Check for
updates

Annealed quantitative estimates for the quadratic 2D-discrete random matching problem

Nicolas Clozeau[1] · Francesco Mattesini[2,3]

© The Author(s) 2024

Abstract

We study a random matching problem on closed compact 2-dimensional Riemannian manifolds (with respect to the squared Riemannian distance), with samples of random points whose common law is absolutely continuous with respect to the volume measure with strictly positive and bounded density. We show that given two sequences of numbers n and $m = m(n)$ of points, asymptotically equivalent as n goes to infinity, the optimal transport plan between the two empirical measures μ^n and ν^m is quantitatively well-approximated by $\big(\mathrm{Id}, \exp(\nabla h^n)\big)_{\#}\mu^n$ where h^n solves a linear elliptic PDE obtained by a regularized first-order linearization of the Monge–Ampère equation. This is obtained in the case of samples of correlated random points for which a stretched exponential decay of the α-mixing coefficient holds and for a class of discrete-time sub-geometrically ergodic Markov chains having a unique absolutely continuous invariant measure with respect to the volume measure.

✉ Francesco Mattesini
  francesco.mattesini@uni-muenster.de

  Nicolas Clozeau
  nicolas.clozeau@ist.ac.at

[1] Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria

[2] Institut für Mathematische Stochastik, Universität Münster, Orleáns Ring 10, 48149 Münster, Germany

[3] Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstrasse 22, 04103 Leipzig, Germany

Keywords Optimal transport · Matching problem · Quantitative estimates

Mathematics Subject Classification 9Q22 · 35R60 · 58J35 · 35J15

Contents

1 Introduction and statement of the main results

1.1 The random matching problem and its asymptotic

The random matching problem is a popular optimization problem at the interface between analysis and probability with applications in many different fields such as statistical physics [16, 47], computer science [43] and economics [21, 25]. Within the mathematical literature, it has been subject of intense studies due to its interactions with many areas, including for instance graph theory [41] and geometric probability [55]. In this paper we focus on one of its simple versions. Let $\{X_k\}_{1 \leq k \leq n}$ and $\{Y_k\}_{1 \leq k \leq m}$ (with possibly $m > n$) be two families of random points on a compact Riemannian manifold $\mathcal{M}$ (endowed with the Riemannian distance d). We are interested in the quadratic matching problem

$$\min_{\pi \in \Pi_{nm}} \sum_{i=1}^{n} \sum_{j=1}^{m} \pi_{ij} \, \mathrm{d}^2(X_i, Y_j) \tag{1.1}$$

where

$$\Pi_{nm} := \left\{ \pi \in [0,1]^{n \times m} \,\middle|\, \sum_{i=1}^{n} \pi_{ij} = \tfrac{1}{m} \quad \text{and} \quad \sum_{j=1}^{m} \pi_{ij} = \tfrac{1}{n} \right\}.$$

Classically, (1.1) can be phrased in terms of a transport problem. Indeed, letting

$$\mu^n := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i} \quad \text{and} \quad \nu^m := \frac{1}{m} \sum_{j=1}^{m} \delta_{Y_j}, \tag{1.2}$$

be the empirical measures associated with the two point clouds, the linear programming problem (1.1) amounts to determine the quadratic Wasserstein distance $W_2^2(\mu^n, \nu^m)$.

In the special case $n = m$ the Birkhoff–von Neumann Theorem provides a correspondence between (1.1) and the usual bipartite matching

$$\min_{\sigma \in \mathcal{S}_n} \sum_{i=1}^{n} d^2(X_i, Y_{\sigma(i)}), \tag{1.3}$$

where $\mathcal{S}_n$ denotes the set of injective maps $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Indeed, since Π_{nn} is a convex polytope, minimizers in (1.1) have to be searched among extremal points. By the Birkhoff–von Neumann Theorem [8, Lemma 2.1.3], the latter are nothing but permutation matrices (up to a factor $\frac{1}{n}$).

A first natural question is to understand the asymptotics of (1.1) as $n, m \uparrow \infty$. For the same number of samples $n = m$ and independently and identically distributed (i.i.d.) on the unit square $[0, 1]^d$, the scaling of the cost (1.1) has been well understood in the mathematical and statistical physics literature. A simple heuristic argument, see for instance [47], suggests that given a point X_i, we can find a point Y_j within a volume of order $O(n^{-1})$ with high probability. For this reason, the typical inter-point distance is of order $O(n^{-\frac{1}{d}})$ suggesting that the scaling of (1.1) is of order $O(n^{-\frac{2}{d}})$. Although attractive, this heuristic turns out to be unfortunately false in low dimension showing a critical behavior when $d = 2$. This critical case is the one on which we focus on in this paper. Ajtai et al. [1] were the first to show that, for i.i.d. uniform samples, a logarithmic correction is needed, deriving[1]

$$\mathbb{E}\big[W_2^2(\mu^n, \nu^n)\big] \sim \frac{\log(n)}{n}, \tag{1.4}$$

extended later by Talagrand [58] for clouds of i.i.d. points which are distributed accordingly to more general common law. A recent breakthrough was obtained within the physics community by Caracciolo et al. [16], and further developed by Caracciolo and Sicuro [17] and by Sicuro in [54], where the asymptotics of the cost are formally derived thanks to a novel PDE approach and optimal transport theory rather than combinatorics. A couple of years later, in general 2-dimensional compact Riemannian manifolds without boundary, the first-order asymptotic has been rigorously justified by Ambrosio et al. [6] for i.i.d. uniform samples and recently extended by Ambrosio et al. [5] for samples distributed accordingly to more general laws which are absolutely continuous (with Hölder continuous density) w.r.t. the volume measure dm, leading

---

[1] We use the notation $A \lesssim B$ if there exists a global constant $C > 0$, which may only depend on d, such that $A \leq CB$. We write $A \sim B$ if both $A \lesssim B$ and $B \lesssim A$ hold.

to

$$\lim_{n \to \infty} \frac{n}{\log(n)} \mathbb{E}\big[W_2^2(\mu^n, \nu^n)\big] = \frac{|\mathcal{M}|}{2\pi}, \tag{1.5}$$

where $|\mathcal{M}|$ denotes the Lebesgue measure of $\mathcal{M}$. The case $n \neq m$ with $n, m \uparrow \infty$ with similar rates is also covered, see [5, Theorem 1.2].

The novel approach introduced in [16], later revised in [11], consists in a linearization of the Monge–Ampère equation that allows for an explicit description of the cost thanks to the linearized proxies (see Sect. 1.2 for more details). The aim of this work is to quantitatively justify the linearization ansatz in terms of convergence of the approximating minimizers of (1.1) towards the optimal ones. In particular, we are interested in the case where the points are identically distributed with a common law $\rho \, dm$ (where we recall that dm denotes the volume measure) and ρ satisfies for some $\lambda, \Lambda > 0$

$$\lambda \leq \rho \leq \Lambda. \tag{1.6}$$

To the best of our knowledge, there are only few results on the asymptotic behavior of the transport map and they are so far limited to the case of i.i.d. uniform samples in the study of the semi-discrete matching problem (that is couplings between μ^n and dm), see the work of Ambrosio et al. [4]. In connection with this work, quantitative estimates on the optimal map for the matching between the Lebesgue measure and Poisson clouds have been obtained by Goldman et al. [31] and Goldman and Huesmann [30].

Our extension in this paper is fourfold: First, we look at more general distribution of points and we consider the case of general densities ρ satisfying (1.6). Second, we do not assume independence and we consider samples which may possess correlations. Third, we do not restrict the analysis to the semi-discrete matching problem and we also investigate the ansatz for the full matching problem (1.1). Finally, we investigate the case where the points are not identically distributed and we extend our result to a class of sub-geometrically ergodic Markov chains.

We finally mention that the effectiveness of the linearization ansatz introduced in [16] is not only limited to the case of i.i.d. distributed points on bounded domains, but it can be employed in many different settings. See for instance [15] for an interesting application to the matching on unbounded domains, [37–39] for an application to Gaussian matching, [35] for an application in random matrix theory, [34, 64–66] for an application to a continuous instance of the matching problem, i.e. when the empirical measure is replaced by the occupation measure of a stochastic process. It is worth to further mention that these techniques can also be employed when considering the matching problem with p-costs in higher dimension, see [32] and when considering a larger class of optimization problems, see [33].

1.2 Linearization ansatz

We now briefly reproduce the linearization ansatz introduced in [16]. For simplicity, we consider the case $\mathcal{M} = \mathbb{T}^2$, $n = m$ and i.i.d. samples with common distribution $\rho\,dm$. Let T^n be an optimal transport map (whose existence is ensured by Brenier's Theorem [14]) between μ^n and ν^n. Based on the transport relation $T^n_\# \mu^n = \nu^n$ and a change of variables, T^n solves (formally) the Monge–Ampère equation

$$(\mu^n \circ T^n)\det(\nabla T^n) = \nu^n. \tag{1.7}$$

Since the cost is quadratic, by [52, Theorem 1.25], there exists a function h^n such that $T^n = \mathrm{Id} + \nabla h^n$. Applying the law of large numbers, we further have the weak convergences $\mu^n, \nu^n \rightharpoonup \rho\,dm$ as $n \uparrow \infty$ so that we expect $T^n \approx \mathrm{Id}$ as $n \uparrow \infty$. Thus, this suggests that the correction ∇h^n is small as $n \uparrow \infty$, allowing to perform (formally) the Taylor expansions

$$\mu^n \circ T^n \approx \mu^n + \nabla\mu^n \cdot \nabla h^n \quad\text{and}\quad \det(\nabla T^n) \approx 1 + \Delta h^n. \tag{1.8}$$

Plugging (1.8) into (1.7), neglecting the higher order terms and replacing μ^n by ρ yields

$$-\nabla \cdot \rho \nabla h^n = \mu^n - \nu^n. \tag{1.9}$$

This formal linearization suggests the following two conjectures

$$\lim_{n\uparrow\infty} \left| \mathbb{E}\big[W_2^2(\mu^n, \nu^n) \big] - \mathbb{E}\left[\int_{\mathbb{T}^2} |\nabla h^n|^2 \rho\,dm \right] \right| = 0 \quad\text{and}$$

$$\lim_{n\uparrow\infty} \int_{\mathbb{T}^2} |T^n - (\mathrm{Id} + \nabla h^n)|^2 = 0 \text{ a.s.} \tag{1.10}$$

Unfortunately, (1.10) cannot hold as it is, since the solution of (1.9) does not belong to H^1 due to the roughness of the source term. To overcome this, following the strategy in [6], a regularization using the heat-semigroup at time $\sim \frac{1}{n}$ (up to logarithmic corrections) is made. Doing so, the first item of (1.10) turns out to be true, leading to the result (1.5) (see for instance [3] for a convergence rate).

1.3 Formulation of the main results

For the remainder of the paper $\mathcal{M}$ denotes a 2-dimensional connected and compact Riemannian manifold without boundary (or the square $[0, 1]^2$) endowed with the Riemannian distance d. For $t > 0$ we denote by p_t the fundamental solution of the heat operator $\partial_t - \Delta$ on $\mathcal{M}$, where Δ denotes the Beltrami-Laplace operator. We define the heat semigroup $(\mathrm{P}_t)_{t>0}$ via its action on probability measures $\mu \in \mathcal{P}(\mathcal{M})$

and square integrable functions $f \in L^2(\mathcal{M})$

$$P_t \mu := \int_{\mathcal{M}} p_t(\cdot, y) d\mu(y) \quad \text{and} \quad P_t f := \int_{\mathcal{M}} p_t(\cdot, y) f(y) dm(y).$$

We first introduce the class of correlated point clouds that we consider for studying the matching problem (1.1). This class concerns point clouds $\{X_i\}_i$ for which the correlations between points decay at an exponential rate, where the correlations are measured in terms of the α-mixing coefficient given by, for any $\ell \geq 1$

$$\alpha_\ell := \sup_{k \geq 1} \sup \left\{ \frac{\text{cov}(f, g)}{\|f\|_{L^\infty} \|g\|_{L^\infty}}, \quad f \in L^\infty(\sigma\{X_j, j \leq k\}) \quad \text{and} \right.$$

$$\left. g \in L^\infty(\sigma\{X_j, j \geq \ell + k\}) \right\}, \tag{1.11}$$

and the β-mixing coefficient given by[2]

$$\beta_{ij} := \sup_{|F| \leq 1} \left| \int_{\mathcal{M} \times \mathcal{M}} F \, d(\mathbb{P}_{(X_i, X_j)} - \mathbb{P}_{X_i} \otimes \mathbb{P}_{X_j}) \right| \quad \text{for any } i, j \geq 1. \tag{1.12}$$

Assumption 1.1 (*Correlated point clouds*) We consider point clouds $\{X_i\}_i \subset \mathcal{M}$ which are identically distributed according to $\rho \, dm$ where ρ satisfies (1.6). We further assume decay of the correlations in the form of

$$\sup_{n \geq 1} \frac{1}{n} \sum_{1 \leq i < j \leq n} \beta_{ij} < \infty, \tag{1.13}$$

and there exist $a, b > 0$ and $\eta \in (0, \infty]$ such that

$$\alpha_\ell \leq a \exp\left(-b\ell^\eta\right) \quad \text{for any } \ell \geq 1. \tag{1.14}$$

Assumption 1.1 is made to ensure good concentration properties of the point clouds. On the one hand, under (1.13), the cost $W_2^2(\mu^n, \nu^n)$ behaves as in the i.i.d. case (1.4) (cf. [12, Theorem 2] and Appendix B). On the other hand, the sub-exponential decay (1.14) of the α-mixing coefficient ensures sub-exponential concentration properties (cf. [45, Theorem 1] and Proposition A.1), which is necessary to run our argument. We refer the reader to Sect. 2 for further technical details.

Our first main result concerns the approximation of transport plans coupling $\{\mu^n\}_n$ and $\{\nu^m\}_m$ defined in (1.2). We justify the formal linearization of the Monge–Ampère equation achieved in Sect. 1.2 in an annealed quantitative way (i.e. in expectation): We show that, a suitable regularization of, the plan $(\text{Id}, \exp(\nabla h^n))_\# \mu^n$, with h^n defined in (1.9), provides a good approximation when measuring the error with respect to the

---

[2] We denote by $\mathbb{P}_X$ the law of a random variable X.

W_2-Wasserstein distance in the product space $\mathcal{M} \times \mathcal{M}$ endowed with the metric

$$\delta^2\big((x, y), (z, w)\big) := d^2(x, z) + d^2(y, w). \tag{1.15}$$

The density ρ will need further regularity in form of fractional Sobolev spaces defined as, for some $\varepsilon > 0$

$$H^\varepsilon := \left\{ f : \mathcal{M} \to \mathbb{R} \,\Big|\, \|f\|_{H^\varepsilon}^2 := \sum_{k \geq 1} \lambda_k^{2\varepsilon} |\hat{f}(k)|^2 < \infty \right\}, \tag{1.16}$$

where $\{\lambda_k, \phi_k\}_k$ denote the eigenvalues and eigenvectors of $-\Delta$ on $\mathcal{M}$ and $\hat{f}(k) = \int_{\mathcal{M}} f \, \phi_k$ denotes the Fourier modes of f. Finally, we denote by $\dot{H}^1$ the L^2-based Sobolev space

$$\dot{H}^1 := \left\{ f : \mathcal{M} \to \mathbb{R} \,\Big|\, \int_{\mathcal{M}} |\nabla f|^2 < \infty \ \text{ with } \int_{\mathcal{M}} f \, dm = 0 \right\}. \tag{1.17}$$

Theorem 1.2 (Approximation of the transport plan) *Let $\rho \in H^\varepsilon$ for some $\varepsilon > 0$ satisfying (1.6) and $\{\mu^n\}_n$ and $\{\nu^m\}_m$ be defined in (1.2) (for $m = m(n)$ with some given increasing map $m : \mathbb{N} \to \mathbb{N}$) with point clouds satisfying Assumption 1.1 and such that there exists $q \in [1, \infty)$ for which $\frac{m(n)}{n} \xrightarrow[n \uparrow \infty]{} q$. We consider[3] $h^{n,t} \in \dot{H}^1$ the weak solution of*

$$-\nabla \cdot \rho \nabla h^{n,t} = \mu^{n,t} - \nu^{m,t}, \tag{1.18}$$

for any $t \in (0, 1)$ with $\mu^{n,t} := P_t \mu^n$ and $\nu^{m,t} := P_t \nu^m$.

