_1^\circ \rangle - \langle I'(f_1^\circ), Q(f_1^\circ, f_1^\circ) \rangle, \\ \lim_{m_2 \rightarrow \infty} \frac{1}{m_2} D_{I,2} &= 0.\end{aligned}$$

The first limit above is negative due to Step 4 of the previous part. Hence, choosing $m_2 > 0$ large enough concludes the proof of the second part of the theorem. $\square$

5.3.2 Dissipation identities for the relative entropy

To go below the threshold $\gamma_* = -\sqrt{8}$, we must bound the lack of dissipation of Fisher information $I(f)$ in the direction

$$L_C f(v) := L_{-3} f(v) = \nabla \cdot \left(\frac{\Pi_v^\perp}{|v|} \nabla f(v) \right),$$

namely, the inter-species linear diffusion operator with Coulomb singularity.

To this end, we consider the weighted spherical Fisher information $J(f)$ that coincides with the dissipation of relative entropy in the direction of $L_C f$. Throughout the section, f will be a local-in-time strong solution to (5.1.10).

Lemma 5.18 (Coercivity of the Boltzmann entropy dissipation along Q_γ). *The dissipation of the Boltzmann relative entropy $H(f|\mathcal{M}_{T^\circ,1})$ in the direction $Q_\gamma(f, f)$ is non-negative and given by*

$$\begin{aligned}D_{H,Q_\gamma}(f) &:= -\langle H'(f|\mathcal{M}_{T^\circ,1}), Q_\gamma(f, f) \rangle \\ &= \frac{1}{2} \int_{\mathbb{R}^6} |v - v^*|^{2+\gamma} |\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)|^2 (f \otimes f) \, dv \, dv^*.\end{aligned}\tag{5.3.8}$$

The same holds in the direction $L_C f$:

$$D_{H,L_C}(f) := -\langle H'(f|\mathcal{M}_{T^\circ,1}), L_C f \rangle = \int_{\mathbb{R}^3} |v|^{-1} |\Pi_v^\perp \nabla \log f|^2 f \, dv.\tag{5.3.9}$$

Proof. By direct computation,

$$\begin{aligned}
 \langle H(f|\mathcal{M}_{T^0,1}), Q_\gamma(f, f) \rangle &= \int_{\mathbb{R}^3} Q_\gamma(f, f) \log f \, dv \\
 &= \int_{\mathbb{R}^6} (\nabla_v - \nabla_{v^*}) \cdot (|v - v^*|^{2+\gamma} \Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) (f \otimes f)) \log f(v) \, dv \, dv^* \\
 &= \int_{\mathbb{R}^6} (\nabla_v - \nabla_{v^*}) \cdot (|v - v^*|^{2+\gamma} \Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) (f \otimes f)) \log f(v^*) \, dv \, dv^* \\
 &= -\frac{1}{2} \int_{\mathbb{R}^6} |v - v^*|^{2+\gamma} |\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)|^2 (f \otimes f) \, dv \, dv^*.
 \end{aligned}$$

For the other dissipation (5.3.9), an integration by parts immediately concludes the proof. $\square$

Lemma 5.19 (Coercivity of the second-order Boltzmann entropy dissipation along L_C). *The dissipation of the functional $J(f) := D_{H,L_C}(f)$ in the direction $L_C f$ is non-negative and given by*

$$D_{J,L_C}(f) := -\langle J'(f), L_C f \rangle = 2 \int_{\mathbb{R}^3} |v|^{-2} \left\| (\Pi_v^\perp \nabla)^2 \log f \right\|_{\text{HS}}^2 f \, dv. \quad (5.3.10)$$

Proof. The proof is analogous to the one of (5.2.12). Notice that the prefactor $|v|^{-1}$ in the integrand of $J(f)$ commutes with spherical derivatives, and thus all the computations can be performed in the same way. $\square$

The dissipation $D_{J,L_C}(f)$ allows us to offset the lack of dissipation of $I(f)$ in the direction of $L_C f$. Thanks to (5.3.3) and the log-Sobolev inequality (5.2.18), we have

$$\begin{aligned}
 \langle I'(f), L_C f \rangle &\leq -2 \int_{\mathbb{R}^3} |v|^{-1} \left(\left\| (\Pi_v^\perp \nabla)^2 \log f \right\|_{\text{HS}}^2 - \frac{9}{4|v|^2} |\Pi_v^\perp \nabla \log f|^2 \right) f \, dv \\
 &\leq 2 \left(\frac{9}{4} - 2 \right) \int_{\mathbb{R}^3} |v|^{-3} |\Pi_v^\perp \nabla \log f|^2 f \, dv.
 \end{aligned}$$

Thus, by interpolation,

$$\begin{aligned}
 \langle I'(f), L_C f \rangle &\leq \frac{1}{2} \left(\int_{\mathbb{R}^3} |v|^{-1} |\Pi_v^\perp \nabla \log f|^2 f \, dv \right)^{\frac{1}{3}} \left(\int_{\mathbb{R}^3} |v|^{-4} |\Pi_v^\perp \nabla \log f|^2 f \, dv \right)^{\frac{2}{3}} \\
 &\leq \frac{1}{2} (D_{H,L_C}(f))^{\frac{1}{3}} \left(\frac{1}{2} D_{J,L_C}(f) \right)^{\frac{2}{3}}, \quad (5.3.11)
 \end{aligned}$$

where we recognized the first of the two integrals as $J(f) = D_{H,L_C}(f)$ given by (5.3.9), and used the fact that the second integral is controlled by (5.3.10) recalling (5.4.8) and (5.4.9);

$$|v|^{-4} |\Pi_v^\perp \nabla \log f|^2 = r^{-6} |\nabla_{\mathbb{S}^2} \log \bar{f}|^2 \leq r^{-6} \Gamma_2(\log \bar{f}, \log \bar{f}) = |v|^{-2} \|(\Pi_v^\perp \nabla)^2 \log f\|_{\text{HS}}^2.$$

5.3.3 A new Lyapunov functional

In this section, we prove [Theorem 5.3](#), which shows that Λ defined in (5.1.12) is a Lyapunov functional for the multi-species Landau equation for singular interactions up to Coulomb.

The standard Fisher information $I(f)$ and the collision operator $Q_\gamma(f, f)$ are both translation invariant, leading to the good dissipation estimate (5.3.2). On the other hand, the entropy dissipation $J(f) = D_{H, L_C}(f)$ breaks the translation invariance as L_C is singular at the origin, and hence does not interact well with $Q_\gamma(f, f)$ by itself. Therefore, we will use the dissipation of J along the spherical diffusion L_C to bound the resulting error terms. This will be achieved with the help of several functional inequalities that we gather in the following two technical lemmas.

Lemma 5.20. *Recall [Definition 5.9](#) for Λ_3^{sym} . Then the following bounds hold:*

$$\int_{\mathbb{R}^3} \left(|v|^{-2} |\Pi_v^\perp \nabla \log f|^4 + (6 + 2\sqrt{8}) |v|^{-4} |\Pi_v^\perp \nabla \log f|^2 \right) f \, dv \leq (3 + \sqrt{8}) D_{J, L_C}(f); \quad (5.3.12)$$

$$\int_{\mathbb{R}^3} \left(|\nabla \cdot (|v|^{-1} \Pi_v^\perp \nabla \log f)|^2 + 2 |v|^{-4} |\Pi_v^\perp \nabla \log f|^2 \right) f \, dv \leq D_{J, L_C}(f); \quad (5.3.13)$$

$$\begin{aligned} \int_{\mathbb{R}^6} |v - v^*|^{2+\gamma} |\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)|^4 (f \otimes f) \, dv \, dv^* \\ \leq 2(6 + 2\sqrt{8}) \bar{C}_{\gamma, \Lambda_3^{\text{sym}}} D_{I, Q_\gamma}(f); \end{aligned} \quad (5.3.14)$$

$$\begin{aligned} \int_{\mathbb{R}^6} |v - v^*|^{2+\gamma} |(\nabla_v - \nabla_{v^*}) \cdot (\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f))|^2 (f \otimes f) \, dv \, dv^* \\ \leq 4 \bar{C}_{\gamma, \Lambda_3^{\text{sym}}} D_{I, Q_\gamma}(f), \end{aligned} \quad (5.3.15)$$

where $\bar{C}_{\gamma, \Lambda_3^{\text{sym}}} := \left(1 - \frac{\gamma^2}{4\Lambda_3^{\text{sym}}}\right)^{-1}$.