There exist an exponent $\kappa > 0$, a deterministic constant C and a random variable $\mathcal{C}_n$ both depending on λ, Λ and $\mathcal{M}$ for which given $t = \frac{\log^\kappa(n)}{n}$ and

$$\gamma^{n,t} := \big(\mathrm{Id}, \exp(\nabla h^{n,t})\big)_{\#} \mu^{n,t}, \tag{1.19}$$

it holds

$$\inf_\pi W_2^2(\pi, \gamma^{n,t}) \leq \mathcal{C}_n \frac{\log(n)}{n} \sqrt{\frac{\log\log(n)}{\log(n)}} \ \text{ with } \ \sup_{n \geq 1} \mathbb{E}[\tfrac{1}{C} \mathcal{C}_n] \leq 1, \tag{1.20}$$

where the inf *runs over all optimal transport plans π between μ^n and ν^m.*

Furthermore, if (1.14) holds with $\eta > 2$, the assumption (1.13) can be dropped and it holds

$$\inf_\pi W_2^2(\pi, \gamma^{n,t}) \leq \mathcal{C}_n \frac{\log(n)}{n} \sqrt{\frac{\log\log(n)}{\log^{1 - \frac{2}{\eta}}(n)}} \ \text{ with } \ \sup_{n \geq 1} \mathbb{E}[\exp(\tfrac{1}{C} \mathcal{C}_n^{\frac{1}{2}})] \leq 2,$$
$$\tag{1.21}$$

---

[3] Where we impose additional Neumann boundary conditions in the case $\mathcal{M} = [0, 1]^2$.

where the inf *runs over all optimal transport plans π between μ^n and ν^m.*

Our second main result concerns the particular case of the semi-discrete matching problem, i.e. optimal coupling between the common law $\rho\,dm$ and $\{\mu^n\}_n$. We know from McCann's theorem [42] that there exists a unique optimal transport map T^n, that is the optimal transport plan π^n can be written as

$$\pi^n = \big(\mathrm{Id}, T^n\big)_{\#}\mu^n.$$

We show that T^n can be approximated in an annealed quantitative way in L^2 by (a suitable regularized version of) the solution of (1.9) with ν^m replaced by ρ.

Theorem 1.3 *Let $\rho \in H^\varepsilon$ for some $\varepsilon > 0$ satisfying (1.6) and $\{\mu^n\}_n$ be defined in (1.2) with a point cloud satisfying Assumption 1.1. We consider $f^{n,t} \in \dot{\mathrm{H}}^1$ the weak solution of*

$$- \nabla \cdot \rho \nabla f^{n,t} = \mu^{n,t} - \rho_t, \tag{1.22}$$

where, for all $t \in (0, 1)$, we recall that $\mu^{n,t} = P_t\mu^n$ and $\rho_t = P_t\rho$. Finally, we denote by T^n the optimal transport map from $\rho\,dm$ to μ^n.

There exist an exponent $\kappa > 0$, a deterministic constant C and a random variable C_n both depending on λ, Λ and $\mathcal{M}$ for which given $t = \frac{\log^\kappa(n)}{n}$, it holds

$$\int_{\mathcal{M}} d^2\big(T^n, exp(\nabla f^{n,t})\big)dm \le C_n \frac{\log(n)}{n}\sqrt{\frac{\log\log(n)}{\log(n)}} \quad with \ \sup_{n\ge 1}\mathbb{E}[\tfrac{1}{C}C_n] \le 1. \tag{1.23}$$

Furthermore, if (1.14) holds with $\eta > 2$, the assumption (1.13) can be dropped and it holds

$$\int_{\mathcal{M}} d^2\big(T^n, exp(\nabla f^{n,t})\big)dm \le C_n \frac{\log(n)}{n}\sqrt{\frac{\log\log(n)}{\log^{1-\frac{2}{\eta}}(n)}} \quad with \ \sup_{n\ge 1}\mathbb{E}[\exp(\tfrac{1}{C}C_n^{\frac{1}{2}})] \le 2. \tag{1.24}$$

We finally mention that in the case where the eigenfunctions $\{\phi_k\}_k$ admit a uniform bound, the conclusions (1.21) and (1.24) can be improved. We comment on the proof at the end of Sect. 3.5.

Remark 1.4 Let $\{\mu^n\}_n$ and $\{\nu^m\}_m$ be as in Theorem 1.2. We assume that the family of eigenfunctions $\{\phi_k\}_k$ satisfies the uniform bound

$$\sup_{k\ge 1} \|\phi_k\|_{\mathrm{L}^\infty} < \infty. \tag{1.25}$$

Then, (1.21) and (1.24) hold true for $\eta > 1$ with a convergence rate $\frac{\log(n)}{n}\sqrt{\frac{\log\log(n)}{\log^{1-\frac{1}{\eta}}(n)}}$ and the same stochastic integrability. Note that (1.25) typically holds when the geometry of $\mathcal{M}$ is flat, see [60].

Theorems 1.2 and 1.3 are not restricted to the case of identically distributed point clouds and we present in the next section and in Appendix C a possible extension, using the same techniques, to a class of sub-geometrically ergodic discrete-time Markov chains.

1.4 Extension to a class of sub-geometrically ergodic Markov chains

We first recall some basic facts on discrete-time Markov chains on $\mathcal{M}$. Such a Markov process is described by its initial distribution $\mu_0 \in \mathcal{P}(\mathcal{M})$ and its transition kernel K, that is a measurable map from $\mathcal{M}$ to the space of probability measures $\mathcal{P}(\mathcal{M})$. We recall that K acts on $\mathcal{P}(\mathcal{M})$ in form of

$$(K\mu)(A) = \int_{\mathcal{M}} K(\cdot, A) \, d\mu \quad \text{for every Borel set } A \subset \mathcal{M} \text{ and } \mu \in \mathcal{P}(\mathcal{M}), \quad (1.26)$$

and likewise on bounded measurable function ψ in form of

$$(K\psi)(x) = \int_{\mathcal{M}} \psi(y) \, K(x, dy) \quad \text{for any } x \in \mathcal{M}. \quad (1.27)$$

Given an initial distribution $\mu_0 \in \mathcal{P}(\mathcal{M})$, we recall that the law of a Markov chain $\{X_n\}_{n \geq 0}$ can be expressed with the help of the transition kernel, namely

$$X_n \sim K^n \mu_0 \quad \text{for any } n \geq 0, \quad (1.28)$$

where K^n stands for the nth-iteration of the kernel K.

We now introduce the class of discrete-time Markov chains we consider.

Assumption 1.5 Let $\mathcal{M}$ be a 2-dimensional compact Riemannian manifold. Let $\mu_0 \in \mathcal{P}(\mathcal{M})$ and $\{X_n\}_{n \geq 1} \subset \mathcal{M}$ be a Markov chain with initial law μ_0. We assume that the chain satisfies the two following conditions:

(i) We assume that there exists a measurable function $k : \mathcal{M} \times \mathcal{M} \to [0, \infty)$ and $\lambda, \Lambda > 0$ such that for any Borel set $A \subset \mathcal{M}$

$$K(x, A) = \int_A k(x, \cdot) dm \quad \text{with } \lambda \leq k \leq \Lambda. \quad (1.29)$$

(ii) We assume that there exist an unique invariant measure μ_∞, i.e. $K\mu_\infty = \mu_\infty$, constants $a, b > 0$ and $\eta \in (0, 1]$ as well as a map $V : \mathcal{M} \to [1, \infty)$ such that $\int_{\mathcal{M}} V d\mu_\infty, \int_{\mathcal{M}} V d\mu_0 < \infty$, for which for any $\ell \geq 1$ and any $\phi \in L^\infty(\mathcal{M})$

$$\|K^\ell \phi - \mu_\infty(\phi)\|_V \leq a \exp(-b\ell^\eta) \|\phi\|_\infty \quad \text{for any } x \in \mathcal{M}, \quad (1.30)$$

where

$$\|\phi\|_V := \sup \frac{|\phi|}{1 + V} \quad \text{and} \quad \mu_\infty(\phi) = \int_{\mathcal{M}} \phi \, d\mu_\infty.$$

We now comment on the consequences of the above assumptions. First, the condition (1.29) ensures that the invariant measure μ_∞ is absolutely continuous with respect to the volume measure with bounded density, namely

$$\mu_\infty = \rho \, \mathrm{dm} \quad \text{with } \lambda \leq \rho \leq \Lambda. \tag{1.31}$$

We briefly give the argument for (1.31). Using the second item of (1.29), we have that the operator

$$\mathcal{I} : f \in \mathrm{L}^1 \mapsto \int_{\mathcal{M}} k(\cdot, y) f(y) \mathrm{dm}(y) \quad \text{is compact.}$$

Furthermore, we note that the closed convex set $\mathcal{C} := \{f \in \mathrm{L}^1 \mid \lambda \leq f \leq \Lambda \text{ and } \int_{\mathcal{M}} f = 1\}$ is invariant under the action of $\mathcal{I}$. Therefore Schauder's fixed point theorem implies that $\mathcal{I}$ admits a fixed point in $\mathcal{C}$. Given such a fixed point ρ, it is clear that μ_∞ defined in (1.31) is an invariant measure according to (1.29).

Second, for irreducible and aperiodic Markov chains, the condition (1.30) is satisfied when the kernel satisfies the following geometric drift condition: There exist a function $V : \mathcal{M} \to [1, \infty)$, a petite set $\mathcal{P}$ and a constant $C > 0$ such that $\sup_{\mathcal{P}} V < \infty$ and

$$KV + \phi \circ V \leq V + C\mathbb{1}_{\mathcal{P}},$$

with for large $x \geq 1$, $\phi(x) = c\frac{x}{\log^\alpha(x)}$ for some $\alpha \geq 0$ and $c > 0$ ((1.30) is then satisfied with $\eta = \frac{1}{1+\alpha}$), see [22, Theorem 2.8 & Section 2.3] and the references therein. Moreover, the assumption (1.30) implies the sub-exponential decay of the β-mixing coefficient (1.12), namely there exists a constant C depending on λ, Λ and $\mathrm{m}(\mathcal{M})$ such that

$$\beta_{ij} \leq C \exp(-b|i - j|^\eta) \quad \text{for any } i, j, \tag{1.32}$$

which ensures that (1.13) and (1.14) hold and ensures good concentration property of the Markov chain, see Proposition A.1. The estimate (1.32) can be seen as a direct consequence of the combination of the estimate on the β-mixing coefficient in [40, Proposition 3] and the assumptions (1.29) and (1.30).

Finally, the condition (1.30) quantifies the weak convergence of the law of the Markov chain to its stationary distribution, namely there exists a constant $C > 0$ such that for any $f \in \mathrm{L}^\infty(\mathcal{M})$

$$|\mathbb{E}[f(X_n)] - \mu_\infty(f)| \leq C \exp(-bn^\eta)\|f\|_{\mathrm{L}^\infty}. \tag{1.33}$$

We shortly give the argument. We first notice that a direct inductive argument together with the semigroup property $K^{n_1+n_2} = K^{n_1} K^{n_2}$ for every $n_1, n_2 > 0$ and Fubini's theorem gives

$$\int_{\mathcal{M}} f \, \mathrm{d}K^n \mu_0 = \int_{\mathcal{M}} K^n f \, \mathrm{d}\mu_0 \quad \text{for any } n \geq 1. \tag{1.34}$$

The combination of (1.28), (1.34) and (1.30) gives

$$|\mathbb{E}[f(X_n)] - \mu_\infty(f)| \overset{(1.28),(1.34)}{=} \left| \int_{\mathcal{M}} (K^n f - \mu_\infty(f)) \, d\mu_0 \right|$$

$$\overset{(1.30)}{\leq} \left(\int_{\mathcal{M}} 1 + V d\mu_0 \right) a \exp(-bn^\eta) \|f\|_\infty.$$

A classical example of a Markov chain satisfying Assumption 1.5 is given by iterated function systems with additive noise. For simplicity, let $\mathcal{M} = \mathbb{T}^2$. Let $\{\theta_n\}_{n\geq 1}$ be i.i.d. random variables with common law $h \, dm$ for some h satisfying $\lambda \leq h \leq \Lambda$. Let $F : \mathbb{T}^2 \to \mathbb{T}^2$ be a contraction, i.e. there exists a constant $L < 1$ such that

$$|F(x) - F(y)| \leq L|x - y| \quad \text{for any } x, y \in \mathbb{T}^d. \tag{1.35}$$

We define the iterated function system $\{X_n\}_{n\geq 1}$ according to the induction

$$X_{n+1} = F(X_n) + \theta_n \quad \text{for any } n \geq 1.$$

The kernel is given by

$$K(x, A) = \int_{\mathbb{T}^2} \mathbb{1}_A(F(x) + \theta) h(\theta) dm(\theta),$$

so that K satisfies (1.29) with

$$k(x, \cdot) = h(\cdot - F(x)).$$

Moreover, the condition (1.35) ensures the validity of (1.30), see for instance [2, Theorem 3.2]. Thus the Markov process $\{X_n\}_{n\geq 1}$ satisfies Assumption 1.5.

We show in Appendix C that the conclusions of Theorem 1.3 and Theorem 1.2 hold true for Markov chains satisfying Assumption 1.5.

1.5 Open problems

We conclude this section with open questions that arise in view of our results. Those concern optimality of our convergence rates, extensions to more general costs and different type of correlated point clouds. For the latter, we mention two possible directions concerning the Ginibre ensemble and Coulomb gases that we think are worth investigating.

(1) *Sharpness of the rates.* The convergence rate in (1.20) and (1.23) match with the one obtained for the case of uniformly distributed samples in [4]. However, even in the latter case, it has not been shown whether the rate is optimal and we suspect the opposite. A possible way to track the optimal rate could be to perform a second-order linearization of the Monge–Ampère equation (1.7). Following the

same type of computations leading to (1.9) in the case $\rho = 1$, a second-order linearization q^n should solve

$$-\Delta q^n = \det(\nabla^2 h^n),$$

where we recall that h^n solves (1.9), providing the conjecture

$$\lim_{n\uparrow\infty} \left| \int_{\mathbb{T}^2} |T^n - (\mathrm{Id} + \nabla h^n)|^2 - \int_{\mathbb{T}^2} |\nabla q^n|^2 \right| = 0.$$

(2) *Extension to p-costs.* A natural question is to investigate if our results hold for different cost functions as p-cost functions for $p > 1$. The behavior of the cost has been optimally quantified in [1, 9, 13, 24]. However, to the best of our knowledge, quantitative estimates on the transport plan in the setting of general p-costs are not known, even for uniformly distributed samples. A possible approach would be to revise the linearization ansatz of [16] for general p-costs. Indeed, if the transport cost between two points x, y is given by $\frac{1}{p}|x-y|^p$ on the torus, Gangbo-McCann's theorem [26, Theorem 1.2] ensures that there exists a map h^n such that the optimal transport map T^n takes the form $T^n = \mathrm{Id} + |\nabla h^n|^{p'-2}\nabla h^n$, where p' denotes the conjugate exponent. Therefore, following the same type of computations leading to (1.9), a first-order linearization should solve the following degenerate p'-Laplace equation

$$-\nabla \cdot \rho |\nabla h^n|^{p'-2}\nabla h^n = \mu^n - \nu^m,$$

and we may expect

$$\lim_{n\uparrow\infty} \int_{\mathbb{T}^2} |T^n - (\mathrm{Id} + |\nabla h^n|^{p'-2}\nabla h^n)|^p = 0.$$

See also [36] for a justification of this linearisation ansatz down to mesoscopic scales based on a large-scale regularity theory for the Monge–Ampère equation.

(3) *Ginibre ensemble.* A (complex) Ginibre ensemble is a non-Hermitian random matrix with independent complex Gaussian entries. Given a $n \times n$ Ginibre ensemble X, we define its empirical spectral distribution as

$$\mu^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i},$$

where $\{\lambda_i\}_{i=1}^n$ are the eigenvalues of the matrix $\frac{X}{\sqrt{n}}$. The so called Circular Law states that, almost-surely, μ^n weakly converges to the uniform distribution on the complex unit disk μ^∞ having Lebesgue density $\frac{1}{\pi}\mathbb{1}_{B_1}$, see for instance [59, Theorem 1.10]. An interesting question is to quantitatively understand the weak convergence of μ^n towards μ^∞ measured in Wasserstein distance. A possible approach to this problem would be to employ the linearization argument discussed

in Sect. 1.2. Note that the Ginibre ensemble is not covered by our setting as it is posed on the whole space $\mathbb{R}^2$ and, in general, the eigenvalues $\{\lambda_i\}_{i=1}^n$ possess long-range correlations and therefore do not satisfy our Assumption 1.1. However, in [50, Théorème 3.1.1], the linearization argument has been shown to be robust enough to derive an upper bound on $\mathbb{E}[W_2(\mu^n, \mu^\infty)]$ (see also [35, Theorem 1.3] for similar results for $\mathbb{E}[W_1(\mu^n, \mu^\infty)]$). The techniques used in [35, 50] avoid a quantification of the correlations between the eigenvalues: This is done in [50] by a decomposition argument together with concentration estimates of the Wasserstein distance around the Moser's coupling and in [35] by using classical tools from non-Hermitian random matrix theory. A challenging question would be to investigate the exact asymptotics of the transport cost $\mathbb{E}[W_2(\mu^n, \mu^\infty)]$ in the case of the Ginibre ensemble complementing, [50, Théorème 3.1.1] and [35, Theorem 1.3] with a lower bound and consequently quantify the convergence of the linearized proxies to the optimal transport map.

(4) *Planar Coulomb gases.* Planar Coulomb gases are many particles systems, in which the particles $\{X_i\}_{i=1}^n$ have repulsively Coulomb interactions and are confined by a potential $V : \mathbb{T}^2 \to \mathbb{R}$. These are modelled by the Hamiltonian

$$\mathcal{H}_n(\{X_i\}_{i=1}^n) = -\sum_{i \neq j} \log |X_i - X_j| + n \sum_{i=1}^n V(X_i) \qquad (1.36)$$

and the Gibbs measure

$$\mathrm{d}\mathbb{P}_{n,\beta} = \frac{1}{Z_{n,\beta}} \exp\left(-\beta \mathcal{H}_n(\{X_i\}_{i=1}^n)\right) \mathrm{d}X_1 \ldots \mathrm{d}X_n, \qquad (1.37)$$

where $Z_{n,\beta}$ is the normalizing constant and β denotes the inverse temperature. In analogy with (1.2), we can define the empirical measure of a Coulomb gas by $\mu^n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$. In this setting the convergence of the empirical measure exhibits a twofold behavior. On the one hand, for small temperature $\beta \gg \frac{1}{n}$, the empirical measure μ^n weakly converges to the *equilibrium measure* μ_{low}^∞ defined as the minimizer

$$\mu_{\text{low}}^\infty := \operatorname*{argmin}_{\mu \in \mathcal{P}(\mathbb{T}^2)} \left\{ -\int_{\mathbb{T}^2 \times \mathbb{T}^2} \log(x - y) \, \mathrm{d}\mu \otimes \mu(x, y) + \int_{\mathbb{T}^2} V \, \mathrm{d}\mu \right\}.$$

On the other hand, for large temperature $\beta \ll \frac{1}{n}$, a correction term is required and the empirical measure μ^n weakly converges to the *thermal equilibrium measure* μ_{high}^∞, that is the minimizer

$$\mu_{\text{high}}^\infty := \operatorname*{argmin}_{\mu \in \mathcal{P}(\mathbb{T}^2)} \left\{ -\int_{\mathbb{T}^2 \times \mathbb{T}^2} \log(x - y) \, \mathrm{d}\mu \otimes \mu(x, y) + \int_{\mathbb{T}^2} V \, \mathrm{d}\mu + \frac{1}{n\beta} \int_{\mathbb{T}^2} \mu \log \mu \right\}.$$

We refer the reader to [53] for a complete exposition on 2D-Coulomb gases. This setting can be seen as an extension of the previous Ginibre ensemble on the torus.

Indeed the law of the spectrum of the Ginibre ensemble is given by the Gibbs measure (1.37) choosing $\beta = 2$ and $V(x) = |x|^2$ in (1.36). Motivated by this observation and the works [35, 50] on the Ginibre ensemble, we expect that the linearization approach could be employed also in this setting. First, we could justify the linearization Ansatz in the spirit of Theorem 1.2 in both temperature regimes, using available results in the literature. Indeed, concentration inequalities around the equilibrium measure have been derived in [27] (see also [18]), that would replace the Bernstein's type inequality of Proposition A.1 and the matching cost estimates in Proposition B.1. Second, a more ambitious question would be to use the linearization argument to derive optimal rates of the convergence to the equilibrium measure in both temperature regimes.

2 Structure of the proof

This section is devoted to describe the main ideas and how are organized the proofs of Theorems 1.2 and 1.3. We mainly focus on the proof of Theorem 1.2 since the proof of Theorem 1.3 follows by the same strategy.

General strategy The proof of Theorem 1.2 follows the strategy employed in [4] to deal with independent and uniformly distributed random points. The main idea is to use the quantitative stability result for transport maps in [4, Theorem 3.2], stating that two transport maps are close in the L^2-topology if the target measures are close in the W_2-topology. We restate below the result for reader's convenience.

Theorem 2.1 (Stability of transport maps) *Let $\nu, \mu_1, \mu_2 \in \mathcal{P}(\mathcal{M})$ such that $\nu \ll m$ and let $T, S : \mathcal{M} \to \mathcal{M}$ be the optimal transport maps respectively for the pairs of measures (ν, μ_1) and (ν, μ_2). We assume that $S = \exp(\nabla f)$ for some $f : \mathcal{M} \to \mathbb{R}$ with $C^{1,1}$-regularity.*
There exists a constant $c > 0$ depending on $\mathcal{M}$ such that, provided

$$\|\nabla f\|_{L^\infty} + \|\nabla^2 f\|_{L^\infty} \le c, \tag{2.1}$$

we have

$$\int_{\mathcal{M}} d^2(S, T)\, d\nu \lesssim W_2^2(\mu_1, \mu_2) + W_2(\mu_1, \mu_2) W_2(\nu, \mu_1).$$

The first step consists of using Theorem 2.1 to deduce a stability estimate (in terms of the quadratic Wasserstein distance) of transport plans in the special case where $\mu_1 = \nu^m$, $\mu_2 = \exp(\nabla h^{n,t})_\# \mu^n$ and $\nu = \mu^n$. In this step, we immediately face the issue of the lack of regularity of $h^{n,t}$ necessary to ensure the condition (2.1): Indeed, recalling that $h^{n,t}$ solves (1.18), it does not have the $C^{1,1}$-regularity condition for non-smooth densities ρ that we consider here. We overcome this issue introducing an additional regularization step: We smooth the operator $-\nabla \cdot \rho\nabla$ and implicitly $\gamma^{n,t}$ in form of

$$\gamma_\delta^{n,t} := \big(\mathrm{Id}, \exp(\nabla h_\delta^{n,t})\big)_\# \mu^{n,t} \quad \text{with} \ -\nabla \cdot \rho_\delta \nabla h_\delta^{n,t} = \mu^{n,t} - \nu^{m,t}, \tag{2.2}$$

for a regularization parameter δ to be optimized and $\rho_\delta := P_\delta \rho$. Classical Schauder's theory ensures that $h_\delta^{n,t}$ owns C^∞-regularity. Doing so, we can use Theorem 2.1 to deduce, provided that

$$\|\nabla h_\delta^{n,t}\|_{L^\infty} + \|\nabla^2 h_\delta^{n,t}\|_{L^\infty} \ll 1, \tag{2.3}$$

a stability error estimate which reads

$$\begin{aligned}
\inf_\pi W_2^2(\pi, \gamma_\delta^{n,t}) &\lesssim W_2^2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\#\mu^{n,t}\big) + W_2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\#\mu^{n,t}\big) W_2(\mu^n, \nu^m) \\
&\quad + W_2^2(\nu^{m,t}, \nu^m) + W_2^2(\mu^{n,t}, \mu^n) + \big(W_2(\nu^{m,t}, \nu^m) \\
&\quad + W_2(\mu^{n,t}, \mu^n)\big) W_2(\mu^n, \nu^m),
\end{aligned} \tag{2.4}$$

where we refer to (3.107) for further details. Our argument then differs from [4] by splitting the error in two parts

$$\inf_\pi W_2^2(\pi, \gamma^{n,t}) \leq \underbrace{W_2^2(\gamma^{n,t}, \gamma_\delta^{n,t})}_{\text{regularization error}} + \underbrace{\inf_\pi W_2^2(\pi, \gamma_\delta^{n,t})}_{\text{stability error}}. \tag{2.5}$$

Our proof then proceeds in two steps, controlling separately the two terms in (2.5).

Control of the regularization error To deal with the regularization error, we look at the difference $e_\delta^{n,t} := h_\delta^{n,t} - h^{n,t}$ which solves, according to (1.18) and (2.2),

$$-\nabla \cdot \rho_\delta \nabla e_\delta^{n,t} = \nabla \cdot (\rho_\delta - \rho)\nabla h^{n,t}. \tag{2.6}$$

Using an energy estimate, we get

$$\int_\mathcal{M} |\nabla e_\delta^{n,t}|^2 \lesssim \int_\mathcal{M} |\rho - \rho_\delta|^2 |\nabla h^{n,t}|^2. \tag{2.7}$$

On the one hand, since $\rho \in L^\infty$, we have $\rho_\delta \to \rho$ as $\delta \downarrow 0$ in every L^q with $q < \infty$. On the other hand, we learn from Meyers' estimate (recalled in Proposition A.3) that there exists $\bar{q} > 2$ such that $\nabla h^{n,t} \in L^{\bar{q}}$. Consequently, we can treat (2.7) using Hölder's inequality which provides

$$\int_\mathcal{M} |\nabla e_\delta^{n,t}|^2 \lesssim \|\rho_\delta - \rho\|_{L^{2(\frac{\bar{q}}{2})'}}^2 \|\nabla h^{n,t}\|_{L^{\bar{q}}}^2. \tag{2.8}$$

The next step is to control the averaged Meyers' norm $\mathbb{E}\big[\|\nabla h^{n,t}\|_{L^{\bar{q}}}^2\big]$, that we show in Proposition 3.3 to be of order of

$$\mathbb{E}\big[\|\nabla h^{n,t}\|_{L^{\bar{q}}}^2\big] \lesssim \frac{|\log(t)| + \log^{\frac{1}{\eta}}(n)}{n}, \tag{2.9}$$

where we recall that η denotes the correlation length, see Assumption 1.1.

The combination of (2.8), (2.9) and local Lipschitzianity of the exponential map yields

$$\mathbb{E}\big[W_2^2(\gamma_\delta^{n,t}, \gamma^{n,t})\big] \lesssim \|\rho_\delta - \rho\|_{L^{2(\frac{\bar{q}}{2})'}}^2 \frac{|\log(t)| + \log^{\frac{1}{\eta}}(n)}{n}, \qquad (2.10)$$

where we refer to (3.104) for further details. We emphasize here that the stretched exponential decay assumption (1.14) of the α-mixing coefficient plays a crucial role in the estimate (2.9). Indeed, the additional contribution on the numerator of the r.h.s. of (2.9) is due to the correlations and is only of logarithmic type $\log^{\frac{1}{\eta}}(n)$ thanks to the stretched exponential decay (1.14). Finally, the latter appears in the estimate (2.10) and can be compensated with the choice of δ in (2.11) and the regularity assumption on ρ in form of (2.16).

Control of the stability error For the stability error, we first need to ensure (2.3). Our strategy follows the idea in [3] which consists of showing that (2.3) is satisfied with very high probability. In our case, a new difficulty comes from our regularization of ρ and the regularization parameter δ has to be carefully optimized. We show that if δ is taken as an inverse power of $\log(n)$, (2.3) becomes very likely as $n \uparrow \infty$. More precisely, we show in Proposition 3.4 that given two exponents κ_1 and $\upsilon \gg \kappa_1$, there exists κ_2 such that given the choices

$$\delta = \frac{1}{\log^{\kappa_1}(n)} \quad \text{and} \quad t = \frac{\log^{\kappa_2}(n)}{n}, \qquad (2.11)$$

we have

$$\mathbb{P}\Big(\|\nabla h_\delta^{n,t}\|_{L^\infty} + \|\nabla^2 h_\delta^{n,t}\|_{L^\infty} > \tfrac{1}{\log^\upsilon(n)}\Big) = o\Big(\frac{1}{n^\ell}\Big) \quad \text{for any } \ell \in \mathbb{N}. \quad (2.12)$$

This result is in the spirit of [3, Theorem 3.3] but, in our setting, the proof relies on Schauder's theory rather than an explicit formula for $h_\delta^{n,t}$ as well as concentration inequalities in form of Proposition A.1 to treat the correlations. For further details on the strategy, we refer to Sect. 3.3. We emphasize here that, to obtain the super-polynomial behaviour (2.12), we crucially use the fact that the concentration inequalities in Proposition A.1 are of sub-exponential type which is itself ensured by the sub-exponential decay assumption on the α-mixing coefficient (1.14). The reason lies in the choices of δ and t in (2.11). Indeed, the only room we have when optimizing t is in the logarithmic growth $\log^{\kappa_2}(n)$ (as already mention in Sect. 1.2, the natural regularization time is $t \sim \frac{1}{n}$ up to logarithmic corrections). Furthermore, the quantity $\|(\nabla h_\delta^{n,t}, \nabla^2 h_\delta^{n,t})\|_{L^\infty}$ can be heuristically estimated by an inverse power of δ and t^{-1} as it involves powers of $\|(\nabla \rho_\delta, \nabla^2 \rho_\delta)\|_{L^\infty}$ and the norm $\|\mu^{n,t} - 1\|_{L^\infty}$ by Schauder's theory. Hence, in the best case scenario where exponential concentration holds, we expect

$$\mathbb{P}\Big(\|\nabla h_\delta^{n,t}\|_{L^\infty} + \|\nabla^2 h_\delta^{n,t}\|_{L^\infty} > \lambda\Big) \lesssim \exp\big(-\lambda n \|(\nabla h_\delta^{n,t}, \nabla^2 h_\delta^{n,t})\|_{L^\infty}^{-1}\big)$$

$$\leq \exp\left(-\lambda n t \delta^{\tilde{\upsilon}}\right) = \exp\left(-\lambda \log^{\kappa_2}(n) \delta^{\tilde{\upsilon}}\right),$$

for some $\tilde{\upsilon} > 0$, which gives a super-polynomial behavior for the choices $\delta = \frac{1}{\log^{\kappa_1}(n)}$ and $\lambda = \frac{1}{\log^{\upsilon}(n)}$ for large κ_2. Weaker properties, as for instance polynomial concentration, would only lead to a decay given by an inverse power of $\log(n)$ which is not enough for our purpose.

With (2.12) in hands, we can restrict the analysis to realizations satisfying $\|\nabla h_\delta^{n,t}\|_{L^\infty} + \|\nabla^2 h_\delta^{n,t}\|_{L^\infty} \leq \frac{1}{\log^{\upsilon}(n)}$ where, for $n \gg 1$, (2.3) is satisfied which puts us in the validity of (2.4). We then treat each terms appearing in (2.4) separately:

- **The optimal control of the cost** $W_2^2(\mu^n, \nu^m)$ has been already well studied and is optimally estimated by

$$\mathbb{E}\left[W_2^2(\mu^n, \nu^m)\right] \lesssim \frac{\log(n)}{n}. \tag{2.13}$$

We refer to Appendix B for a detailed statement, references and extensions to the cases of Assumption 1.1 and Assumption 1.5.

- **The smoothing errors** $W_2^2(\mu^n, \mu^{n,t})$ and $W_2^2(\nu^m, \nu^{m,t})$. Classical contractivity estimates are known and are usually applied to deal with these errors, see for instance [23, Theorem 3], which bound the errors by t. However, due to the choice of t in (2.11), this result is of no use since t is much larger than the magnitude of the cost, namely $t \gg \frac{\log(n)}{n}$. Instead, we follow the approach in [3], where the authors showed that in the particular case of empirical measures in dimension 2, we can improve the rate and obtain the bound $\frac{\log\log(n)}{n} \ll \frac{\log(n)}{n}$. We extend this result to our setting of non-constant densities and correlated points. In Proposition 3.5, we derive

$$\mathbb{E}\left[W_2^2(\mu^n, \mu^{n,t})\right] + \mathbb{E}\left[W_2^2(\nu^m, \nu^{m,t})\right] \lesssim \frac{\log\log(n)}{n} + t\|\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}\|_{L^1}. \tag{2.14}$$

As opposed to [3], our approach uses Fourier analysis together with additional cares to handle the correlations and non-constant densities.