Proof. For a function $\bar{g}$ on $\mathbb{S}^2$, consider the vector field $X := |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 \nabla_{\mathbb{S}^2} \bar{g}$. Then the divergence theorem yields

$$\begin{aligned} 0 &= \int_{\mathbb{S}^2} \nabla_{\mathbb{S}^2} \cdot X \\ &= \int_{\mathbb{S}^2} \left(2 \nabla_{\mathbb{S}^2}^2 \log \bar{g} : (\nabla_{\mathbb{S}^2} \log \bar{g} \otimes \nabla_{\mathbb{S}^2} \log \bar{g}) + \Delta_{\mathbb{S}^2} \log \bar{g} |\nabla_{\mathbb{S}^2} \log \bar{g}|^2 + |\nabla_{\mathbb{S}^2} \log \bar{g}|^4 \right) \bar{g}. \end{aligned}$$

Note that $|\Delta_{\mathbb{S}^2} \log \bar{g}| \leq \sqrt{2} \|\nabla_{\mathbb{S}^2}^2 \log \bar{g}\|_{\text{HS}}$. By Cauchy–Schwartz inequality, for any $0 < \delta < (1 + 1/\sqrt{2})^{-1}$,

$$\int_{\mathbb{S}^2} |\nabla_{\mathbb{S}^2} \log \bar{g}|^4 \bar{g} \leq \frac{(1 + \frac{1}{\sqrt{2}}) \delta^{-1}}{1 - (1 + \frac{1}{\sqrt{2}}) \delta} \int_{\mathbb{S}^2} \|\nabla_{\mathbb{S}^2}^2 \log \bar{g}\|_{\text{HS}}^2 \bar{g}, \quad (5.3.16)$$

and $\delta := \frac{1}{2+\sqrt{2}}$ minimizes the constant on the right-hand side of (5.3.16) to be $6 + 2\sqrt{8}$. Now (5.3.12) follows from (5.3.10) recalling (5.4.8) and (5.4.9).

The second estimate (5.3.13) follows from $|\Delta_{\mathbb{S}^2} \log \bar{g}| \leq \sqrt{2} \|\nabla_{\mathbb{S}^2}^2 \log \bar{g}\|_{\text{HS}}$.

For the third estimate (5.3.14), recall the lower bound (5.3.2) for dissipation of $I(f)$ in the direction of $Q_\gamma(f, f)$ given by Proposition 5.17,

$$D_{I, Q_\gamma}(f) \geq 8 \left(1 - \frac{\gamma^2}{4\Lambda_3^{\text{sym}}}\right) \int_{\mathbb{R}^6} |z|^{2+\gamma} \left\| (\Pi_z^\perp \nabla_z)^2 \log \underline{f}_\otimes \right\|_{\text{HS}}^2 \underline{f}_\otimes \, dz \, dz^*.$$

Note that $(\nabla_v - \nabla_{v^*}) \mapsto 2\nabla_z$, and thus,

$$|\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)|^4 = 16 |\Pi_z^\perp \nabla_z \log \underline{f}_\otimes|^4.$$

With this at hand, the inequalities (5.3.14) and (5.3.15) follow from the same argument as the first two inequalities. $\square$

Lemma 5.21. *Let $f \in L_{2l}^1 \cap L_l^2 \cap L \log L$ for some $l \geq 7$ and $\Lambda(f) < \infty$. Then*

$$\langle J'(f), Q_\gamma(f, f) \rangle \leq C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} [\Lambda(f)^{2\theta_{\gamma, \delta}} D_{J, L_C}(f) D_{I, Q_\gamma}(f)]^{\frac{1}{2}} \quad (5.3.17)$$

for some positive constants $\theta_{\gamma, \delta}$ and $C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta}$ which are computed explicitly in the proof.

Proof. We compute the Gateaux derivative,

$$\begin{aligned} & \langle J'(f), Q_\gamma(f, f) \rangle \\ &= \int_{\mathbb{R}^3} |v|^{-1} |\Pi_v^\perp \nabla_v \log f|^2 Q_\gamma(f, f) + 2|v|^{-1} \Pi_v^\perp \nabla_v \log f(v) \cdot \nabla_v (f^{-1} Q_\gamma(f, f)) f \, dv \\ &= - \int_{\mathbb{R}^3} |v|^{-1} |\Pi_v^\perp \nabla_v \log f|^2 Q_\gamma(f, f) \, dv - 2 \int_{\mathbb{R}^3} \nabla_v \cdot (|v|^{-1} \Pi_v^\perp \nabla_v \log f) Q_\gamma(f, f) \, dv \\ &= - \int_{\mathbb{R}^6} \left(|v|^{-1} |\Pi_v^\perp \nabla_v \log f|^2 (\nabla_v - \nabla_{v^*}) \right. \\ & \quad \left. \cdot (|v - v^*|^{2+\gamma} \Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) (f \otimes f)) \right) dv \, dv^* \\ & \quad - 2 \int_{\mathbb{R}^6} \nabla_v \cdot (|v|^{-1} \Pi_v^\perp \nabla_v \log f) (\nabla_v - \nabla_{v^*}) \cdot (|v - v^*|^{2+\gamma} \Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) (f \otimes f)), \end{aligned}$$

where we integrate by parts in the second equality; we omit $dv dv^*$ in the last two integrals. We recall the relation

$$\begin{aligned} & \frac{(\nabla_v - \nabla_{v^*}) \cdot (|v - v^*|^{2+\gamma} \Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) (f \otimes f))}{f \otimes f} \\ &= |v - v^*|^{2+\gamma} \left((\nabla_v - \nabla_{v^*}) \cdot (\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)) \right. \\ & \quad \left. + |\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)|^2 \right). \end{aligned}$$

Hölder and triangle inequalities then yield

$$\begin{aligned} & \langle J'(f), Q_\gamma(f, f) \rangle \\ & \leq \left[\left(\int_{\mathbb{R}^3} |v|^{-2} |\Pi_v^\perp \nabla_v \log f(v)|^4 \left(\int_{\mathbb{R}^3} |v - v^*|^{2+\gamma} f(v^*) dv^* \right) f(v) dv \right)^{\frac{1}{2}} \right. \\ & \quad \left. + 2 \left(\int_{\mathbb{R}^3} |\nabla_v \cdot (|v|^{-1} \Pi_v^\perp \nabla_v \log f(v))|^2 \left(\int_{\mathbb{R}^3} |v - v^*|^{2+\gamma} f(v^*) dv^* \right) f(v) dv \right)^{\frac{1}{2}} \right] \\ & \times \left[\left(\int_{\mathbb{R}^6} |v - v^*|^{2+\gamma} |(\nabla_v - \nabla_{v^*}) \cdot (\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f))|^2 (f \otimes f) dv dv^* \right)^{\frac{1}{2}} \right. \\ & \quad \left. + \left(\int_{\mathbb{R}^6} |v - v^*|^{2+\gamma} |\Pi_{v-v^*}^\perp (\nabla_v - \nabla_{v^*}) \log(f \otimes f)|^4 (f \otimes f) dv dv^* \right)^{\frac{1}{2}} \right] \\ & \leq C_{\gamma, \Lambda_3^{\text{sym}}} \left\| \int_{\mathbb{R}^3} |\cdot - v^*|^{2+\gamma} f(v^*) dv^* \right\|_{L^\infty}^{\frac{1}{2}} (D_{J, L_C}(f))^{\frac{1}{2}} (D_{I, Q_\gamma}(f))^{\frac{1}{2}}, \end{aligned} \quad (5.3.18)$$

where in the last inequality we used (5.3.12)–(5.3.15) from Lemma 5.20. The constant is given by

$$C_{\gamma, \Lambda_3^{\text{sym}}} := \left(2 + \sqrt{3 + \sqrt{8}} \right) \times \left(2 + \sqrt{12 + 4\sqrt{8}} \right) \times \sqrt{\bar{C}_{\gamma, \Lambda_3^{\text{sym}}}}.$$