- **The error in the Moser coupling** $W_2^2\left(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\#\mu^{n,t}\right)$. We follow the strategy in [3]. This error can be related to a Moser coupling between $\mu^{n,t}$ and $\nu^{m,t}$ (see for instance [62, Appendix p. 16]): The equation (1.18) gives a natural coupling between $\mu^{n,t}$ and $\nu^{m,t}$ which can be formulated using Benamou–Brenier's theorem [10],

$$\nu^{m,t} = \phi(1, \cdot)_\#\mu^{n,t} \quad \text{with} \phi \text{being the flow induced by} s \mapsto \frac{\rho_\delta \nabla h_\delta^{n,t}}{s\mu^{n,t} + (1-s)\nu^{m,t}}.$$

Then, using the transport plan $\left(\phi(1, \cdot), \exp(\nabla h_\delta^{n,t})\right)_\#\mu^{n,t}$ as a competitor, we get

$$W_2^2\left(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\#\mu^{n,t}\right) = W_2^2\left(\phi(1, \cdot)_\#\mu^{n,t}, \exp(\nabla h_\delta^{n,t})_\#\mu^{n,t}\right)$$

$$\leq \int_{\mathcal{M}} \mathrm{d}^2\big(\phi(1, \cdot), \exp(\nabla h_\delta^{n,t})\big),$$

that we combine with a quantitative stability result for flows of vector fields, proved in [3, Proposition A.1], leading to

$$W_2^2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\big) \lesssim \big(\|\rho_\delta - \rho\|_{L^{2(\frac{\bar q}{2})'}}^2 + \|\rho_t - \rho\|_{L^{2(\frac{\bar q}{2})'}}^2 + \tfrac{1}{\log^\upsilon(n)}\big)\|\nabla h_\delta^{n,t}\|_{L^{\bar q}}^2,$$

where we recall that $\bar q$ denotes the Meyers' exponent introduced in (2.7). For further details, we refer to (3.115). It then remains to appeal to Meyers' estimate, see Proposition A.3, to (2.6) together with (2.9) in form of

$$\mathbb{E}\big[\|\nabla h_\delta^{n,t}\|_{L^{\bar q}}^2\big] \lesssim \mathbb{E}\big[\|\nabla h^{n,t}\|_{L^{\bar q}}^2\big] \overset{(2.9)}{\lesssim} \frac{|\log(t)| + \log^{\frac{1}{\eta}}(n)}{n},$$

which finally yields

$$\mathbb{E}\big[W_2^2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\big)\big]$$
$$\lesssim \big(\|\rho_\delta - \rho\|_{L^{2(\frac{\bar q}{2})'}}^2 + \|\rho_t - \rho\|_{L^{2(\frac{\bar q}{2})'}}^2 + \tfrac{1}{\log^\upsilon(n)}\big)\frac{|\log(t)| + \log^{\frac{1}{\eta}}(n)}{n}. \quad (2.15)$$

To conclude, we see that all bounds involve errors in terms of the approximation of ρ using the heat-semigroup, that we need to quantify. This is where the assumption $\rho \in H^\varepsilon$ plays a role, in form of the quantitative estimate

$$\|\rho_s - \rho\|_{L^2} \lesssim \|\rho\|_{H^\varepsilon} s^\varepsilon \quad \text{for any } s > 0, \quad (2.16)$$

see (3.94) for a proof. Combining (2.4), (2.5), (2.10), (2.13), (2.14), (2.15), (2.16) with the choices of δ and t in (2.11), we obtain Theorem 1.2. The proof of Theorem 1.3 is obtained using the same strategy where the first step is simpler, since we apply directly Theorem 2.1 with $\mu_1 = \rho$, $\nu = \mu^n$ and $\mu_2 = \exp(\nabla f_\delta^{n,t})_\# \mu^n$ where $f^{n,t}$ solves $-\nabla \cdot \rho_\delta \nabla f^{n,t} = \mu^{n,t} - \rho_t$.

3 Proofs

3.1 Notations and preliminary results

We provide in this section some notations and preliminary results needed in the proofs of Theorems 1.2 and 1.3. We recall that throughout the paper, we denote by $\mathcal{M}$ a 2-dimensional compact connected Riemannian manifold (or the square $[0, 1]^2$) endowed with the Riemannian distance d.

Wasserstein distance Given $\mu, \nu \in \mathcal{P}(\mathcal{M})$, we define the quadratic Wasserstein distance as

$$W_2^2(\mu, \nu) := \min_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{M} \times \mathcal{M}} \mathrm{d}^2(x, y) \, \mathrm{d}\pi(x, y), \quad (3.1)$$

where $\Pi(\mu, \nu)$ is the set of couplings between μ and ν, that is the set of $\pi \in \mathcal{P}(\mathcal{M} \times \mathcal{M})$ having μ and ν as first and second marginal, respectively. We refer the reader to the monographs [61, 62] for a detailed overview on the subject of optimal transport. We recall the following simple, but useful Lipschitz contraction property of the Wasserstein distance.

Lemma 3.1 (Lipschitz property of the Wasserstein metric) *Let (D, d_D) be a complete and separable metric space, let $\mu, \nu \in \mathcal{P}(\mathcal{M})$ and let $T : \mathcal{M} \to D$ be a L-Lipschitz map. It holds*

$$W_2^2(T_\# \mu, T_\# \nu) \leq L^2 W_2^2(\mu, \nu). \tag{3.2}$$

Proof For any coupling $\pi \in \Pi(\mu, \nu)$, the push-forward $(T, T)_\# \pi$ is a coupling between $T_\# \mu$ and $T_\# \nu$. Moreover, it holds

$$\int_{D \times D} \mathrm{d}^2(x, y) \, \mathrm{d}((T, T)_\# \pi)(x, y) = \int_{\mathcal{M} \times \mathcal{M}} \mathrm{d}^2(T(x), T(y)) \, \mathrm{d}\pi(x, y)$$
$$\leq L^2 \int_{\mathcal{M} \times \mathcal{M}} \mathrm{d}^2(x, y) \, \mathrm{d}\pi(x, y).$$

Taking the infimum among all possible couplings $\pi \in \Pi(\mu, \nu)$ leads to (3.2). $\qquad\square$

Heat semigroup and heat kernel We recall some basic facts on the heat semigroup and its generator, we refer the reader to [19, Chapter VI] for a more detailed overview. For $t > 0$, we denote by p_t the fundamental solution of the heat operator $\partial_t - \Delta$ on $\mathcal{M}$, where Δ denotes the Beltrami-Laplace operator. Classical Schauder's theory ensures that p_t is smooth and it is known, see for instance [57] and [3, Appendix B], that p_t and its derivatives satisfy for some $C > 0$ depending on $\mathcal{M}$

$$|\nabla^N p_t(x, y)| \lesssim t^{-1-\frac{N}{2}} \exp\left(-\frac{1}{C} \frac{\mathrm{d}^2(x, y)}{t}\right) \quad \text{for any } t > 0, N \geq 1 \text{ and } x, y \in \mathcal{M}. \tag{3.3}$$

The kernel p_t admits the spectral decomposition

$$p_t(x, y) := \sum_{k \geq 1} e^{-t\lambda_k} \phi_k(x) \phi_k(y) \quad \text{for any } t > 0 \text{ and } x, y \in \mathcal{M}, \tag{3.4}$$

converging in $\mathrm{L}^2(\mathcal{M} \times \mathcal{M})$, where we recall that $\{\lambda_k, \phi_k\}$ denotes the eigenvalues and eigenvectors of $-\Delta$ on $\mathcal{M}$. Specifying (3.4) on the diagonal and using $\|\phi_k\|_{\mathrm{L}^2(\mathcal{M})} = 1$, we obtain the trace formula

$$\sum_{k \geq 1} e^{-t\lambda_k} = \int_{\mathcal{M}} p_t(x, x) \mathrm{dm}(x) \quad \text{for any } t > 0. \tag{3.5}$$

We recall that $\{\phi_k\}_{k \geq 1}$ can be estimated in terms of the eigenvalues,

$$\|\phi_k\|_{\mathrm{L}^\infty} \lesssim \lambda_k^{\frac{1}{2}}. \tag{3.6}$$

We briefly recall the argument. Applying the Gagliardo-Nirenberg's interpolation inequality [7, Theorem 3.70], it holds

$$\|\phi_k\|_{L^\infty} \lesssim \|\nabla^2\phi_k\|_{L^2}^{\frac{1}{2}}\|\phi_k\|_{L^2}^{\frac{1}{2}} = \|\nabla^2\phi_k\|_{L^2}^{\frac{1}{2}}.$$

In combination with $-\Delta\phi_k = \lambda_k\phi_k$ and elliptic regularity, in form of

$$\|\nabla^2\phi_k\|_{L^2} \lesssim \lambda_k\|\phi_k\|_{L^2} = \lambda_k,$$

we obtain (3.6).

We recall that $(P_t)_{t>0}$ admits the spectral gap property, that is there exists a constant $C_{sg} > 0$ such that

$$\|P_t f\|_{L^2(\mathcal{M})} \leq e^{-C_{sg}t}\|f\|_{L^2(\mathcal{M})} \quad \text{for any } f \in L^2 \text{ with } \int_{\mathcal{M}} f \, dm = 0. \quad (3.7)$$

Note that equivalently, (3.7) can be formulated in terms of the eigenvalues $\{\lambda_k\}_{k\geq 1}$ in form of

$$\inf_{k\geq 1} \lambda_k \geq C_{sg}, \quad (3.8)$$

simply by specifying (3.7) on $\{\phi_k\}_{k\geq 1}$. Finally, we recall the Riesz-transform bound

$$\int_{\mathcal{M}} |\nabla f|^4 \, dm \lesssim \int_{\mathcal{M}} \left|(-\Delta)^{\frac{1}{2}}f\right|^4 \, dm \quad \text{for any } f \text{ such that } \int_{\mathcal{M}} |\nabla f|^4 \, dm < \infty \quad (3.9)$$

where $(-\Delta)^{\frac{1}{2}}$ can be defined via its spectral representation

$$(-\Delta)^{\frac{1}{2}}f = \sum_{k\geq 1} \sqrt{\lambda_k}\,(\phi_k, f)_{L^2}\phi_k \quad \text{for any } f \in H^{\frac{1}{2}}, \quad (3.10)$$

with $(\cdot, \cdot)_{L^2}$ the inner product in L^2. We refer to the monograph [63] for a discussion of the inequalities (3.7) and (3.9), see Chapter 1 for the case of a Riemannian manifold without boundary and Chapter 2 for the case of a Riemannian manifold with (convex) boundary. In connection with the Wasserstein metric, the heat semigroup satisfies the following classical contraction property.

Lemma 3.2 (Semigroup contraction for absolutely continuous measures) *Let $\rho \in L^\infty$ satisfying* (1.6). *Given $\rho_t := P_t\rho$, it holds*

$$W_2^2(\rho_t \, dm, \rho \, dm) \lesssim t\|\rho_t - \rho\|_{L^1} \quad \text{for any } t > 0. \quad (3.11)$$

Proof Using g defined via $-\Delta g = \rho_t - \rho$ together with (1.6), [3, Corollary 4.4] yields

$$W_2^2(\rho_t \, dm, \rho \, dm) \lesssim \int_{\mathcal{M}} |\nabla g|^2 \, dm = \int_{\mathcal{M}} (\rho_t - \rho)g \, dm.$$

Writing $g = \int_0^\infty P_\tau(\rho_t - \rho)\, d\tau = -\int_0^t \rho_\tau\, d\tau$ together with $|\int_0^t \rho_\tau\, d\tau| \leq \|\rho\|_{L^\infty} t$ gives (3.11). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.2 L^q-type estimates

As we have seen in (2.8), we need a sharp control of the averaged Meyers' norm $\mathbb{E}\big[\|\nabla h^{n,t}\|_{L^{\bar{q}}}^2\big]$. This will be obtained as an immediate corollary of the following proposition, for more details see (3.101).

Proposition 3.3 (L^q-estimates) *Let $\{\mu^n\}_n$ be defined in (1.2) with point clouds satisfying Assumption 1.1. Let $\bar{q}$ be the Meyers exponent given in Theorem A.3 for the operator $-\nabla \cdot \rho\nabla$. The solution[4] $f^{n,t} \in \dot{H}^1$ of*

$$- \nabla \cdot \rho\nabla f^{n,t} = \mu^{n,t} - \rho_t, \tag{3.12}$$

satisfies:

$$\left(\int_{\mathcal{M}} |\nabla f^{n,t}|^q\, dm\right)^{\frac{2}{q}} \leq \mathcal{C}_{n,t}\frac{|\log(t)| + \log^{\frac{1}{\eta}}(n)}{n} \quad \text{for any } q \in [2, \min\{\bar{q}, 4\}], \tag{3.13}$$

and a random variable $\mathcal{C}_{n,t}$ satisfying for some generic constant C_q depending on q,

$$\sup_{n,t} \mathbb{E}[\tfrac{1}{C_q}\mathcal{C}_{n,t}] \leq 1. \tag{3.14}$$

Furthermore, if (1.14) holds with $\eta \geq 1$ then the assumption (1.13) can be dropped and the stochastic integrability can be improved up to losing a $\log(n)$ factor, namely

$$\left(\int_{\mathcal{M}} |\nabla f^{n,t}|^q\, dm\right)^{\frac{2}{q}}$$
$$\leq \mathcal{D}_{n,t}\left(\frac{\log^{\frac{1}{\eta}}(n)|\log(t)|}{n} + \frac{t^{-1}(1 + \log^2(n)\mathbb{1}_{\eta\neq\infty})}{n^2}\right) \quad \text{for any } q \in [2, \min\{\bar{q}, 4\}], \tag{3.15}$$

and a random variable $\mathcal{D}_{n,t}$ satisfying for some generic constant D_q depending on q,

$$\sup_{n,t} \mathbb{E}[\exp(\tfrac{1}{D_q}\mathcal{D}_{n,t}^{\frac{1}{2}})] \leq 2. \tag{3.16}$$

Proof We proceed in four steps. In the first step, we prove a representation formula for $(-\Delta)^{\frac{1}{2}}$ that we will use as the core tool in the next steps. In the second step, we compare the two operators $-\nabla \cdot \rho\nabla$ and $-\Delta$, with help of Meyers' estimate recalled in Theorem A.3. Doing so, we then have to bound the L^q-norms ($2 \leq q \leq \bar{q}$) of the gradient of the solution[5] to the Poisson equation with r.h.s. $\mu^{n,t} - \rho_t$ and

---

[4] With Neumann boundary conditions in case $\mathcal{M}$ has a boundary.

[5] Which belongs to any L^q for any $q < \infty$ from Calderón-Zygmund' theory, see for instance [28].

Neumann boundary conditions. We control all the norms by the L^4-norm that, in turn, we estimate using the Riesz-transform bound (3.9) and following some ideas from [6, Lemma 3.17]. In the third and fourth steps, we control the bound previously obtained in expectation where our main tool is Assumption 1.1 and the concentration inequalities in Proposition A.1.

STEP 1. A REPRESENTATION FORMULA FOR $(-\Delta)^{\frac{1}{2}}$. We show that given $f \in C^2$ such that $n_{\mathcal{M}} \cdot \nabla f = 0$ on $\partial \mathcal{M}$ we have

$$(-\Delta)^{\frac{1}{2}} f = \frac{1}{\sqrt{\pi}} \int_0^\infty \tau^{-\frac{1}{2}} \Delta P_\tau f \, d\tau. \tag{3.16}$$

Note that $(-\Delta)^{\frac{1}{2}} f$, defined in (3.10), is well defined in L^2. Indeed, using the fact that $(\phi_k, f)_{L^2} = \frac{1}{\lambda_k} (\phi_k, \Delta f)_{L^2}$ and (3.8), we have for any $N \leq M < \infty$

$$\left\| \sum_{N \leq n \leq M} \sqrt{\lambda_n} \, (\phi_n, f)_{L^2} \phi_n \right\|_{L^2}^2 = \sum_{N \leq n \leq M} \lambda_n |(\phi_n, f)_{L^2}|^2 = \sum_{N \leq n \leq M} \frac{1}{\lambda_n} |(\phi_n, \Delta f)_{L^2}|^2$$
$$\leq \frac{1}{C_{sg}} \sum_{n \geq N} |(\phi_n, \Delta f)_{L^2}|^2,$$

which vanishes as $N \uparrow \infty$ uniformly in M.

We now justify (3.16). Observe that since $n_{\mathcal{M}} \cdot \nabla f = 0$ on $\partial \mathcal{M}$,

$$\Delta P_s f = P_s \Delta f \quad \text{for any } s \in (0, \infty), \tag{3.17}$$

which is a direct consequence of two integration by parts using the heat-kernel representation $P_s f = \int_{\mathcal{M}} p_s(\cdot, y) f(y) \, dm(y)$. Therefore

$$\int_0^\infty \tau^{-\frac{1}{2}} \Delta P_\tau f \, d\tau = \int_0^\infty \tau^{-\frac{1}{2}} P_\tau \Delta f \, d\tau, \tag{3.18}$$

where the last integral is well-defined in L^2 since from (3.7)

$$\int_0^\infty t^{-\frac{1}{2}} \|P_\tau \Delta f\|_{L^2} \, d\tau \leq \int_0^\infty t^{-\frac{1}{2}} e^{-C_{sg}\tau} \|\Delta f\|_{L^2} \, d\tau < \infty.$$

We then use the spectral decomposition of the heat semigroup (3.4) to get

$$P_\tau \Delta f = \sum_n e^{-\lambda_n \tau} (\phi_n, \Delta f)_{L^2} \phi_n \quad \text{in } L^2. \tag{3.19}$$

The combination of (3.18) and (3.19) yields for any $\eta \in L^2$

$$\left(\int_0^\infty \tau^{-\frac{1}{2}} \Delta P_\tau f, \eta \right)_{L^2} d\tau = \int_0^\infty \tau^{-\frac{1}{2}} (P_\tau \Delta f, \eta)_{L^2} \, d\tau$$

$$\overset{(3.19)}{=} \int_0^\infty \tau^{-\frac{1}{2}} \sum_n e^{-\lambda_n \tau} (\phi_n, \Delta f)_{L^2} (\phi_n, \eta)_{L^2} \, d\tau. \tag{3.20}$$

Using (3.8), we have

$$\int_0^\infty \tau^{-\frac{1}{2}} \sum_n e^{-\lambda_n \tau} |(\phi_n, \Delta f)_{L^2}||(\phi_n, \eta)_{L^2}| \, d\tau$$

$$\leq \left(\int_0^\infty \tau^{-\frac{1}{2}} e^{-C_{sg}\tau} \, d\tau \right) \sum_n |(\phi_n, \Delta f)_{L^2}||(\phi_n, \eta)_{L^2}| < \infty,$$

so that we can exchange integration and summation in (3.20) to obtain

$$\left(\int_0^\infty \tau^{-\frac{1}{2}} \Delta P_\tau f, \eta \right)_{L^2} d\tau = \sum_n \left(\int_0^\infty \tau^{-\frac{1}{2}} e^{-\lambda_n \tau} \, d\tau \right) (\phi_n, \Delta f)_{L^2} (\phi_n, \eta)_{L^2}$$

$$= \sqrt{\pi} \sum_n \frac{1}{\sqrt{\lambda_n}} (\phi_n, \Delta f)_{L^2} (\phi_n, \eta)_{L^2}$$

$$= \sqrt{\pi} \sum_n \sqrt{\lambda_n} (\phi_n, f)_{L^2} (\phi_n, \eta)_{L^2} \overset{(3.10)}{=} \sqrt{\pi} ((-\Delta f)^{\frac{1}{2}}, \eta)_{L^2},$$

which gives (3.16) by arbitrariness of η. Finally, note that the r. h. s. integral in (3.16) is absolutely convergent thanks to the integration by parts $\Delta P_\tau f = P_\tau \Delta f$ and the bounds on the heat kernel (3.3), so that it defines a function in C^0.

STEP 2. COMPARISON WITH THE SOLUTION OF THE POISSON EQUATION. We claim that for any $q \in [2, \min\{\bar{q}, 4\}]$ and $p < \infty$

$$\mathbb{E}\left[\left(\int_{\mathcal{M}} |\nabla f^{n,t}|^q \, dm \right)^{\frac{2p}{q}} \right]^{\frac{1}{p}} \lesssim_q \left(\int_{\mathcal{M}} dm \left(\int_0^\infty ds \, \mathbb{E}\left[\left((-s\Delta)^{\frac{1}{2}} P_{s+t} (\mu^n - \rho) \right)^{2p} \right]^{\frac{1}{p}} \right)^2 \right)^{\frac{1}{2}}. \tag{3.21}$$

Let $g^{n,t} \in \dot{H}^1$ be the solution of the following Poisson equation

$$-\Delta g^{n,t} = \mu^{n,t} - \rho_t. \tag{3.22}$$

Re-expressing the right-hand side of (3.12) as $\nabla \cdot \nabla g^{n,t}$, we apply Meyers' estimate recalled in Theorem A.3 and Hölder's inequality to obtain:

$$\left(\int_{\mathcal{M}} |\nabla f^{n,t}|^q \, dm \right)^{\frac{2}{q}}$$

$$\lesssim \left(\int_{\mathcal{M}} |\nabla g^{n,t}|^q \, dm \right)^{\frac{2}{q}} \lesssim_q \left(\int_{\mathcal{M}} |\nabla g^{n,t}|^4 \, dm \right)^{\frac{1}{2}} \quad \text{for any } q \in [2, \min\{\bar{q}, 4\}], \tag{3.23}$$

We now introduce the Paley–Littlewood functional

$$\mathcal{L}(g) := \left(\int_0^\infty s (\partial_s \mathrm{P}_s g)^2 \, \mathrm{d}s \right)^{\frac{1}{2}} \quad \text{for any } g \in \mathrm{L}^4 \text{ and } \int_{\mathcal{M}} g = 0.$$

We recall that the inverse of $\mathcal{L}$ is continuous, see [56], namely

$$\|g\|_{\mathrm{L}^4} \lesssim \|\mathcal{L}(g)\|_{\mathrm{L}^4}. \tag{3.24}$$

Combining the Riesz transform bound (3.9) with (3.24) yields

$$\int_{\mathcal{M}} |\nabla g^{n,t}|^4 \, \mathrm{dm} \lesssim \int_{\mathcal{M}} \left(\mathcal{L}\left((-\Delta)^{\frac{1}{2}} g^{n,t} \right) \right)^4 \mathrm{dm}$$
$$= \int_{\mathcal{M}} \mathrm{dm} \left(\int_0^\infty \mathrm{d}s \, s (\partial_s \mathrm{P}_s (-\Delta)^{\frac{1}{2}} g^{n,t})^2 \right)^2. \tag{3.25}$$

We now claim that

$$\partial_s \mathrm{P}_s (-\Delta)^{\frac{1}{2}} g^{n,t} = (-\Delta)^{\frac{1}{2}} \mathrm{P}_{s+t} (\mu^n - \rho) \quad \text{for any } s \geq 0, \tag{3.26}$$

which requires a special care when $\mathcal{M}$ has a boundary. We use the definition of P_s in form of $\partial_s \mathrm{P}_s = \Delta \mathrm{P}_s$ to get

$$\partial_s \mathrm{P}_s (-\Delta)^{\frac{1}{2}} g^{n,t} = \Delta \mathrm{P}_s (-\Delta)^{\frac{1}{2}} g^{n,t}. \tag{3.27}$$

Recalling that $n_{\mathcal{M}} \cdot \nabla g^{n,t} = 0$, (3.17) implies that $\Delta \mathrm{P}_\tau g^{n,t} = \mathrm{P}_\tau \Delta g^{n,t}$ which, combined with (3.22) and (3.16), gives

$$(-\Delta)^{\frac{1}{2}} g^{n,t} = \frac{1}{\sqrt{\pi}} \int_0^\infty \tau^{-\frac{1}{2}} \mathrm{P}_\tau (\mu^{n,t} - \rho_t) \, \mathrm{d}\tau. \tag{3.28}$$

In particular, it implies that $n_{\mathcal{M}} \cdot \nabla (-\Delta)^{\frac{1}{2}} g^{n,t} = 0$. Therefore, one can use once more (3.17) and, together with (3.28), (3.27) turns into

$$\partial_s \mathrm{P}_s (-\Delta)^{\frac{1}{2}} g^{n,t} = \frac{1}{\sqrt{\pi}} \int_0^\infty \tau^{-\frac{1}{2}} \mathrm{P}_s \Delta \mathrm{P}_\tau (\mu^{n,t} - \rho_t) \, \mathrm{d}\tau. \tag{3.29}$$

Using a last time (3.17) in form of $\mathrm{P}_s \Delta \mathrm{P}_\tau (\mu^{n,t} - \rho_t) = \Delta \mathrm{P}_s \mathrm{P}_\tau (\mu^{n,t} - \rho_t)$ that we combine with the semigroup property $\mathrm{P}_t \mathrm{P}_{t'} = \mathrm{P}_{t+t'}$ yields

$$\mathrm{P}_s \Delta \mathrm{P}_\tau (\mu^{n,t} - \rho_t) = \Delta \mathrm{P}_\tau \mathrm{P}_{s+t} (\mu^n - \rho),$$

which, together with (3.29) and (3.16) leads to (3.26).

The combination of (3.23), (3.25) and (3.26) together with Minkowski's inequality gives

$$\mathbb{E}\left[\left(\int_{\mathcal{M}}|\nabla f^{n,t}|^q\,\mathrm{dm}\right)^{\frac{2p}{q}}\right]^{\frac{1}{p}} \overset{(3.23)}{\lesssim} \mathbb{E}\left[\left(\int_{\mathcal{M}}\mathrm{dm}\left(\int_0^\infty \mathrm{d}s\left((-s\Delta)^{\frac{1}{2}}\mathrm{P}_{s+t}(\mu^n-\rho)\right)^2\right)^2\right)^{\frac{p}{2}}\right]^{\frac{1}{p}}$$

$$\leq\left(\int_{\mathcal{M}}\mathrm{dm}\left(\int_0^\infty \mathrm{d}s\,\mathbb{E}\left[\left((-s\Delta)^{\frac{1}{2}}\mathrm{P}_{s+t}(\mu^n-\rho)\right)^{2p}\right]^{\frac{1}{p}}\right)^2\right)^{\frac{1}{2}},$$

which is (3.21).

STEP 3. PROOF OF (3.13). The estimate (3.13) is a consequence of (3.21) applied with $p=1$ and

$$\mathbb{E}\left[\left((-s\Delta)^{\frac{1}{2}}\mathrm{P}_{s+t}(\mu^n-\rho)(x)\right)^2\right]$$
$$\lesssim \frac{\zeta(s,t)}{n}(1+\log^{\frac{1}{\eta}}(n)s^{\frac{1}{\log(n)}}\zeta(s,t)^{1-\frac{1}{\log(n)}})\quad\text{for any }s\in(0,\infty),\quad(3.30)$$

with

$$\zeta(s,t):=\min\left\{(s+t)^{-1},(s+t)^{-2}\right\}.\tag{3.31}$$

Indeed, plugging (3.30) in (3.21) gives

$$\mathbb{E}\left[\left(\int_{\mathcal{M}}|\nabla f|^q\,\mathrm{dm}\right)^{\frac{2}{q}}\right]\lesssim\frac{1}{n}\int_0^\infty\zeta(s,t)(1+\log^{\frac{1}{\eta}}(n)s^{\frac{1}{\log(n)}}\zeta(s,t)^{1-\frac{1}{\log(n)}})\,\mathrm{d}s.$$
$$\tag{3.32}$$

Using that

$$\int_0^\infty\zeta(s,t)\,\mathrm{d}s\leq t^{-1}\int_0^t\mathrm{d}s+\int_t^1 s^{-1}\,\mathrm{d}s+\int_1^\infty s^{-2}\,\mathrm{d}s\lesssim|\log(t)|,$$

and analogously

$$\int_0^\infty s^{\frac{1}{\log(n)}}\zeta(s,t)^{1-\frac{1}{\log(n)}}\,\mathrm{d}s\lesssim 1,$$

(3.32) implies (3.13).

We now show (3.30). First, using the definition (1.2) of μ^n, we have

$$(-s\Delta)^{\frac{1}{2}}\mathrm{P}_{s+t}(\mu^n-\rho)=\frac{1}{n}\sum_{k=1}^n\omega_{s+t}(\cdot,X_k)\quad\text{with}\quad\omega_{s+t}(\cdot,y):=(-s\Delta)^{\frac{1}{2}}\mathrm{P}_{s+t}(\delta_y-\rho),$$
$$\tag{3.33}$$

such that expanding the square gives

$$\left((-s\Delta)^{\frac{1}{2}} P_{s+t}(\mu^n - \rho)(x)\right)^2$$

$$= \frac{1}{n^2} \sum_{k=1}^{n} \left(\omega_{s+t}(x, X_k)\right)^2 + \frac{2}{n^2} \sum_{1 \le \ell < \ell' \le n} \omega_{s+t}(x, X_\ell)\omega_{s+t}(x, X_{\ell'}).$$

The estimate (3.30) is then a consequence of

$$\mathbb{E}[|\omega_{s+t}(x, X_1)|^2] \lesssim \min\left\{(s+t)^{-1}, (s+t)^{-2}\right\} \quad \text{and}$$

$$\sup_{y \in \mathcal{M}} |\omega_{s+t}(x, y)| \lesssim \sqrt{s} \min\left\{(s+t)^{-\frac{3}{2}}, (s+t)^{-2}\right\}. \tag{3.34}$$

Indeed, the first item of (3.34) immediately implies

$$\mathbb{E}\left[\sum_{k=1}^{n} \left(\omega_{s+t}(\cdot, X_k)\right)^2\right] \lesssim n\zeta(s, t),$$

while, using the assumption (1.14) and that $t^{-\frac{3}{\log(n)}} \lesssim 1$,

$$\sum_{1 \le \ell < \ell' \le n} \mathbb{E}[\omega_{s+t}(\cdot, X_\ell)\omega_{s+t}(\cdot, X_{\ell'})]$$

$$= \sum_{1 \le \ell < \ell' \le n} \mathbb{E}[\omega_{s+t}(\cdot, X_\ell)\omega_{s+t}(\cdot, X_{\ell'})]^{\frac{1}{\log(n)}} \mathbb{E}[\omega_{s+t}(\cdot, X_\ell)\omega_{s+t}(\cdot, X_{\ell'})]^{1-\frac{1}{\log(n)}}$$

$$\le \sum_{1 \le \ell < \ell' \le n} \alpha_{\ell-\ell'}^{\frac{1}{\log(n)}} \left(\sup_{y \in \mathcal{M}} |\omega_{s+t}(x, y)|\right)^{\frac{2}{\log(n)}} \mathbb{E}[|\omega_{s+t}(x, X_1)|^2]^{1-\frac{1}{\log(n)}} \tag{3.35}$$

$$\overset{(3.34),(1.14)}{\lesssim} n \log^{\frac{1}{\eta}}(n) s^{\frac{1}{\log(n)}} t^{-\frac{3}{\log(n)}} \zeta(s, t)^{1-\frac{1}{\log(n)}} \lesssim \log^{\frac{1}{\eta}}(n) s^{\frac{1}{\log(n)}} \zeta(s, t)^{1-\frac{1}{\log(n)}}.$$

It remains to prove (3.34) and we start with the first item. Here, we use the fact that

$$\|(-s\Delta)^{\frac{1}{2}} P_{\frac{s}{2}} f\|_{L^2} \lesssim \|f\|_{L^2} \quad \text{for all } f \in C^2 \text{ and uniformly in } s. \tag{3.36}$$

This can be seen from (3.10): Using that $P_{\frac{s}{2}}$ is an auto-adjoint operator in L^2 and that from (3.19) we learn $P_{\frac{s}{2}} \phi_n = e^{-\lambda_n \frac{s}{2}} \phi_n$, we have

$$\|(-s\Delta)^{\frac{1}{2}} P_{\frac{s}{2}} f\|_{L^2}^2 = \sum_n (s\lambda_n) |(\phi_n, P_{\frac{s}{2}} f)_{L^2}|^2 = \sum_n (s\lambda_n) e^{-s\lambda_n} |(\phi_n, f)_{L^2}|^2 \lesssim \|f\|_{L^2}^2.$$

Hence, using the semi-group property $P_{s+t} = P_{\frac{s}{2}} P_{\frac{s}{2}+t}$, we deduce

$$\mathbb{E}[|\omega_{s+t}(x, X_1)|^2] \lesssim \|(-s\Delta)^{\frac{1}{2}} P_{s+t}(\delta_y - \rho)\|_{L^2}^2 = \|(-s\Delta)^{\frac{1}{2}} P_{\frac{s}{2}} P_{\frac{s}{2}+t}(\delta_y - \rho)\|_{L^2}^2$$

$$\overset{(3.36)}{\lesssim} \|P_{\frac{s}{2}+t}(\delta_y - \rho)\|_{L^2}^2.$$

We now bound $\|P_{\frac{s}{2}+t}(\delta_y - \rho)\|_{L^2}^2$ in two different ways. First, using the bounds (3.3) of the heat-kernel, we have by the triangle inequality

$$\|P_{\frac{s}{2}+t}(\delta_y - \rho)\|_{L^2}^2 \lesssim \|p_{\frac{s}{2}+t}(\cdot, y)\|_{L^2}^2 \overset{(3.3)}{\lesssim} (s+t)^{-1},$$

yielding to the first alternative in the first item of (3.34). Second, applying Poincaré's inequality yields

$$\|P_{\frac{s}{2}+t}(\delta_y - \rho)\|_{L^2}^2 \lesssim \|\nabla p_{\frac{s}{2}+t}(\cdot, y)\|_{L^2}^2 \overset{(3.3)}{\lesssim} (s+t)^{-2},$$

yielding to the second alternative in the first item of (3.34).