Recall that $-3 \leq \gamma < -2$. Therefore, by Young's inequality, splitting $|\cdot - v^*|$ into two regions, one close and one away from the origin, we get

$$\begin{aligned} & \left\| \int_{\mathbb{R}^3} \frac{f(v^*)}{|\cdot - v^*|^{-(2+\gamma)}} dv^* \right\|_{L^\infty} \leq 1 + \left\| |\cdot|^{2+\gamma} \mathbf{1}_{|\cdot| < 1} \right\|_{L^{-(3-\delta)/(2+\gamma)}} \|f\|_{L^{(3-\delta)/(5+\gamma-\delta)}} \\ & = 1 + C_{\gamma, \delta} \|f\|_{L^{(3-\delta)/(5+\gamma-\delta)}} \end{aligned} \quad (5.3.19)$$

for any $\delta > 0$ small enough ($0 < \delta < 5 + \gamma$) and

$$C_{\gamma, \delta} := \left(\frac{4\pi}{\delta} \right)^{-\frac{(2+\gamma)}{(3-\delta)}},$$

where the choice of exponents stems from the criticality $|\cdot|^{-1}\mathbb{1}_{|\cdot|\leq 1} \notin L^3$. Indeed, the exponent of the Lebesgue norm of f should be as small as possible to control the norm itself by Fisher information with a small exponent, through the Sobolev inequality [22] and the $L^1 \cap L^3$ interpolation. Let C_S be the optimal constant of the Sobolev inequality, then

$$\|f\|_{L^{(3-\delta)/(5+\gamma-\delta)}} \leq \|f\|_{L^3}^{2\theta_{\gamma,\delta}} = \|\sqrt{f}\|_{L^6}^{4\theta_{\gamma,\delta}} \leq \left(C_S^2 \|\sqrt{f}\|_{\dot{H}^1}^2\right)^{2\theta_{\gamma,\delta}} = \left(\frac{C_S^2}{4} I(f)\right)^{2\theta_{\gamma,\delta}},$$

where

$$\theta_{\gamma,\delta} := -\frac{3(2+\gamma)}{4(3-\delta)}.$$

Combining with (5.3.18) and the upper bound $I(f) \leq \Lambda(f)$ gives the estimate (5.3.17) with

$$C_{S,\gamma,\Lambda_3^{\text{sym}},\delta} := C_{\gamma,\Lambda_3^{\text{sym}}} \times \left(1 + C_{\gamma,\delta} \left(\frac{C_S^2}{4}\right)^{2\theta_{\gamma,\delta}}\right)^{\frac{1}{2}}.$$

□

We are now set to prove [Theorem 5.3](#).

Proof of Theorem 5.3. We compute the time derivative of $\Lambda(f)$:

$$\begin{aligned} \frac{d}{dt}\Lambda(f) &= \frac{d}{dt}(I(f) + aJ(f) + RH(f|\mathcal{M}_{T^0,1})) \\ &= -D_{I,Q_\gamma}(f) + c_{ei}\langle I'(f), L_C(f) \rangle + a\langle J'(f), Q_\gamma(f, f) \rangle \\ &\quad - ac_{ei}D_{J,L_C}(f) - RD_{H,Q_\gamma}(f) - Rc_{ei}D_{H,L_C}(f). \end{aligned} \quad (5.3.20)$$

Courtesy of (5.3.11) and Young's inequality, we obtain for any $\eta > 0$,

$$\begin{aligned} \langle I'(f), L_C f \rangle &\leq 2^{-\frac{5}{3}} (\eta^{-2} D_{H,L_C}(f))^{\frac{1}{3}} (\eta D_{J,L_C}(f))^{\frac{2}{3}} \\ &\leq 2^{-\frac{5}{3}} \left(\frac{1}{3} \eta^{-2} D_{H,L_C}(f) + \frac{2}{3} \eta D_{J,L_C}(f) \right). \end{aligned} \quad (5.3.21)$$

Similarly for any $\kappa > 0$, using (5.3.17) we obtain,

$$\langle J'(f), Q_\gamma(f, f) \rangle \leq C_{S,\gamma,\Lambda_3^{\text{sym}},\delta} \Lambda(f)^{\theta_{\gamma,\delta}} \left(\frac{\kappa}{2} D_{J,L_C}(f) + \frac{1}{2\kappa} D_{I,Q_\gamma}(f) \right). \quad (5.3.22)$$

Now, we can control the evolution of the functional Λ . For this, we make the bootstrap assumption,

$$\Lambda(f) \leq C^\circ R,$$

for some constant $C^\circ > 0$. Note that this holds initially at time $t = 0$ with

$$C^\circ := I(f^\circ) + J(f^\circ) + H(f^\circ | \mathcal{M}_{T^\circ, 1}),$$

provided that $a \leq 1$ and $R \geq 1$. Then, inserting estimates (5.3.21) and (5.3.22) in (5.3.20) yields

$$\begin{aligned} & \frac{d}{dt} \Lambda(f) \\ & \leq -D_{I, Q_\gamma}(f) + 2^{-\frac{5}{3}} c_{ei} \left(\frac{1}{3} \eta^{-2} D_{H, L_C}(f) + \frac{2}{3} \eta D_{J, L_C}(f) \right) \\ & \quad + a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} \Lambda(f)^{\theta_{\gamma, \delta}} \frac{\kappa}{2} D_{J, L_C}(f) + a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} \Lambda(f)^{\theta_{\gamma, \delta}} \frac{1}{2\kappa} D_{I, Q_\gamma}(f) \\ & \quad - a c_{ei} D_{J, L_C}(f) - R D_{H, Q_\gamma}(f) - R c_{ei} D_{H, L_C}(f) \\ & \leq \left(\frac{a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} (C^\circ R)^{\theta_{\gamma, \delta}}}{2\kappa} - 1 \right) D_{I, Q_\gamma}(f) + \left(\frac{2^{-5/3}}{3} c_{ei} \eta^{-2} - R c_{ei} \right) D_{H, L_C}(f) \\ & \quad - R D_{H, Q_\gamma}(f) + \left(\frac{2^{-2/3}}{3} c_{ei} \eta + \frac{a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} (C^\circ R)^{\theta_{\gamma, \delta}} \kappa}{2} - a c_{ei} \right) D_{J, L_C}(f). \end{aligned}$$

Case 1: $-3 < \gamma < -2$. We choose the parameters κ and η such that all prefactors are non-positive. For this, we first set

$$\kappa := a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} (C^\circ R)^{\theta_{\gamma, \delta}},$$

so that the coefficient of $D_{I, Q_\gamma}(f, f)$, becomes non-positive, i.e.,

$$\frac{a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta} (C^\circ R)^{\theta_{\gamma, \delta}}}{2\kappa} - 1 = \frac{1}{2} - 1 = -\frac{1}{2} \leq 0.$$

Next, we set $\eta := (2^{2/3} 3R)^{-\frac{1}{2}}$, so that the coefficient of $D_{H, L_C}(f)$ becomes non-positive;

$$\frac{2^{-5/3}}{3} c_{ei} \eta^{-2} - R c_{ei} = \frac{R c_{ei}}{2} - R c_{ei} = -\frac{R c_{ei}}{2} \leq 0.$$

Lastly, we need to ensure that the coefficient of $D_{J, L_C}(f)$ is non-positive. Substituting κ and η , this amounts to

$$\frac{c_{ei}}{6\sqrt{3}} \frac{1}{\sqrt{R}} + \frac{a^2 C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta}^2 (C^\circ R)^{2\theta_{\gamma, \delta}}}{2} - a c_{ei} \leq 0.$$

We can absorb the first term setting $\frac{a}{4} := \frac{1}{6\sqrt{3}} \frac{1}{\sqrt{R}}$ so that it remains to verify

$$a C_{S, \gamma, \Lambda_3^{\text{sym}}, \delta}^2 (C^\circ R)^{2\theta_{\gamma, \delta}} \leq \frac{c_{ei}}{2}. \quad (5.3.23)$$

In the case where $-3 < \gamma < -2$, we can choose $0 < \delta < 3(3 + \gamma)$ for which $\theta_{\gamma, \delta} < 1/4$, hence (5.3.23) holds for $R > 0$ large enough. This choice satisfies in particular $\delta < 5 + \gamma$ and the L^∞ -estimate (5.3.19) is thus valid. Then

$$\frac{d}{dt} \Lambda(f) \leq -\frac{1}{2} D_{I, Q_\gamma}(f, f) - \frac{ac_{ei}}{2} D_{I, L_S}(f) - \frac{Rc_{ei}}{2} D_{H, L_C}(f) - R D_{H, L_S}(f) \leq 0.$$

Recall that we assumed a priori that $\Lambda(f) \leq C^\circ R$, which holds at time $t = 0$. This holds a posteriori for all time since Λ is non-increasing given $\Lambda(f^\circ) \leq C^\circ R$, which concludes the proof in that case.