We now turn to the second item of (3.34). Here, we make use of the representation formula (3.16) applied to $f = P_{s+t}(\delta_y - \rho)$. Using the semi-group property $P_\tau P_{s+t} = P_{\tau+s+t}$, this takes the form

$$\omega_{s+t}(x, y) = \sqrt{\frac{s}{\pi}} \int_0^\infty \tau^{-\frac{1}{2}} \Delta P_{\tau+s+t}(\delta_y - \rho)(x) \, d\tau.$$

A direct application of the heat-kernel bounds (3.3) leads to

$$\sup_{y \in \mathcal{M}} |\omega_{s+t}(x, y)| \lesssim \sqrt{s} \int_0^\infty \tau^{-\frac{1}{2}} (\tau + s + t)^{-2} \, d\tau \lesssim \sqrt{s} \, (s+t)^{-\frac{3}{2}},$$

which is the first alternative in the second item of (3.34). For the second alternative, we write

$$\Delta P_{\tau+s+t}(\delta_y - \rho)(x) = \int_{\mathcal{M}} \rho \big(\Delta p_{\tau+s+t}(x, y) - \Delta p_{\tau+s+t}(x, \cdot) \big) \, dm,$$

so that, using (3.3),

$$\sup_{y \in \mathcal{M}} |\omega_{s+t}(x, y)| \lesssim \sqrt{s} \int_0^\infty \tau^{-\frac{1}{2}} \sup_{y \in \mathcal{M}} |\nabla \Delta p_{\tau+s+t}(x, y)| \, d\tau$$

$$\overset{(3.3)}{\lesssim} \int_0^\infty \tau^{-\frac{1}{2}} (\tau + s + t)^{-\frac{5}{2}} \, d\tau \lesssim \sqrt{s} \, (s+t)^{-2}.$$

STEP 4. PROOF OF (3.15). Following the same computations done in the previous step, the estimate (3.15) is a consequence of (3.21) and the moment bounds

$$\mathbb{E}\left[\left((-s\Delta)^{\frac{1}{2}} P_{s+t}(\mu^n - \rho)(x) \right)^{2p} \right]^{\frac{1}{p}}$$

$$\lesssim p^2 \left(\frac{\zeta(s,t)}{n}(1 + \log^{\frac{1}{\eta}}(n)s^{\frac{1}{2\log(n)}}\zeta(s,t)^{1-\frac{1}{\log(n)}}) + \frac{\sup_{y \in \mathcal{M}}|\omega_{s+t}(x,y)| + \log^2(n)}{n^2} \right)$$
$$\text{for any } s \in (0, \infty), \tag{3.37}$$

together with the second item of (3.34) and Lemma A.2, where we recall that ζ is defined in (3.31).

We now show (3.37). Using (3.33) and the assumption $\eta \geq 1$, we apply the concentration inequality in Proposition A.1 to the effect of: for any $\lambda \geq 0$

$$\mathbb{P}\left(\left| (-s\Delta)^{\frac{1}{2}} P_{s+t}(\mu^n - \rho)(x) \right| \geq \lambda \right)$$
$$\lesssim \exp\left(-\frac{1}{C} \frac{n^2\lambda^2}{nv^2 + \left(\sup_{y \in \mathcal{M}}|\omega_{s+t}(x,y)| \right)^2 + n\lambda \left(\sup_{y \in \mathcal{M}}|\omega_{s+t}(x,y)| \right)\log^2(n)} \right), \tag{3.38}$$

with

$$v^2 := \mathbb{E}[|\omega_{s+t}(x,X_1)|^2] + 2 \sum_{1 \leq \ell < \ell' \leq n} \left| \mathbb{E}[\omega_{s+t}(x,X_\ell)\omega_{s+t}(x,X_{\ell'})] \right|.$$

We then obtain (3.37) combining (3.38) with (3.34) and (3.35), together with an application of the Layer-cake formula. □

3.3 Fluctuation estimates

This section is devoted to justify (2.12) needed to ensure the condition (2.3) with very high probability. Our result is in the spirit of [3, Theorem 3.3]. However, our strategy differs from [3, Theorem 3.3] and is based on Schauder's theory, with an additional special care on the dependences on δ. We briefly sketch the main ingredients of the fluctuation estimates (2.12). By linearity, it is enough to show (2.12) for $f_\delta^{n,t} \in \dot{\mathrm{H}}^1$ being the solution of

$$-\nabla \cdot \rho_\delta \nabla f_\delta^{n,t} = \mu^{n,t} - \rho_t. \tag{3.39}$$

We then make use of the chain rule to expand the equation in

$$-\Delta f_\delta^{n,t} = \frac{1}{\rho_\delta}\nabla\rho_\delta \cdot \nabla f_\delta^{n,t} + \frac{1}{\rho_\delta}(\mu^{n,t} - \rho_t), \tag{3.40}$$

and we define the auxiliary problem

$$-\Delta u_\delta^{n,t} = \frac{1}{\rho_\delta}(\mu^{n,t} - \rho_t), \tag{3.41}$$

such that the difference $v_\delta^{n,t} := f_\delta^{n,t} - u_\delta^{n,t}$ solves $-\Delta v_\delta^{n,t} = \frac{1}{\rho_\delta}\nabla\rho_\delta \cdot \nabla f_\delta^{n,t}$. Doing so, on the one hand, we can control $u_\delta^{n,t}$ which can be handled using the explicit formula

in terms of the heat-kernel, the explicit bounds on the latter, cf (3.3), and the regularity of $\frac{1}{\rho_\delta}$. On the other hand, we use Schauder's estimates to control $v_\delta^{n,t}$ and $f_\delta^{n,t}$. Using the fact that those estimates depend polynomially on $\|\rho_\delta\|_{C^{0,\alpha}}$, we can keep track on the dependences on δ that we can optimize later on. We finally mention that in the case where $\mathcal{M}$ has a boundary, $u_\delta^{n,t}$ cannot be directly defined by (3.41) since the r. h. s. does not have zero mean. In order to also include this case, we add a zero order term in the equation, see (3.42).

Proposition 3.4 (Fluctuation estimates) *Let $\{\mu^n\}_n$ be defined in (1.2) with point clouds satisfying Assumption 1.1. For any parameter $\delta \in (0,1)$ and $\upsilon > 0$, we define[6] $u_\delta^{n,t} \in H^1$ weak solution of*

$$u_\delta^{n,t} - \Delta u_\delta^{n,t} = \tfrac{1}{\rho_\delta}(\mu^{n,t} - \rho_t), \tag{3.42}$$

and the two events

$$\mathcal{A}_n := \left\{ \|\mu^{n,t} - \rho_t\|_{L^\infty} \leq \tfrac{1}{\log^\upsilon(n)} \right\} \quad and \quad \mathcal{B}_{\delta,n} := \left\{ \|(\nabla u_\delta^{n,t}, \nabla^2 u_\delta^{n,t})\|_{L^\infty} \leq \tfrac{1}{\log^\upsilon(n)} \right\}. \tag{3.43}$$

There exists $\kappa > 0$ such that for any $\kappa_1 > 0$ and the choice

$$\delta = \delta_n := \tfrac{1}{\log^{\kappa_1}(n)}, \tag{3.44}$$

the solution $f_\delta^{n,t} \in \dot{H}^1$ of (3.39) satisfies

$$\mathbb{1}_{\mathcal{A}_n \cap \mathcal{B}_{\delta,n}} \left\|(\nabla f_\delta^{n,t}, \nabla^2 f_\delta^{n,t})\right\|_{L^\infty} \leq \frac{1}{\log^{\upsilon-(\kappa+2)\kappa_1}(n)}. \tag{3.45}$$

Furthermore, there exists $\kappa_2 > 0$ depending on υ such that for the choice $t = t_n := \frac{\log^{\kappa_2}(n)}{n}$, we have

$$\mathbb{P}(\mathcal{A}_n^c \cup \mathcal{B}_{\delta,n}^c) = o(\tfrac{1}{n^\ell}) \quad for\ any\ \ell \in \mathbb{N}. \tag{3.46}$$

Proof The proof of Proposition 3.4 is split into two steps. In the first step, we prove (3.45) where our main tool is Schauder's theory and an explicit formula for $u_\delta^{n,t}$ defined in (3.40). In the second step, we show (3.46), where our main tool is the concentration inequalities in Proposition A.1 and the explicit formula for $u_\delta^{n,t}$ used in the first step.

STEP 1. PROOF OF (3.45). We define the difference $v_\delta^{n,t} := f_\delta^{n,t} - \left(u_\delta^{n,t} - \int_{\mathcal{M}} u_\delta^{n,t}\right) \in \dot{H}^1$ and note that from (3.40) and (3.42), it solves

$$-\Delta v_\delta^{n,t} = \frac{1}{\rho_\delta}\nabla\rho_\delta \cdot \nabla f_\delta^{n,t} - \left(u_\delta^{n,t} - \int_{\mathcal{M}} u_\delta^{n,t}\right). \tag{3.47}$$

---

[6] With Neumann boundary conditions in case $\mathcal{M}$ has a boundary.

We claim that there exists $\kappa > 0$ such that with the choices of δ in (3.44), we have

$$\mathbb{1}_{\mathcal{A}_n \cap \mathcal{B}_{\delta,n}} \left\| (\nabla v_\delta^{n,t}, \nabla^2 v_\delta^{n,t}) \right\|_{\mathrm{L}^\infty} \lesssim \frac{1}{\log^{\upsilon - (\kappa+2)\kappa_1}(n)}. \tag{3.48}$$

The estimate (3.45) then follows from (3.48) and the triangle inequality.

We now prove (3.48). We apply Schauder's estimate (see for instance [29] and [48, Chapter 10]) to (3.47) to obtain

$$\left\| (\nabla v_\delta^{n,t}, \nabla^2 v_\delta^{n,t}) \right\|_{\mathrm{L}^\infty} \lesssim \left\| \tfrac{1}{\rho_\delta} \nabla \rho_\delta \cdot \nabla f_\delta^{n,t} \right\|_{\mathrm{C}^{0,\alpha}} + \left\| u_\delta^{n,t} - \int_{\mathcal{M}} u_\delta^{n,t} \right\|_{\mathrm{C}^{0,\alpha}}. \tag{3.49}$$

While the second r. h. s. term is directly of order of $\frac{1}{\log^\upsilon(n)}$ in $\mathcal{B}_{\delta,n}$, the first r. h. s. term requires additional attention. Using (1.6) and (3.3), ρ_δ satisfies

$$\rho_\delta \geq \lambda \quad \text{and} \quad \delta^{-1} \|\nabla \rho_\delta\|_{\mathrm{L}^\infty} + \|\nabla^2 \rho_\delta\|_{\mathrm{L}^\infty} \lesssim \delta^{-2}, \tag{3.50}$$

and together with the algebraic property of $\|\cdot\|_{\mathrm{C}^{0,\alpha}}$, we have

$$\| \tfrac{1}{\rho_\delta} \nabla \rho_\delta \cdot \nabla f_\delta^{n,t} \|_{\mathrm{C}^{0,\alpha}} \leq \| \tfrac{1}{\rho_\delta} \nabla \rho_\delta \|_{\mathrm{C}^{0,\alpha}} \| \nabla f_\delta^{n,t} \|_{\mathrm{C}^{0,\alpha}} \lesssim \delta^{-2} \| \nabla f_\delta^{n,t} \|_{\mathrm{C}^{0,\alpha}}.$$

The latter is bounded using Schauder's estimate applied this time on (3.39) (knowing that the dependence on $\|\rho_\delta\|_{\mathrm{C}^{0,\alpha}}$ is at most polynomial): there exists $\kappa > 0$ such that

$$\| \nabla f_\delta^{n,t} \|_{\mathrm{C}^{0,\alpha}} \lesssim \| \rho_\delta \|_{\mathrm{C}^{0,\alpha}}^\kappa \| \mu^{n,t} - \rho_t \|_{\mathrm{L}^\infty} \overset{(3.50)}{\lesssim} \delta^{-\kappa} \| \mu^{n,t} - \rho_t \|_{\mathrm{L}^\infty}$$

$$\overset{(3.44)}{\lesssim} \log^{\kappa\kappa_1}(n) \| \mu^{n,t} - \rho_t \|_{\mathrm{L}^\infty},$$

which yields the following control of the first r. h. s. term of (3.49)

$$\| \tfrac{1}{\rho_\delta} \nabla \rho_\delta \cdot \nabla f_\delta^{n,t} \|_{\mathrm{C}^{0,\alpha}} \lesssim \log^{(\kappa+2)\kappa_1}(n) \| \mu^{n,t} - \rho_t \|_{\mathrm{L}^\infty}, \tag{3.51}$$

which is of order of $\frac{1}{\log^{\upsilon - (\kappa+2)\kappa_1}(n)}$ in $\mathcal{A}_n$.