Case 2: $\gamma = -3$. For $\lambda > 0$, consider the rescaled function f_λ given by

$$f_\lambda(v) = \lambda^{-3} f(v/\lambda).$$

Pick $\lambda > 0$ such that $I(f_\lambda^\circ) = 1$. By scale-invariance of the equation, it suffices to prove the claim for f_λ , and we consider

$$C_\lambda^\circ := I(f_\lambda^\circ) + J(f_\lambda^\circ) + H(f_\lambda^\circ | \mathcal{M}_{\lambda^2 T^\circ, 1}). \quad (5.3.24)$$

Further, we remark that for $\gamma = -3$ we have $\theta_{-3, \delta} = \frac{1}{4} \frac{3}{3-\delta}$. Now we pick $R = 1$ and fix the other parameters such that the prefactors of $D_{I, Q_\gamma}(f)$ and $D_{H, L_C}(f)$ vanish; that is,

$$\kappa := \frac{a C_{S, -3, \Lambda_3^{\text{sym}}, \delta} (C_\lambda^\circ)^{\theta_{\gamma, \delta}}}{2}, \quad \text{and} \quad \eta := \frac{1}{\sqrt{2^{5/3} 3}}.$$

It remains to determine constants such that

$$\frac{c_{ei}}{6\sqrt{6}} + \frac{a^2 C_{S, -3, \Lambda_3^{\text{sym}}, \delta}^2 (C_\lambda^\circ)^{2\theta_{\gamma, \delta}}}{4} - ac_{ei} \leq 0.$$

We set $a := \frac{1}{3\sqrt{6}}$ and the last inequality holds as soon as

$$\frac{1}{6\sqrt{6}} (C_\lambda^\circ)^{\frac{3}{6-2\delta}} C_{S, -3, \Lambda_3^{\text{sym}}, \delta}^2 \leq c_{ei}, \quad (5.3.25)$$

Hence, we obtain monotonicity of the functional $\Lambda(f)$ if the inequality above holds for some $0 < \delta < 2$. We finally obtain the well-posedness criterion (5.1.11) by picking $\delta = \frac{3}{2}$, and recasting the condition on f_λ° in terms of f° . $\square$

Remark 5.22. We can obtain an explicit value for the constant K in (5.1.11), by inserting $\delta = \frac{3}{2}$ in (5.3.25) which gives $K \approx 124$.

5.4 Appendix

5.4.1 Spherical calculus

In this appendix, we gather some spherical calculus identities that will be used throughout the chapter.

We decompose the gradient as

$$\nabla_z = (\hat{z} \otimes \hat{z}) \nabla_z + \Pi_z^\perp \nabla_z = \hat{z} \partial_{|z|} + \frac{1}{|z|} \nabla_{\hat{z}}, \quad \hat{z} := \frac{z}{|z|}, \quad \partial_{|z|} := \hat{z} \cdot \nabla_z, \quad \nabla_{\hat{z}} := |z| \Pi_z^\perp \nabla_z.$$

If we further perform the change of variables $z \mapsto r\omega$ with $(r, \omega) := (|z|, \hat{z})$, then for scalar functions we have

$$\nabla_z \mapsto \nabla_{\mathbb{S}^2}, \quad \partial_{|z|} \mapsto \partial_r, \quad \nabla_z \mapsto \omega \partial_r + \frac{1}{r} \nabla_{\mathbb{S}^2}, \quad (5.4.1)$$

We denote by $\bar{g}(r, \omega) := g(z)$ a smooth function on $\mathbb{R}^3$ in spherical coordinates. The following handy relations hold:

$$\partial_{|z|} \hat{z} = 0, \quad \nabla_{\hat{z}} |z| = 0, \quad \nabla_{\hat{z}} \hat{z} = \Pi_z^\perp, \quad \partial_{|z|} \nabla_{\hat{z}} g = \nabla_{\hat{z}} \partial_{|z|} g.$$

Most of the relations in the chapter contain $\log$; we write the sequel of the appendix that way, but these relations are of course general.

Lemma 5.23 (Spherical Hessian and Laplacian identities). *The following holds:*

$$\frac{1}{r^2} \nabla_{\mathbb{S}^2}^2 \log \bar{g} = \Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp - \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp. \quad (5.4.2)$$

In particular,

$$\frac{1}{r^2} \Delta_{\mathbb{S}^2} \log \bar{g} = \Pi_z^\perp : \left(\Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp - \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp \right). \quad (5.4.3)$$

Proof. We have (see e.g. [1]) that

$$\left(\frac{1}{r^2} \nabla_{\mathbb{S}^2}^2 \log \bar{g} \right) \sigma = (\Pi_z^\perp \nabla_z^2 \log g) \sigma + \Pi_z^\perp D_\sigma \Pi_z^\perp \nabla_z \log g, \quad \sigma \in T_{\hat{z}}(\mathbb{S}^2) \subset \mathbb{R}^3.$$

Now, since $\Pi_z^\perp \sigma = \sigma$,

$$(\Pi_z^\perp \nabla_z^2 \log g) \sigma = (\Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp) \sigma,$$

and, using $D_\sigma \hat{z} = (\nabla_z \hat{z})\sigma = \frac{1}{r}\Pi_z^\perp \sigma = \frac{1}{r}\sigma$, we obtain

$$\begin{aligned}\Pi_z^\perp D_\sigma \Pi_z^\perp \nabla_z \log g &= -\frac{1}{|z|} \Pi_z^\perp (\sigma \otimes \hat{z} + \hat{z} \otimes \sigma) \nabla_z \log g \\ &= -\frac{1}{|z|} (\sigma \otimes \hat{z}) \nabla_z \log g \\ &= -\frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp \sigma.\end{aligned}$$

(5.4.3) follows noticing that $\Pi_z^\perp = \sigma_1 \otimes \sigma_1 + \sigma_2 \otimes \sigma_2$ for any orthonormal basis $\{\sigma_1, \sigma_2\}$ of $T_{\hat{z}}\mathbb{S}^2 \subset \mathbb{R}^3$. $\square$

Lemma 5.24 (Decomposition of the Euclidean Hessian). *The following holds:*

$$\begin{aligned}\nabla_z^2 \log g &= \partial_{|z|}^2 \log g \hat{z} \otimes \hat{z} + \frac{2}{|z|} (\hat{z} \otimes_{\text{sym}} \nabla_{\hat{z}} \partial_{|z|} \log g) - \frac{2}{|z|^2} (\hat{z} \otimes_{\text{sym}} \nabla_{\hat{z}} \log g) \\ &\quad + \left(\Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp - \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp \right) + \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp.\end{aligned}$$

Proof. We decompose the Euclidean Hessian as

$$\nabla_z^2 \log g = \nabla_z^2 \log g (\hat{z} \otimes \hat{z}) + \nabla_z^2 \log g \Pi_z^\perp.$$

One checks that

$$\nabla_z^2 \log g (\hat{z} \otimes \hat{z}) = \partial_{|z|}^2 \log g \hat{z} \otimes \hat{z} + \frac{1}{|z|} \nabla_{\hat{z}} \partial_{|z|} \log g \otimes \hat{z} - \frac{1}{r^2} \nabla_{\hat{z}} \log g \otimes \hat{z}.$$

Similarly,

$$\nabla_z^2 \log g \Pi_z^\perp = (\hat{z} \otimes \hat{z}) \nabla_z^2 \log g \Pi_z^\perp + \Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp,$$

where

$$(\hat{z} \otimes \hat{z}) \nabla_z^2 \log g \Pi_z^\perp = \frac{1}{|z|} \hat{z} \otimes \nabla_{\hat{z}} \partial_{|z|} \log g - \frac{1}{r^2} \hat{z} \otimes \nabla_{\hat{z}} \log g.$$

We conclude recalling (5.4.2). $\square$

Lemma 5.25. *The following holds:*

$$\begin{aligned}\nabla_z (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log g) &= -2|z|^{2+\gamma} (\hat{z} \otimes_{\text{sym}} \nabla_{\hat{z}} \log g) + (2+\gamma) |z|^\gamma \hat{z} \otimes \nabla_{\hat{z}} \log g \\ &\quad + |z|^{1+\gamma} \hat{z} \otimes \nabla_{\hat{z}} \partial_{|z|} \log g + |z|^{2+\gamma} \left(\Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp - \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp \right).\end{aligned}\tag{5.4.4}$$

In particular,

$$\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log g) = r^\gamma \Delta_{\mathbb{S}^2} \log \bar{g}\tag{5.4.5}$$

Proof. We have

$$\begin{aligned}\nabla_z (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log g) &= \left(\hat{z} \partial_{|z|} + \frac{1}{|z|} \nabla_{\hat{z}} \right) (|z|^{1+\gamma} \nabla_{\hat{z}} \log g) \\ &= (1+\gamma) |z|^\gamma \hat{z} \otimes \nabla_{\hat{z}} \log g + |z|^{1+\gamma} \hat{z} \otimes \nabla_{\hat{z}} \partial_{|z|} \log g + |z|^\gamma \nabla_{\hat{z}}^2 \log g\end{aligned}$$

with

$$\nabla_{\hat{z}}^2 \log g = |z|^2 \left(\Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp - \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp \right) - \nabla_{\hat{z}} \log g \otimes \hat{z}, \quad (5.4.6)$$

where we used that

$$\nabla_{\hat{z}} \Pi_z^\perp = -3 (\Pi_z^\perp \otimes \hat{z})_{\text{sym}} + \hat{z} \otimes \Pi_z^\perp.$$

Hence, noting that all the tensor products in (5.4.4) are orthogonal and using $(\Pi_z^\perp)^2 = \Pi_z^\perp$,

$$\begin{aligned}\nabla_z \cdot (|z|^{2+\gamma} \Pi_z^\perp \nabla_z \log g) &= |z|^{2+\gamma} \Pi_z^\perp : \left(\Pi_z^\perp \nabla_z^2 \log g \Pi_z^\perp - \frac{1}{|z|} \partial_{|z|} \log g \Pi_z^\perp \right) \\ &= r^\gamma \Delta_{\mathbb{S}^2} \log \bar{g},\end{aligned}$$

where the last equality is (5.4.3). □

Remark 5.26. We emphasize that $\nabla_{\hat{z}}^2$, as a matrix, is not (a multiple of) the extrinsic spherical Hessian as shown by (5.4.2) and (5.4.6). They differ by the spherical gradient tensorized with the associated unit vector. Nevertheless, in view of (5.4.3),

$$\Delta_{\mathbb{S}^2} \log \bar{g} = \Pi_z^\perp : \nabla_{\hat{z}}^2 \log g. \quad (5.4.7)$$

Moreover, a direct computation yields

$$\|\nabla_{\hat{z}}^2 \log g\|_{\text{HS}}^2 = |z|^4 \left\| \left(\Pi_z^\perp \nabla_z \right)^2 \log g \right\|_{\text{HS}}^2 = \Gamma_2 (\log \bar{g}, \log \bar{g}), \quad (5.4.8)$$

where the iterated carré du champs Γ_2 of the spherical Laplacian on $\mathbb{S}^2$ is given by

$$\Gamma_2 (\log \bar{g}, \log \bar{g}) = \|\nabla_{\mathbb{S}^2}^2 \log \bar{g}\|_{\text{HS}}^2 + |\nabla_{\mathbb{S}^2} \log \bar{g}|^2. \quad (5.4.9)$$

Bibliography

- [1] P.-A. Absil, Robert Mahony, and Jochen Trumpf. “An extrinsic look at the Riemannian Hessian”. In: *Geometric science of information*. Vol. 8085. Lecture Notes in Comput. Sci. Springer, Heidelberg, 2013, pp. 361–368.
- [2] Andrei Agrachev, Ugo Boscain, Robert Neel, and Luca Rizzi. “Intrinsic random walks in Riemannian and sub-Riemannian geometry via volume sampling”. In: *ESAIM: COCV* 24.3 (2018), pp. 1075–1105.
- [3] Raphaël V. Alexandre and Cédric Villani. “On the Landau approximation in plasma physics”. In: *Ann. Inst. H. Poincaré C Anal. Non Linéaire* 21.1 (2004), pp. 61–95.
- [4] Marcos M. Alexandrino and Renato G. Bettiol. *Lie Groups and Geometric Aspects of Isometric Actions*. en. Cham: Springer International Publishing, 2015.
- [5] Marc Arnaudon, Anton Thalmaier, and Feng-Yu Wang. “Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below”. en. In: *Bulletin des Sciences Mathématiques* 130.3 (Apr. 2006), pp. 223–233.
- [6] Ata Deniz Aydın and Mikaela Iacobelli. “Asymptotic quantization of measures on Riemannian manifolds via covering growth estimates”. en. In: *Advances in Mathematics* 474 (July 2025), p. 110311.
- [7] Fabrice Baudoin, Nizar Demni, and Jing Wang. *Stochastic Areas, Horizontal Brownian Motions, and Hypocoelliptic Heat Kernels*. en. 1st ed. Vol. 37. EMS Tracts in Mathematics. EMS Press, Aug. 2024.
- [8] Mikhail Belkin, Jian Sun, and Yusu Wang. “Discrete laplace operator on meshed surfaces”. en. In: *Proceedings of the twenty-fourth annual symposium on Computational geometry*. College Park MD USA: ACM, June 2008, pp. 278–287.
- [9] Emmanuel Boissard. “Simple Bounds for the Convergence of Empirical and Occupation Measures in 1-Wasserstein Distance”. en. In: *Electron. J. Probab.* 16.none (Jan. 2011).
- [10] Emmanuel Boissard and Thibaut Le Gouic. “On the mean speed of convergence of empirical and occupation measures in Wasserstein distance”. en. In: *Ann. Inst. H. Poincaré Probab. Statist.* 50.2 (May 2014).

- [11] François Bolley, Arnaud Guillin, and Cédric Villani. “Quantitative Concentration Inequalities for Empirical Measures on Non-compact Spaces”. en. In: *Probab. Theory Relat. Fields* 137.3-4 (Mar. 2007), pp. 541–593.
- [12] François Bolley and Cédric Villani. “Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities”. en. In: *Annales de la Faculté des sciences de Toulouse : Mathématiques* 14.3 (2005), pp. 331–352.
- [13] Ugo Boscain, Robert Neel, and Luca Rizzi. “Intrinsic random walks and sub-Laplacians in sub-Riemannian geometry”. en. In: *Advances in Mathematics* 314 (July 2017), pp. 124–184.
- [14] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. en. Oxford University Press, Feb. 2013.
- [15] Marc Briant and Esther S. Daus. “The Boltzmann equation for a multi-species mixture close to global equilibrium”. In: *Arch. Ration. Mech. Anal.* 222.3 (2016), pp. 1367–1443.
- [16] Rene Cabrera, Maria Pia Gualdani, and Nestor Guillen. “Regularization estimates of the Landau-Coulomb diffusion”. In: *Nonlinear Anal.* 251 (2025), Paper No. 113695, 8.
- [17] Eric A. Carlen. “Superadditivity of Fisher’s information and logarithmic Sobolev inequalities”. In: *J. Funct. Anal.* 101.1 (1991), pp. 194–211.
- [18] Kleber Carrapatoso, Laurent Desvillettes, and Lingbing He. “Estimates for the Large Time Behavior of the Landau Equation in the Coulomb Case”. In: *Arch. Ration. Mech. Anal.* 224.2 (2017), pp. 381–420.
- [19] José A. Carrillo, Jingwei Hu, and Samuel Q. Van Fleet. “A particle method for the multispecies Landau equation”. In: *Acta Appl. Math.* 194 (2024), Paper No. 6, 28.
- [20] Joe P Chen and Patrícia Gonçalves. “Asymptotic behavior of density in the boundary-driven exclusion process on the Sierpinski gasket”. In: *Mathematical Physics, Analysis and Geometry* 24.3 (2021), p. 24.
- [21] Joe P. Chen. *Local ergodicity in the exclusion process on an infinite weighted graph*. 2017.
- [22] Andrea Cianchi, Nicola Fusco, Francesco Maggi, and Aldo Pratelli. “The sharp Sobolev inequality in quantitative form”. In: *J. Eur. Math. Soc. (JEMS)* 11.5 (2009), pp. 1105–1139.
- [23] Esther S. Daus, Ansgar Jüngel, Clément Mouhot, and Nicola Zamponi. “Hypocoercivity for a linearized multispecies Boltzmann system”. In: *SIAM J. Math. Anal.* 48.1 (2016), pp. 538–568.
- [24] Jérôme Dedecker, Aurélie Fischer, and Bertrand Michel. *Concentration of the empirical measure in Wasserstein distance: bounds involving the covering dimension*. en. arXiv:2601.11133 [math]. Jan. 2026.