STEP 2. PROOF OF (3.46). We provide the arguments for (3.46) in case that (1.14) holds for $\eta < 1$ and the event $\mathcal{B}_{\delta,n}^c$ that we reduce to $\{ \| \nabla^2 u_\delta^{n,t} \|_{\mathrm{L}^\infty} \leq \frac{1}{\log^\upsilon(n)} \}$, the other cases as well as the case of $\nabla u_\delta^{n,t}$ follow by a straightforward adaptation. The estimate (3.46) is a consequence of

$$\sup_{x \in \mathcal{M}} \mathbb{P}\left(|\partial_{ij}^2 u_\delta^{n,t}(x)| \geq \tfrac{1}{2\log^\upsilon(n)} \right) \leq n \exp\left(-\frac{1}{C_1} \left(\frac{nt}{\log^\upsilon(n)} \right)^\eta \right)$$

$$+ \exp\left(-\frac{1}{C_2} \frac{n^2 t \log^{-2\upsilon}(n)}{t^{-1} + n\delta^{-4} \log^{\frac{1}{\eta}}(n)} \right) \tag{3.52}$$

$$+ \exp\left(-\frac{1}{C_3} \frac{nt^2}{\log^\upsilon(n)} \exp\left(\frac{1}{C_4} \left(\frac{nt}{\log^\upsilon(n)} \right)^{\eta(1-\eta)} \log^{-1}\left(\frac{nt}{\log^\upsilon(n)} \right) \right) \right).$$

for some constants $C_1, C_2, C_3, C_4 > 0$. To see this, let η_n be defined by

$$\eta_n = \frac{1}{2 \log^\upsilon(n) \|\nabla \partial_{ij}^2 u_\delta^{n,t}\|_{L^\infty}}. \tag{3.53}$$

By compactness of $\mathcal{M}$, we can find a η_n-net $\{x_k\}_{1 \leq k \leq N} \subset \mathcal{M}$ with $N \lesssim \eta_n^{-2}$. We note that

$$\left\{ \|\partial_{ij}^2 u_\delta^{n,t}\|_{L^\infty} > \tfrac{1}{\log^\upsilon(n)} \right\} \subset \bigcup_{k=1}^N \left\{ |\partial_{ij}^2 u_\delta^{n,t}(x_k)| > \tfrac{1}{2 \log^\upsilon(n)} \right\}. \tag{3.54}$$

Indeed, if for any $k \in \{1, \cdots, N\}$, $|\partial_{ij}^2 u_\delta^{n,t}(x_k)| \leq \tfrac{1}{2 \log^\upsilon(n)}$ then for any $x \in \mathcal{M}$ there exists $j \in \{1, \cdots, N\}$ such that

$$|\partial_{ij}^2 u_\delta^{n,t}(x)| \leq \eta_n \|\nabla \partial_{ij}^2 u_\delta^{n,t}\|_{L^\infty} + |\partial_{ij}^2 u_\delta^{n,t}(x_j)| \overset{(3.53)}{\lesssim} \frac{1}{\log^\upsilon(n)}.$$

Applying $\mathbb{P}$ on (3.54) yields

$$\mathbb{P}\left(\|\partial_{ij}^2 u_\delta^{n,t}\|_{L^\infty} \geq \tfrac{1}{\log^\upsilon(n)} \right) \leq \sum_{k=1}^N \mathbb{P}\left(|\partial_{ij}^2 u_\delta^{n,t}(x_k)| \geq \tfrac{1}{2 \log^\upsilon(n)} \right)$$

$$\lesssim \eta_n^{-2} \sup_{x \in \mathcal{M}} \mathbb{P}\left(|\partial_{ij}^2 u_\delta^{n,t}(x)| \geq \tfrac{1}{2 \log^\upsilon(n)} \right).$$

Using (3.52) and

$$\eta_n^{-2} \lesssim \log^{2\upsilon}(n) t^{-3},$$

which can be proven following the arguments leading to the second item of (3.59) below, this yields $\mathbb{P}(\{\|\nabla^2 u_\delta^{n,t}\|_{L^\infty} \leq \tfrac{1}{\log^\upsilon(n)}\}) = o(\tfrac{1}{n^m})$ for any $m \in \mathbb{N}$ for the choice $\kappa_2 > 4\kappa_1 + \tfrac{1}{\eta} + 2\upsilon$.

We now prove (3.52). We first exploit the following explicit representation formula[7] for $u_\delta^{n,t}$,

$$u_\delta^{n,t} = \int_0^\infty e^{-s} \, \mathrm{P}_s\left(\tfrac{1}{\rho_\delta}(\mu^{n,t} - \rho_t) \right) \mathrm{d}s, \tag{3.55}$$

that we expand, using the definition (1.2) of μ^n, in form of

$$u_\delta^{n,t} = \frac{1}{n} \sum_{k=1}^n \omega(\cdot, X_k) \quad \text{with } \omega(\cdot, y) := \int_0^\infty e^{-s} \, \mathrm{P}_s\left(\tfrac{1}{\rho_\delta}(p_t(\cdot, y) - \rho_t) \right) \mathrm{d}s. \tag{3.56}$$

---

[7] A simple change of variable gives that the kernel $\tilde{p}_s$ associated to $1 - \Delta$ is given by $\tilde{p}_s = e^{-s} p_s$.

Then applying the concentration inequalities in Proposition A.1, we get

$$
\mathbb{P}\Big(|\partial_{ij}^2 u_\delta^{n,t}(x)| \ge \tfrac{1}{\log^\upsilon(n)}\Big) \le n \exp\Big(-\frac{1}{C_1}\Big(\frac{n}{M \log^2(n)}\Big)^\beta\Big) + \exp\Big(-\frac{1}{C_2}\frac{n^2 \log^{-2\upsilon}(n)}{M^2 + nv^2}\Big)
$$
$$
+ \exp\Big(-\frac{1}{C_3}\frac{n\lambda}{M^2 \log^\upsilon(n)} \exp\Big(\frac{1}{C_4}\Big(\frac{n}{M \log^\upsilon(n)}\Big)^{\beta(1-\beta)} \log^{-1}\big(\tfrac{n}{M \log^\upsilon(n)}\big)\Big)\Big),
$$
(3.57)

where

$$
M := \sup_{y \in \mathcal{M}} |\partial_{ij}^2 \omega(x, y)| \quad \text{and}
$$
$$
v^2 := \mathbb{E}[|\partial_{ij}^2 \omega(x, X_1)|^2] + 2 \sup_{\ell \ge 1} \sum_{k > \ell} |\mathbb{E}[\partial_{ij}^2 \omega(x, X_\ell) \partial_{ij}^2 \omega(x, X_k)]|. \quad (3.58)
$$

The estimate (3.52) then follows from the three following estimates

$$
\mathbb{E}[|\partial_{ij}^2 \omega(x, X_1)|^2] \lesssim \delta^{-4} t^{-1}, \quad \sup_{y \in \mathcal{M}} |\partial_{ij}^2 \omega(x, y)| \lesssim t^{-1} \quad \text{and} \quad v^2 \lesssim \log^{\frac{1}{\eta}}(n) \delta^{-4} t^{-1},
$$
(3.59)

that we prove separately in the next three sub-steps.

SUB–STEP 2.1. PROOF OF THE FIRST ITEM OF (3.59). Splitting the time integral into $\int_0^t + \int_t^\infty$ and subtracting and adding back $\frac{1}{\rho_\delta(x)}$ in the first integral as well as using the semigroup property of $\{P_s\}_s$ in form of

$$
\int_0^t ds\, e^{-s}\, P_s(p_t(\cdot, y) - \rho_t) = \int_t^{2t} ds\, e^{-(s-t)}(p_s(\cdot, y) - \rho_s),
$$

we decompose ω into a regular-part $\mathcal{J}_1$ and a singular-part $\mathcal{J}_2$:

$$
\omega(x, y) = \underbrace{\frac{1}{\rho_\delta(x)} \int_t^{2t} ds\, e^{-s}(p_s(\cdot, y) - \rho_s)(x) + \int_t^\infty ds\, e^{-s} P_s\Big(\frac{1}{\rho_\delta}(p_t(\cdot, y) - \rho_t)\Big)(x)}_{=:\mathcal{J}_1(x, y)}
$$
$$
+ \underbrace{\int_0^t ds\, e^{-s} P_s\Big(\big(\frac{1}{\rho_\delta} - \frac{1}{\rho_\delta(x)}\big)(p_t(\cdot, y) - \rho_t)\Big)(x)}_{=:\mathcal{J}_2(x, y)}.
$$
(3.60)

For the regular-part $\mathcal{J}_1$, we apply directly the heat-kernel bounds (3.3) and the first item of (3.50), to obtain from the triangle inequality and Minkowski's inequality

$$
\mathbb{E}[|\partial_{ij}^2 \mathcal{J}_1(x, X_1)|^2] \lesssim \|\partial_{ij}^2 \mathcal{J}_1(x, \cdot)\|_{L^2}^2 \lesssim (\|\nabla \rho_\delta\|_{L^\infty}^4 + \|\nabla^2 \rho_\delta\|_{L^\infty}^2)\Big(\int_t^{2t} \|p_s(x, \cdot)\|_{L^2} ds\Big)^2
$$

$$+ \|\nabla \rho_\delta\|_{L^\infty}^2 \underbrace{\left(\int_t^{2t} \|\nabla p_s(x, \cdot)\|_{L^2} \, ds \right)^2 + \left(\int_t^{2t} \|\nabla^2 p_s(x, \cdot)\|_{L^2} \, ds \right)^2}_{=: \mathcal{R}_1(x)}$$

$$+ \underbrace{\left(\int_t^\infty ds \left(\int_{\mathcal{M}} dm(y) \left(\partial_{ij}^2 P_s \left(\tfrac{1}{\rho_\delta} (p_t(\cdot, y) - \rho_t) \right)(x) \right)^2 \right)^{\frac{1}{2}} \right)^2}_{=: \mathcal{R}_2(x)}.$$

$$(3.61)$$

The first r. h. s. term $\mathcal{R}_1(x)$ is dominated using directly the heat-kernel bounds (3.3) and (3.50)

$$\mathcal{R}_1(x) \lesssim \delta^{-4} t^{\frac{1}{2}} + \delta^{-2} + t^{-1} \lesssim t^{-1}. \tag{3.62}$$

For the second r. h. s. side term $\mathcal{R}_2(x)$, we first simplify the y-integral. Using that

$$\int_{\mathcal{M}} |p_t(z, w) - \rho_t(z)| \, dm(w) \lesssim 1 \quad \text{and}$$

$$\int_{\mathcal{M}} |p_t(w, y) - \rho_t(w)| \, dm(w) \lesssim 1 \quad \text{for any } z, y \in \mathcal{M}, \tag{3.63}$$

we have by Jensen's inequality and the heat-kernel bounds (3.3)

$$\int_{\mathcal{M}} \left(\partial_{ij}^2 P_s \left(\tfrac{1}{\rho_\delta} (p_t(\cdot, y) - \rho_t) \right)(x) \right)^2 dm(y)$$

$$= \int_{\mathcal{M}} \left(\int_{\mathcal{M}} dw \, \partial_{ij}^2 p_s(x, w) \tfrac{1}{\rho_\delta(w)} (p_t(w, y) - \rho_t(w)) \right)^2 dm(y)$$

$$\overset{(3.63)}{\lesssim} \int_{\mathcal{M}} |\partial_{ij}^2 p_s(x, w)|^2 \, dm(w) \lesssim s^{-3}.$$

Thus,

$$\mathcal{R}_2(x) \lesssim t^{-1}. \tag{3.64}$$

The combination of (3.61), (3.62) and (3.63) yields

$$\mathbb{E}[|\partial_{ij}^2 \mathcal{J}_1(x, X_1)|^2] \lesssim t^{-1}. \tag{3.65}$$

We now turn to the singular-part $\mathcal{J}_2$. We first apply Minkowski's inequality in form of

$$\mathbb{E}[|\mathcal{J}_2(x, X_1)|^2] \lesssim \|\partial_{ij}^2 \mathcal{J}_2(x, \cdot)\|_{L^2}^2$$

$$\lesssim \left(\int_0^t ds \left(\int_{\mathcal{M}} dm(y) \left(\partial_{ij}^2 P_s \left((\tfrac{1}{\rho_\delta} - \tfrac{1}{\rho_\delta(x)})(p_t(\cdot, y) - \rho_t) \right)(x) \right)^2 \right)^{\frac{1}{2}} \right)^2.$$

$$(3.66)$$

We then simplify the y-integral. To this aim, we bound the integrand in L^∞ using the heat-kernel bounds (3.3), (3.50) and

$$\left| \frac{1}{\rho_\delta(y)} - \frac{1}{\rho_\delta(x)} \right| \overset{(3.50)}{\lesssim} \delta^{-1} d(x, y) \quad \text{for any } x, y \in \mathcal{M}, \tag{3.67}$$

in form of

$$
\left| \partial_{ij}^2 P_s \left((\tfrac{1}{\rho_\delta} - \tfrac{1}{\rho_\delta(x)})(p_t(\cdot, y) - \rho_t) \right)(x) \right| \lesssim t^{-1} \int_{\mathcal{M}} \left| \partial_{ij}^2 \left(p_s(x, w)(\tfrac{1}{\rho_\delta(w)} - \tfrac{1}{\rho_\delta(x)}) \right) \right| dm(w)
$$
$$
\overset{(3.50),(3.67)}{\lesssim} t^{-1}\delta^{-1} \left(\int_{\mathcal{M}} |\nabla^2 p_s(x, w)| d(x, w) \, dm(w) + \int_{\mathcal{M}} |\nabla p_s(x, w)| \, dm(w) \right)
$$
$$
+ t^{-1}\delta^{-2} \int_{\mathcal{M}} |p_s(x, w)| \, dm(w)
$$
$$
\overset{(3.3)}{\lesssim} t^{-1}\delta^{-1}s^{-\frac{1}{2}} + t^{-1}\delta^{-2}. \tag{3.68}
$$

This yields together with (3.66)

$$\mathbb{E}[|\mathcal{J}_2(x, X_1)|^2] \lesssim \delta^{-4} t^{-1}. \tag{3.69}$$

To conclude, the combination of (3.60), (3.65) and (3.69) shows the first item of (3.59).

SUB–STEP 2.2. PROOF OF THE SECOND ITEM OF (3.59). We use the decomposition (3.60). For the regular-part $\mathcal{J}_1$, we argue as in (3.61) for the first term whereas the second-term is estimated using the heat-kernel bounds (3.3) in form of

$$
\left| \partial_{ij}^2 P_s \left(\tfrac{1}{\rho_\delta}(p_t(\cdot, y) - \rho_t) \right)(x) \right| = \left| \int_{\mathcal{M}} dm(w) \, \partial_{ij}^2 p_s(x, w) \tfrac{1}{\rho_\delta(w)}(p_t(w, y) - \rho_t(w)) \right|
$$
$$
\overset{(3.3),(3.50)}{\lesssim} s^{-2} \int_{\mathcal{M}} dm(w) \, |p_t(w, y) - \rho_t(w)|
$$
$$
\lesssim s^{-2},
$$

so that

$$
\left| \int_t^\infty ds \, e^{-s} \partial_{ij}^2 P_s \left(\tfrac{1}{\rho_\delta}(p_t(\cdot, y) - \rho_t) \right)(x) \right| \lesssim t^{-1}.
$$

Hence,

$$\sup_{y \in \mathcal{M}} |\mathcal{J}_1(x, y)| \lesssim t^{-1}. \tag{3.70}$$

For the singular-part $\mathcal{J}_2$, we use the bound (3.68) which directly yields

$$\sup_{y \in \mathcal{M}} |\mathcal{J}_2(x, y)| \lesssim \delta^{-1} t^{-\frac{1}{2}}. \tag{3.71}$$

The combination of (3.60), (3.70) and (3.71) gives the second item of (3.59).

SUB- STEP 2.3. PROOF OF THE THIRD ITEM OF (3.59). According to the first item of (3.59), it suffices to give the argument for the second term in the definition (3.58) of v^2. We use the assumption (1.14) together with the two first items of (3.59) in form of

$$
\begin{aligned}
\sum_{k>\ell} &|\mathbb{E}[\partial_{ij}^2\omega(x, X_i)\partial_{ij}^2\omega(x, X_j)]| \\
&= \sum_{k>\ell} |\mathbb{E}[\partial_{ij}^2\omega(x, X_\ell)\partial_{ij}^2\omega(x, X_k)]|^{\frac{1}{\log(n)}} |\mathbb{E}[\partial_{ij}^2\omega(x, X_\ell)\partial_{ij}^2\omega(x, X_k)]|^{1-\frac{1}{\log(n)}} \\
&\lesssim \sum_{k>\ell} \alpha_{k-\ell}^{\frac{1}{\log(n)}} (\sup_{y\in\mathcal{M}} |\partial_{ij}^2\omega(x, y)|)^{\frac{2}{\log(n)}} (\mathbb{E}[|\partial_{ij}^2\omega(x, X_1)|^2])^{1-\frac{1}{\log(n)}} \\
&\overset{(3.59),(1.14)}{\lesssim} \delta^{-4(1-\frac{1}{\log(n)})} t^{-1-\frac{1}{\log(n)}} \sum_{k=0}^{\infty} \exp(-b\frac{k^\eta}{\log(n)}) \\
&\lesssim \delta^{\frac{4}{\log(n)}} t^{-\frac{1}{\log(n)}} \log^{\frac{1}{\eta}}(n)\delta^{-4}t^{-1},
\end{aligned}
$$

which concludes since $\limsup_{n\uparrow\infty} \delta^{\frac{4}{\log(n)}} t^{-\frac{1}{\log(n)}} \lesssim 1$. □

3.4 Contractivity estimates

This section is devoted to the control of the smoothing errors $W_2^2(\mu^{n,t}, \mu^n)$ and $W_2^2(v^{m,t}, v^m)$ for the particular choice of t given in Proposition 3.4. The first result is in the spirit of [3, Theorem 5.2] that we extend in the case of non-uniformly distributed and correlated points. This extension requires a finer analysis of the error and the proof relies on Berry–Esseen type inequalities in the spirit of [12, Theorem 5].