- [25] Jérôme Dedecker and Florence Merlevède. "Behavior of the empirical Wasserstein distance in $\mathbb{R}^d$ under moment conditions". en. In: *Electron. J. Probab.* 24.none (Jan. 2019).
- [26] Anna DeMasi. *Mathematical Methods for Hydrodynamic Limits*. en. Lecture Notes in Mathematics Ser v.1501. Berlin, Heidelberg: Springer Berlin / Heidelberg, 1992.
- [27] Laurent Desvillettes. "On asymptotics of the Boltzmann equation when the collisions become grazing". In: *Transport Theory Statist. Phys.* 21.3 (1992), pp. 259–276.
- [28] Laurent Desvillettes, William Golding, Maria Pia Gualdani, and Amelie Loher. "Production of the Fisher information for the Landau-Coulomb equation with L1 initial data". In: *arXiv preprint arXiv:2410.10765* (2024).
- [29] Laurent Desvillettes and Cédric Villani. "On the spatially homogeneous Landau equation for hard potentials. I. Existence, uniqueness and smoothness". In: *Comm. Partial Differential Equations* 25.1-2 (2000), pp. 179–259.
- [30] Laurent Desvillettes and Cédric Villani. "On the spatially homogeneous Landau equation for hard potentials. II. H-theorem and applications". In: *Comm. Partial Differential Equations* 25.1-2 (2000), pp. 261–298.
- [31] R. M. Dudley. "The Speed of Mean Glivenko-Cantelli Convergence". en. In: *Ann. Math. Statist.* 40.1 (Feb. 1969), pp. 40–50.
- [32] Mitia Duerinckx and Corentin Le Bihan. "Lenard-Balescu thermalization: rigorous derivation from a toy model". In: *arXiv preprint arXiv:2511.10778* (2025).
- [33] Mitia Duerinckx and Raphael Winter. "Well-posedness of the Lenard-Balescu equation with smooth interactions". In: *Arch. Ration. Mech. Anal.* 247.4 (2023), Paper No. 71, 52.
- [34] Manh Hong Duong, Boris Golubkov, and Zihui He. "On a fuzzy Landau Equation: Part III. The grazing collision limit". In: *arXiv preprint arXiv:2512.04713* (2025).
- [35] Manh Hong Duong and Zihui He. "On a fuzzy Landau Equation: Part I. A variational approach". In: *arXiv preprint arXiv:2504.07666* (2025).
- [36] Manh Hong Duong and Zihui He. "On a fuzzy Landau Equation: Part II. Solvability results". In: *arXiv preprint arXiv:2507.10288* (2025).
- [37] Alessandra Faggionato. "Hydrodynamic limit of simple exclusion processes in symmetric random environments via duality and homogenization". In: *Probability Theory and Related Fields* 184.3–4 (Sept. 2022), pp. 1093–1137.
- [38] Matthew Folz. "Volume growth and stochastic completeness of graphs". en. In: *Trans. Amer. Math. Soc.* 366.4 (Sept. 2013), pp. 2089–2119.
- [39] Nicolas Fournier. "Convergence of the empirical measure in expected wasserstein distance: non-asymptotic explicit bounds in d ". en. In: *ESAIM: PS* 27 (2023), pp. 749–775.

- [40] Nicolas Fournier and Arnaud Guillin. "On the rate of convergence in Wasserstein distance of the empirical measure". en. In: *Probab. Theory Relat. Fields* 162.3-4 (Aug. 2015), pp. 707–738.
- [41] Tingran Gao. "The diffusion geometry of fibre bundles: Horizontal diffusion maps". en. In: *Applied and Computational Harmonic Analysis* 50 (Jan. 2021), pp. 147–215.
- [42] Nicolás García Trillos and Dejan Slepčev. "On the Rate of Convergence of Empirical Measures in -transportation Distance". en. In: *Can. j. math.* 67.6 (Dec. 2015), pp. 1358–1383.
- [43] Cristian Giardinà and Frank Redig. *Duality for Markov Processes: A Lie Algebraic Approach*. en. Vol. 365. Grundlehren der mathematischen Wissenschaften. Cham: Springer Nature Switzerland, 2025.
- [44] William Golding, Maria Gualdani, and Amélie Loher. "Global smooth solutions to the Landau-Coulomb equation in $L^{3/2}$ ". In: *Arch. Ration. Mech. Anal.* 249.3 (2025), Paper No. 34, 54.
- [45] Maria Gordina and Thomas Laetsch. "A convergence to Brownian motion on sub-Riemannian manifolds". en. In: *Trans. Amer. Math. Soc.* 369.9 (Mar. 2017), pp. 6263–6278.
- [46] Maria Gordina and Thomas Laetsch. "Sub-Laplacians on Sub-Riemannian Manifolds". en. In: *Potential Anal* 44.4 (May 2016), pp. 811–837.
- [47] Alexander Grigor'yan. "Heat kernels on weighted manifolds and applications". en. In: *Contemporary Mathematics*. Ed. by Jay Jorgenson and Lynne Walling. Vol. 398. Providence, Rhode Island: American Mathematical Society, 2006, pp. 93–191.
- [48] Maria Gualdani and Nestor Guillen. "On A_p weights and the Landau equation". In: *Calc. Var. Partial Differential Equations* 58.1 (2018), p. 17.
- [49] Maria Pia Gualdani, Nestor Guillen, Nataša Pavlović, Maja Tasković, and Nicola Zamponi. "The fuzzy Landau equation: global well-posedness and Fisher information". In: *arXiv preprint arXiv:2507.08689* (2025).
- [50] Maria Pia Gualdani and Raphael Winter. "A blow-down mechanism for the Landau-Coulomb equation". In: *J. Funct. Anal.* 288.7 (2025), Paper No. 110816, 36.
- [51] Maria Pia Gualdani and Nicola Zamponi. "Spectral gap and exponential convergence to equilibrium for a multi-species Landau system". In: *Bull. Sci. Math.* 141.6 (2017), pp. 509–538.
- [52] Nestor Guillen and Luis Silvestre. "The Landau equation does not blow up". In: *Acta Math.* 234.2 (2025), pp. 315–375.
- [53] Yan Guo. "Decay and continuity of the Boltzmann equation in bounded domains". In: *Arch. Ration. Mech. Anal.* 197.3 (2010), pp. 713–809.

- [54] Theodore E Harris. “Additive set-valued Markov processes and graphical methods”. In: *The Annals of Probability* (1978), pp. 355–378.
- [55] Ling-Bing He, Richard M Höfer, Jie Ji, and Raphael Winter. “Runaway avalanches in plasmas with external electric fields: spatially inhomogeneous case in a perturbation framework”. In: *arXiv preprint arXiv:2605.14520* (2026).
- [56] Matthias Hein, Jean-Yves Audibert, and Ulrique von Luxburg. “Graph Laplacians and their Convergence on Random Neighborhood Graphs”. en. In: *The Journal of Machine Learning Research* 8 (Dec. 2007), pp. 1325–1370.
- [57] Elton P. Hsu. *Stochastic analysis on manifolds*. eng. Graduate studies in mathematics volume 38. Providence, Rhode Island: American Mathematical Society, 2002.
- [58] Mikaela Iacobelli. “Asymptotic quantization for probability measures on Riemannian manifolds”. In: *ESAIM: COCV* 22.3 (July 2016), pp. 770–785.
- [59] Cyril Imbert, Luis Silvestre, and Cédric Villani. “On the monotonicity of the Fisher information for the Boltzmann equation”. In: *Acta Math.* (2025).
- [60] Sehyun Ji. “Bounds for the optimal constant of the Bakry-Émery Γ_2 criterion inequality on $\mathbb{R}P^{d-1}$ ”. In: *arXiv preprint arXiv:2408.13954* (2024).
- [61] Sehyun Ji. “Dissipation estimates of the Fisher information for the Landau equation”. In: *arXiv preprint arXiv:2410.09035* (2024).
- [62] Erik Jorgensen. “The central limit problem for geodesic random walks”. en. In: 32 (1975), pp. 1–64.
- [63] Jonathan Junné, Frank Redig, and Rik Versendaal. *Hydrodynamic limit of the symmetric exclusion process on complete Riemannian manifolds and principal bundles*. en. arXiv:2410.20167v2. 2024.
- [64] Jonathan Junné, Frank Redig, and Rik Versendaal. “Invariance Principle for Lifts of Geodesic Random Walks”. en. In: *J Theor Probab* 39.2 (June 2026), p. 21.
- [65] Jonathan Junné, Raphael Winter, and Havva Yoldaş. *On the existence of solutions to the multi-species Landau equation*. en. arXiv:2512.17082 [math.AP]. Dec. 2025.
- [66] Wilfrid S. Kendall. “The Radial Part of Brownian Motion on a Manifold: A Semimartingale Property”. In: *Ann. Probab.* 15.4 (Oct. 1987).
- [67] Claude Kipnis and Claudio Landim. *Scaling Limits of Interacting Particle Systems*. en. Ed. by S. S. Chern et al. Vol. 320. Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999.
- [68] Oldřich Kowalski and Masami Sekizawa. “On the geometry of orthonormal frame bundles”. en. In: *Mathematische Nachrichten* 281.12 (Dec. 2008), pp. 1799–1809.
- [69] Lev Davidovich Landau. *The Kinetic Equation in the Case of Coulomb Interaction*. Defense Technical Information Center, 1958.

- [70] Günter Last and Mathew Penrose. *Lectures on the Poisson Process*. en. 1st ed. Cambridge University Press, Oct. 2017.
- [71] Corentin Le Bihan and Raphael Winter. “The grazing collisions limit from the linearized Boltzmann equation to the Landau equation for short-range potentials”. In: *Kinet. Relat. Models* 16.5 (2023), pp. 654–675.
- [72] John M. Lee. *Introduction to Riemannian Manifolds*. en. Vol. 176. Graduate Texts in Mathematics. Cham: Springer International Publishing, 2018.
- [73] E.M. Lifschitz and Lev P. Pitajewski. “Physical kinetics”. In: *Textbook of theoretical physics. 10*. 1983, pp. 168–209.
- [74] Thomas M. Liggett, ed. *Interacting Particle Systems*. en. Repr. of the 1985 Ed. Classics in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.
- [75] Yu-Chu Lin, Haitao Wang, and Kung-Chien Wu. “Smoothing effects and decay estimate of the solution of the linearized two species Landau equation”. In: *Commun. Math. Sci.* 16.8 (2018), pp. 2261–2300.
- [76] Anning Liu, Jian-Guo Liu, and Yulong Lu. “On the rate of convergence of empirical measure in ∞ -Wasserstein distance for unbounded density function”. en. In: *Quart. Appl. Math.* 77.4 (May 2019), pp. 811–829.
- [77] John Lott and Cedric Villani. “Ricci curvature for metric-measure spaces via optimal transport”. en. In: *Ann. Math.* 169.3 (May 2009), pp. 903–991.
- [78] Ronald Meester and Rahul Roy. *Continuum Percolation*. 1st ed. Cambridge University Press, June 1996.
- [79] Peter Michor. *Topics in Differential Geometry*. en. Vol. 93. Graduate Studies in Mathematics. Providence, Rhode Island: American Mathematical Society, July 2008.
- [80] Kam-Ping Mok. “On the differential geometry of frame bundles of Riemannian manifolds.” In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 1978.302 (Sept. 1978), pp. 16–31.
- [81] Clément Mouhot and Lukas Neumann. “Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus”. In: *Nonlinearity* 19.4 (2006), pp. 969–998.
- [82] P.T. Nagy. “Geodesies of a principal bundle with Kaluza—Klein metric”. In: *Acta Universitatis szegediensis* 48 (1985), pp. 351–323.
- [83] Barrett O’Neill. *Semi-Riemannian geometry: with applications to relativity*. Pure and applied mathematics 103. New York: Academic Press, 1983.
- [84] Barrett O’Neill. “The fundamental equations of a submersion.” In: *Michigan Math. J.* 13.4 (Dec. 1966).
- [85] Shin-ichi Ohta. “On the measure contraction property of metric measure spaces”. In: *Comment. Math. Helv.* 82.4 (Dec. 2007), pp. 805–828.

-
- [86] Patricia Reynaud-Bouret. "Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities". In: *Probability Theory and Related Fields* 126.1 (2003), pp. 103–153.
 - [87] Shigeo Sasaki. "On the differential geometry of tangent bundles of Riemannian manifolds, II". In: *Tohoku Math. J. (2)* 14.2 (Jan. 1962).
 - [88] Pieter P. J. M. Schram. *Kinetic theory of gases and plasmas*. Vol. 46. Fundamental Theories of Physics. Kluwer Academic Publishers Group, Dordrecht, 1991.
 - [89] Mikhail Shubin. "Essential self-adjointness for semi-bounded magnetic Schrödinger operators on non-compact manifolds". In: *Journal of Functional Analysis* 186.1 (2001), pp. 92–116.
 - [90] Luis Silvestre. "Upper bounds for parabolic equations and the Landau equation". In: *J. Differential Equations* 262.3 (2017), pp. 3034–3055.
 - [91] Oleg G. Smolyanov, Heinrich V. Weizsäcker, and Olaf Wittich. "Chernoff's Theorem and Discrete Time Approximations of Brownian Motion on Manifolds". en. In: *Potential Anal* 26.1 (Jan. 2007), pp. 1–29.
 - [92] Alexander Sotirov and Shih-Hsien Yu. "On the Solution of a Boltzmann System for Gas Mixtures". In: *Arch. Ration. Mech. Anal.* 195.2 (2010), pp. 675–700.
 - [93] Frank Spitzer. "Interaction of Markov processes". en. In: *Advances in Mathematics* 5.2 (Oct. 1970), pp. 246–290.
 - [94] Herbert Spohn. *Large Scale Dynamics of Interacting Particles*. en. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991.
 - [95] Robert M. Strain. "On the linearized Balescu-Lenard equation". In: *Comm. Partial Differential Equations* 32.10-12 (2007), pp. 1551–1586.
 - [96] Karl-Theodor Sturm. "A curvature-dimension condition for metric measure spaces". en. In: *Comptes Rendus. Mathématique* 342.3 (Dec. 2005), pp. 197–200.
 - [97] Toshikazu Sunada. "Geodesic Flows and Geodesic Random Walks". en. In: University of Tokyo, 1984, pp. 47–85.
 - [98] Giuseppe Toscani and Cédric Villani. "On the trend to equilibrium for some dissipative systems with slowly increasing a priori bounds". In: *J. Statist. Phys.* 98.5-6 (2000), pp. 1279–1309.
 - [99] Bart Van Ginkel and Frank Redig. "Hydrodynamic Limit of the Symmetric Exclusion Process on a Compact Riemannian Manifold". en. In: *J Stat Phys* 178.1 (Jan. 2020), pp. 75–116.
 - [100] Veeravalli Seshadri Varadhan. "On the Convergence of Sample Probability Distributions". In: 19 (Feb. 1958), pp. 23–26.
 - [101] Rik Versendaal. "Large deviations for geodesic random walks". In: *Electron. J. Probab.* 24.none (Jan. 2019).

- [102] Cédric Villani. “A Review of Mathematical Topics in Collisional Kinetic Theory”. en. In: *Handbook of Mathematical Fluid Dynamics*. Vol. 1. Elsevier, 2002, pp. 71–74.
- [103] Cédric Villani. “Fisher information in kinetic theory”. In: *arXiv preprint arXiv:2501.00925* (2025).
- [104] Cédric Villani. “On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations”. In: *Arch. Rational Mech. Anal.* 143.3 (1998), pp. 273–307.
- [105] Cédric Villani. *Optimal Transport*. Ed. by M. Berger et al. Vol. 338. Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.
- [106] Jaak Vilms. “Totally geodesic maps”. In: *J. Differential Geom.* 4.1 (Jan. 1970).
- [107] Feng-Yu Wang. *Analysis for Diffusion Processes on Riemannian Manifolds*. en. Vol. 18. Advanced Series on Statistical Science & Applied Probability. WORLD SCIENTIFIC, Nov. 2013.
- [108] Feng-Yu Wang. “Precise limit in Wasserstein distance for conditional empirical measures of Dirichlet diffusion processes”. en. In: *Journal of Functional Analysis* 280.11 (June 2021), p. 108998.
- [109] Feng-Yu Wang. “Wasserstein convergence rate for empirical measures on noncompact manifolds”. en. In: *Stochastic Processes and their Applications* 144 (Feb. 2022), pp. 271–287.
- [110] Feng-Yu Wang and Bingyao Wu. *Wasserstein Convergence for Empirical Measures of Subordinated Diffusions on Riemannian Manifolds*. en. arXiv:2107.11568 [math]. July 2021.
- [111] Jonathan Weed and Francis Bach. “Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance”. en. In: *Bernoulli* 25.4A (Nov. 2019).
- [112] Kung-Chien Wu. “Global in time estimates for the spatially homogeneous Landau equation with soft potentials”. In: *J. Funct. Anal.* 266.5 (2014), pp. 3134–3155.

Summary

This thesis is concerned with the broad question:

How is the emergence of deterministic macroscopic laws from random microscopic dynamics influenced by the geometry of the underlying space?

The rigorous derivation of macroscopic equations from microscopic models is a central theme of mathematical physics. In the probabilistic setting, interacting particle systems provide tractable microscopic models made of many particles evolving according to stochastic rules. Under suitable scaling limits, the empirical conserved quantities of these systems are expected to converge to deterministic solutions of partial differential equations. Such results are called hydrodynamic limits.

The first model studied in this thesis is the symmetric exclusion process (SEP). In the classical Euclidean setting, the SEP is defined on a lattice and its hydrodynamic limit, under diffusive scaling, is governed by the heat equation. On a Riemannian manifold, however, there is no canonical lattice. We therefore discretize the manifold by random neighborhood graphs (RNG), obtained by sampling random vertices and connecting nearby points. This construction allows the microscopic jump rates to encode the geometry of the underlying space.

Using this approach, we prove the hydrodynamic limit for the SEP on complete Riemannian manifolds. The limiting equation is the weighted heat equation associated to a Gibbs measure.

The non-compact setting creates additional difficulties: heat may escape to infinity, and the RNG is infinite, so the existence of the SEP is not automatic. Lower Ricci curvature bounds are used to control volume growth, ensure conservativeness of the limiting diffusion, and rule out explosion of the underlying random walk (RW).

A key reason why the SEP is tractable is its self-duality with random walks. Although the particles interact through the exclusion rule, the evolution of the expected occupation variables is governed by the generator of a single RW. Thus, the hydrodynamic limit can be reduced to the convergence of graph Laplacians towards the corresponding geometric Laplacians.

The geometric setting also makes it possible to obtain anisotropic limiting equations.

When the space has the structure of a principal bundle, horizontal and vertical directions can be distinguished. By choosing jump rates at two different scales, we derive sub-elliptic diffusion equations with drift as hydrodynamic limits of the SEP. The limiting operator is then a horizontal Laplacian rather than the full Laplace–Beltrami operator.

This horizontal Laplacian is naturally related to the Laplace–Beltrami operator on the base manifold. We give a probabilistic interpretation of this relation by studying geodesic random walks and their horizontal lifts. We prove that suitably lifted geodesic random walks converge to horizontal Brownian motion, while their projections converge to Brownian motion on the base manifold. In the limit, this recovers the classical identity relating the horizontal Laplacian to the Laplace–Beltrami operator.

A second class of models studied in this thesis consists of McKean–Vlasov particle systems on manifolds. These are continuum particle systems describing aggregation–diffusion dynamics under mean-field interactions. In the many-particle limit, one expects propagation of chaos; the particles become asymptotically independent and their common law solves a nonlinear Fokker–Planck equation. For applications and simulations, it is important to quantify how close the empirical measure of the particle system is to this limiting law.

This leads to quantitative estimates for empirical measures on manifolds. We prove bounds in expectation and concentration inequalities in Wasserstein distance. The methods are metric in nature and depend on covering estimates for balls. Geometry therefore enters through volume growth, and in particular through curvature bounds. The same curvature assumptions also enter the stability estimates for the McKean–Vlasov dynamics via comparison theory and couplings of diffusions on manifolds. Combining both ingredients yields quantitative propagation-of-chaos estimates in curved spaces.

The final part of the thesis concerns the spatially homogeneous multi-species Landau equation, a kinetic equation from plasma physics describing a mixture of several species undergoing grazing collisions. The unknown is a vector of velocity distributions where the different species may have different masses. Although this part is not about hydrodynamic limits, it is still connected to the geometric theme of the thesis. In the Coulomb case, the collision operator contains an angular diffusion acting on the sphere of relative velocities. This spherical structure is closely related to Fisher information methods.

For the single-species Coulomb Landau equation, it was recently proved that the Fisher information is a Lyapunov functional. In the multi-species setting, the different masses break part of the symmetry used in the single-species argument. We show that, below a certain threshold of singularity in the interactions, the Fisher information ceases to be a Lyapunov functional. Nevertheless, in the limiting ion–electron regime for two species, we recover well-posedness for Coulomb interactions by introducing a new Lyapunov functional built from Fisher information, relative entropy, and entropy dissipation along the inter-species collision operator.

Acknowledgments

Frank, you not only gave me the opportunity to conduct a thesis in this fascinating field at the intersection of probability, geometry, and analysis, but also shaped my journey throughout the PhD. I am deeply grateful for your teaching, your guidance, and the kindness with which you answered my many questions. You always supported my projects and gave me the freedom to develop my own ideas. I am especially grateful for the trust you placed in my ability to work independently and obtain results, while still being present to both assert ideas and guide me back whenever I risked losing focus or spreading myself too thin. You also gave me several opportunities to present my work, and you created an active seminar in geometry and probability at Delft, which fostered connections between researchers that might otherwise not have developed. You were inspiring as group leader and contributed greatly to the warm and stimulating atmosphere within the probability group; I thank you for this as well. In music, you also surprised me with your vast knowledge and abilities which were inspiring.

Rik, you were always available whenever I wanted to discuss mathematics or needed your opinion on my work. I very much appreciated the spontaneous nature of these discussions, which often turned an informal question into a helpful mathematical exchange.

Fabrice, Mark, Nathael, Patrícia, and Richard, I would like to thank you all to partake as external jury members for my dissertation defense. Fabrice, I would like to thank you for the hospitality in Aarhus and the common mathematical interest which now leads to an active collaboration. Richard, I would like to thank you for the interesting mathematical and political discussions.

Berend, Jasper, and Victor, thank you for sharing the office with me, for the not-so-serious atmosphere that came with it, and for the countless coffee breaks. Ardjen, Hidde, Serena, and Vicente, thank you for the many fun evenings, the conversations, and all the moments we shared throughout the PhD. You made these years much richer and more enjoyable. Yago, thank you for keeping French alive a few more years within the PhDs.

Brieuc, Colm, Guilhem, Maxime, you already know everything... I am sincerely grateful to have you in my life. Laura, your support is beyond what anyone can expect.

Eliott, Guillaume, Louan, Lucas, Maxime, Pacôme, and Pierre, thank you for being there ever since our bachelor's years. I always look forward to our monthly get-togethers even if I

cannot always be there; the variety of our conversations and the kind of laughter that is unique to this group are refreshing.

Benoît, Edouard, and Pierre-François, I am happy that, after more than ten years, we still take the time to go on holiday together and enjoy these moments of friendship.

Finally, I would like to thank my family for the support throughout the journey. Maman, you have always been present for me, and I know you always will be. Having your presence by my side is invaluable. Gauthier, your love, kindness, and curiosity are an infinite source of joy and inspiration.

Curriculum Vitae

Jonathan Junné was born in Woluwe-Saint-Lambert, Belgium. He completed his secondary education at Centre Scolaire Saint-Michel in Etterbeek.

He studied physics and mathematics at Université Libre de Bruxelles, where he obtained his Bachelor degrees in 2019 and 2020 respectively, both with Magna Cum Laude. He then continued with a master's degree in pure mathematics at ETH Zürich. He graduated in 2022 with a thesis entitled "On the Uniqueness of the Vlasov-Poisson system", written under the supervision of Prof.dr. Mikaela Iacobelli.

In July 2022, he started his PhD in the Delft Institute of Applied Mathematics at Delft University of Technology, under the supervision of Prof.dr. Frank Redig and Dr. Rik Versendaal. This PhD position was funded by the Dutch Research Council (NWO) through the project *Interacting particle systems and Riemannian geometry* with project number OCENW.M20.251.

Publications

Published

- [1] M. Iacobelli and J. Junné. Stability estimates for the Vlasov–Poisson system in p -kinetic Wasserstein distances. *Bulletin of the London Mathematical Society*, 56:2250–2267, 2024.
- [2] J. Junné, F. Redig, and R. Versendaal. Invariance Principle for Lifts of Geodesic Random Walks. *Journal of Theoretical Probability*, 39:21, 2026.

Preprints

- [3] J. Junné, F. Redig, and R. Versendaal. Hydrodynamic limit of the symmetric exclusion process on complete Riemannian manifolds and principal bundles. arXiv preprint arXiv:2410.20167, 2024.
- [4] J. Junné and A. Rege. Stability estimates in kinetic Wasserstein distances for the Vlasov–Poisson system with Yudovich density. arXiv preprint arXiv:2503.07122, 2025. *To appear in Kinetic and Related Models*.
- [5] I. Ben Porat, A. Gagnebin, M. Iacobelli, and J. Junné. Propagation of velocity moments for the magnetized Vlasov–Poisson system with space-time dependent magnetic fields. arXiv preprint arXiv:2510.22753, 2025.
- [6] J. Junné, R. Winter, and H. Yoldaş. On the existence of solutions to the multi-species Landau equation. arXiv preprint arXiv:2512.17082, 2025.
- [7] F. Baudoin, J. Junné, and D. Tewodrose. Asymptotic mean value Laplacian on equiregular sub-Riemannian manifolds. arXiv preprint arXiv:2609.02710, 2026.
- [8] J. Junné, R. Winter, and H. Yoldaş. A counterexample to McKean’s conjecture for the Landau–Coulomb equation. arXiv preprint arXiv:2609.03847, 2026.