Proposition 3.5 (Semigroup contraction for empirical measures) *Let $\{\mu^n\}_n$ be defined in* (1.2) *with point clouds satisfying Assumption 1.1. Given t such that Proposition 3.4 holds, we have*

$$
W_2^2(\mu^{n,t}, \mu^n) \leq \mathcal{C}_n \frac{\log\log(n)}{n} + t\|\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}\|_{L^1}, \tag{3.72}
$$

for some random variable $\mathcal{C}_n$ satisfying for $C < \infty$

$$
\sup_{n\geq 1} \mathbb{E}[\tfrac{1}{C}\mathcal{C}_n] \leq 1.
$$

Furthermore, if (1.14) *holds with $\eta \geq 1$ then the assumption* (1.13) *can be dropped and the stochastic integrability can be improved up to losing a $\log(n)$ factor, namely*

$$
W_2^2(\mu^{n,t}, \mu^n) \leq \mathcal{D}_n \frac{\log^{\frac{1}{\eta}}(n)\log\log(n)}{n} + t\|\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}\|_{L^1}, \tag{3.73}
$$

for some random variable $\mathcal{D}_n$ satisfying for $D < \infty$

$$\sup_{n \geq 1} \mathbb{E}\left[\exp\left(\tfrac{1}{D} \mathcal{D}_n^{\frac{1}{2}}\right) \right] \leq 2.$$

Proof According to the fluctuation estimates in Proposition 3.4 together with $W_2^2(\mu^{n,t}, \mu^n) \leq (\mathrm{diam}(\mathcal{M}))^2$, we can restrict the analysis in $\mathcal{A}_n$ defined in (3.43). Note that for n large enough, (1.6) yields

$$\frac{\lambda}{2} \leq \mu^{n,t} \leq \Lambda + 1 \quad \text{in } \mathcal{A}_n. \tag{3.74}$$

We split the proof into three steps. In the first step, we prove a Berry–Esseen type smoothing inequality for $W_2^2(\mu^{n,t}, \mu^n)$ which decomposes the error in a deterministic part involving ρ and a random part involving the Fourier coefficients $\{\widehat{\mu^n}(k)\}_k$ of μ^n. In the second step, we prove (3.72). In the third step, we control the fluctuations of $\{\widehat{\mu^n}(k)\}_k$ using the concentration inequalities in Proposition A.1 and deduce (3.73).

STEP 1. BERRY- - ESSEEN TYPE INEQUALITY. Recalling that we denote by $\{\lambda_k, \phi_k\}_k$ the eigenvalues and eigenfunctions of $-\Delta$ respectively, we prove that

$$W_2^2(\mu^{n,t}, \mu^n) \lesssim \frac{1}{n} + \sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1\right)^2 |\widehat{\mu^n}(k) - \widehat{\rho}(k)|^2 + t\|\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}\|_{\mathrm{L}^1}, \tag{3.75}$$

where

$$\widehat{\mu^n}(k) := \int_{\mathcal{M}} \phi_k \, d\mu^n \quad \text{and} \quad \widehat{\rho}(k) := \int_{\mathcal{M}} \rho \, \phi_k \, dm.$$

We first apply the triangle inequality and use the classical contractivity estimate in [23, Theorem 3] to get

$$W_2^2(\mu^{n,t}, \mu^n) \lesssim W_2^2(\mu^{n,\frac{1}{n}}, \mu^n) + W_2^2(\mu^{n,t+\frac{1}{n}}, \mu^{n,t}) + W_2^2(\mu^{n,t+\frac{1}{n}}, \mu^{n,\frac{1}{n}})$$

$$\lesssim \frac{1}{n} + W_2^2(\mu^{n,t+\frac{1}{n}}, \mu^{n,\frac{1}{n}}). \tag{3.76}$$

We then apply Peyre's estimate [49] to the second r. h. s., which takes the form

$$W_2^2(\mu^{n,t+\frac{1}{n}}, \mu^{n,\frac{1}{n}})$$

$$\leq 4 \sup \left\{ \left| \int_{\mathcal{M}} (\mu^{n,t+\frac{1}{n}} - \mu^{n,\frac{1}{n}}) f \, dm \right|^2 \quad \text{with} \quad \int_{\mathcal{M}} \mu^{n,t+\frac{1}{n}} |\nabla f|^2 \, dm \leq 1 \right\}. \tag{3.77}$$

Now, given an arbitrary f such that

$$\int_{\mathcal{M}} \mu^{n,t+\frac{1}{n}} |\nabla f|^2 \, dm \leq 1, \tag{3.78}$$

we split

$$\int_{\mathcal{M}} (\mu^{n,t+\frac{1}{n}} - \mu^{n,\frac{1}{n}}) f \, dm = \int_{\mathcal{M}} \left(\mu^{n,t+\frac{1}{n}} - \mu^{n,\frac{1}{n}} - (\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}) \right) f \, dm$$
$$+ \int_{\mathcal{M}} f (\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}) \, dm. \qquad (3.79)$$

For the first r. h. s. term of (3.79), we expand the integral using (3.4). Thus, together with the semigroup property of $\{P_t\}_{t>0}$ and Cauchy-Schwarz's inequality, we get

$$\int_{\mathcal{M}} \left(\mu^{n,t+\frac{1}{n}} - \mu^{n,\frac{1}{n}} - (\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}) \right) f \, dm$$
$$= \int_{\mathcal{M}} d(\mu^{n,t} - \mu^n - (\rho_t - \rho))(y) \int_{\mathcal{M}} dm(x) \, f(x) p_{\frac{1}{n}}(x, y)$$
$$\overset{(3.4)}{=} \sum_{k \geq 1} e^{-\frac{1}{n}\lambda_k} \int_{\mathcal{M}} d(\mu^{n,t} - \mu^n - (\rho_t - \rho))(y) \int_{\mathcal{M}} dm(x) \, f(x) \phi_k(x) \phi_k(y)$$
$$= \sum_{k \geq 1} e^{-\frac{1}{n}\lambda_k} \widehat{f}(k) \left(\widehat{\mu^{n,t}}(k) - \widehat{\mu^n}(k) - (\widehat{\rho_t}(k) - \widehat{\rho}(k)) \right)$$
$$\leq \left(\sum_{k \geq 1} \lambda_k |\widehat{f}(k)|^2 \right)^{\frac{1}{2}} \left(\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} |\widehat{\mu^{n,t}}(k) - \widehat{\mu^n}(k) - (\widehat{\rho_t}(k) - \widehat{\rho}(k))|^2 \right)^{\frac{1}{2}}$$

Using that from (3.74) we have $\mu^{n,t+\frac{1}{n}} \geq \frac{\lambda}{2}$ and recalling (3.78), we get

$$\left(\sum_{k \geq 1} \lambda_k |\widehat{f}(k)|^2 \right)^{\frac{1}{2}} \leq \left(\int_{\mathcal{M}} |\nabla f|^2 \, dm \right)^{\frac{1}{2}} \leq \frac{2}{\lambda} \left(\int_{\mathcal{M}} \mu^{n,t+\frac{1}{n}} |\nabla f|^2 \, dm \right)^{\frac{1}{2}} \leq \frac{2}{\lambda}.$$

Furthermore, noticing that $P_t \phi_k = e^{-t\lambda_k} \phi_k$ which implies that, since P_t is self-adjoint

$$\widehat{\mu^{n,t}}(k) = e^{-t\lambda_k} \widehat{\mu^n}(k) \quad \text{and} \quad \widehat{\rho_t}(k) = e^{-t\lambda_k} \widehat{\rho}(k),$$

we obtain

$$\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} |\widehat{\mu^{n,t}}(k) - \widehat{\mu^n}(k) - (\widehat{\rho_t}(k) - \widehat{\rho}(k))|^2 = \sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1 \right)^2 |\widehat{\mu^n}(k) - \widehat{\rho}(k)|^2.$$

This leads to

$$\int_{\mathcal{M}} d(\mu^{n,t+\frac{1}{n}} - \mu^{n,\frac{1}{n}} - (\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}})) f \lesssim \left(\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1 \right)^2 |\widehat{\mu^n}(k) - \widehat{\rho}(k)|^2 \right)^{\frac{1}{2}}.$$
$$(3.80)$$

For the second r. h. s. of (3.79), we introduce $w \in \dot{H}^1$ satisfying

$$-\Delta w = \rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}},$$

so that from an integration by parts, Cauchy-Schwarz' inequality and the combination of (3.74) and (3.78), we obtain

$$\int_{\mathcal{M}} f(\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}) \, dm = \int_{\mathcal{M}} \nabla f \cdot \nabla w \, dm \lesssim \left(\int_{\mathcal{M}} |\nabla w|^2 \, dm \right)^{\frac{1}{2}}.$$

Using then the explicit formula $w = -\int_{\frac{1}{n}}^{t+\frac{1}{n}} \rho_\tau \, d\tau$ together with $|\int_{\frac{1}{n}}^{t+\frac{1}{n}} \rho_\tau \, d\tau| \overset{(1.6)}{\leq} \Lambda t$, we get

$$\int_{\mathcal{M}} |\nabla w|^2 \, dm = \int (\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}) w \, dm \lesssim t \| \rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}} \|_{L^1},$$

so that

$$\left| \int_{\mathcal{M}} f(\rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}}) \, dm \right| \lesssim \sqrt{t} \| \rho_{t+\frac{1}{n}} - \rho_{\frac{1}{n}} \|_{L^1}^{\frac{1}{2}}. \tag{3.81}$$

The combination of (3.79), (3.80), (3.81) and (3.76) leads to (3.75).

STEP 2. PROOF OF (3.72). According to (3.75), it remains to show that

$$\mathbb{E}\left[\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1 \right)^2 |\widehat{\mu^n}(k) - \widehat{\rho}(k)|^2 \right] \lesssim \frac{\log\log(n)}{n}. \tag{3.82}$$

Writing

$$\widehat{\mu^n}(k) - \widehat{\rho}(k) = \frac{1}{n} \sum_{\ell=1}^{n} (\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)]), \tag{3.83}$$

we first expand the square in form of

$$\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1 \right)^2 |\widehat{\mu^n}(k) - \widehat{\rho}(k)|^2$$

$$= \frac{1}{n^2} \sum_{\ell=1}^{n} \sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1 \right)^2 \left(\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)] \right)^2 \tag{3.84}$$

$$+ \frac{2}{n^2} \sum_{1 \leq \ell < \ell' \leq n} \sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1 \right)^2 \left(\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)] \right) \left(\phi_k(X_{\ell'}) - \mathbb{E}[\phi_k(X_{\ell'})] \right).$$

For the first r. h. s. term of (3.84), we use the normalisation $\|\phi_k\|_{L^2} = 1$ together with (1.6) to the effect of

$$\mathbb{E}\big[|\phi_k(X_1) - \mathbb{E}[\phi_k(X_1)]|^2\big] \leq \Lambda, \tag{3.85}$$

and get

$$\mathbb{E}\left[\sum_{\ell=1}^{n}\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2(\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)])^2\right] \leq \Lambda n \sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2. \tag{3.86}$$

For the second r. h. s. term of (3.84), we use the definition of the β-mixing coefficient (1.12) together with the assumption (1.13) in form of

$$\mathbb{E}\left[\sum_{1\leq\ell<\ell'\leq n}\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2\big(\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)]\big)\big(\phi_k(X_{\ell'}) - \mathbb{E}[\phi_k(X_{\ell'})]\big)\right]$$

$$\leq \sum_{1\leq\ell<\ell'\leq n}\beta_{\ell\ell'}\sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2\phi_k(x)\phi_k(y)\right| \tag{3.87}$$

$$\overset{(1.13)}{\lesssim} n \sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2\phi_k(x)\phi_k(y)\right|.$$

The combination of (3.84), (3.86) and (3.87) yields

$$\mathbb{E}\left[\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2|\widehat{\mu^n}(k) - \widehat{\rho}(k)|^2\right]$$

$$\lesssim \frac{1}{n}\left(\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2 + \sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2\phi_k(x)\phi_k(y)\right|\right).$$

It remains to show that

$$\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2 + \sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\big(e^{-t\lambda_k} - 1\big)^2\phi_k(x)\phi_k(y)\right| \lesssim \log\log(n). \tag{3.88}$$

We only treat the second l. h. s. term of (3.88), the first term is controlled the same way. For any $x, y \in \mathcal{M}$, we expand

$$\sum_{k\geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\left(e^{-t\lambda_k}-1\right)^2 \phi_k(x)\phi_k(y)$$

$$= \sum_{k\geq 1}\frac{1}{\lambda_k}e^{-2(\frac{1}{n}+t)\lambda_k}\phi_k(x)\phi_k(y) - 2\sum_{k\geq 1}\frac{1}{\lambda_k}e^{-(\frac{2}{n}+t)\lambda_k}\phi_k(x)\phi_k(y)$$

$$+ \sum_{k\geq 1}\frac{1}{\lambda_k}e^{-\frac{2}{n}\lambda_k}\phi_k(x)\phi_k(y).$$

We then disintegrate using the spectral decomposition of the heat kernel (3.4) in form of

$$\sum_{k\geq 1}\frac{1}{\lambda_k}e^{-s\lambda_k}\phi_k(x)\phi_k(y) = \int_s^1 \mathrm{d}s\, p_s(x,y) - \sum_{k\geq 1}\frac{1}{\lambda_k}e^{-\lambda_k}\phi_k(x)\phi_k(y) \quad \text{for any } s>0,$$

so that we obtain

$$\sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\left(e^{-t\lambda_k}-1\right)^2\phi_k(x)\phi_k(y)\right|$$

$$= \sup_{x,y\in\mathcal{M}}\left|\int_{2(\frac{1}{n}+t)}^{\frac{2}{n}+t}p_s(x,y)\,\mathrm{d}s + \int_{\frac{2}{n}}^{\frac{2}{n}+t}p_s(x,y)\,\mathrm{d}s\right|$$

$$\lesssim \int_{2(\frac{1}{n}+t)}^{\frac{2}{n}+t}s^{-1}\,\mathrm{d}s + \int_{\frac{2}{n}}^{\frac{2}{n}+t}s^{-1}\,\mathrm{d}s$$

$$\lesssim \log\log(n).$$

STEP 3. PROOF OF (3.73). It is a consequence of the following fluctuation estimates

$$\mathbb{E}\big[|\widehat{\mu^n}(k)-\widehat{\rho}(k)|^{2p}\big]^{\frac{1}{p}} \lesssim p^2\frac{1}{n}\left(\log^{\frac{1}{\eta}}(n)\lambda_k^{\frac{1}{\log(n)}} + \frac{\lambda_k(1+\log^2(n))}{n}\right) \quad \text{for any } p<\infty,$$

$$\tag{3.89}$$

together with Lemma A.2. Indeed, applying Minkowski's inequality followed by (3.89) and $\lambda_k^{\frac{1}{\log(n)}}e^{-\frac{2}{n}\lambda_k}\lesssim e^{-\frac{1}{n}\lambda_k}$ yield

$$\mathbb{E}\left[\left(\sum_{k\geq 1}\frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k}\left(e^{-t\lambda_k}-1\right)^2|\widehat{\mu^n}(k)-\widehat{\rho}(k)|^2\right)^p\right]^{\frac{1}{p}}$$

$$\lesssim p^2\frac{1}{n}\sum_{k\geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}\left(e^{-t\lambda_k}-1\right)^2\left(\log^{\frac{1}{\eta}}(n) + \frac{\lambda_k(1+\log^2(n))}{n}\right),$$

and (3.73) follows from

$$\sum_{k \geq 1} \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1\right)^2 \lesssim \log\log(n) \quad \text{and} \quad \sum_{k \geq 1} e^{-\frac{1}{n}\lambda_k} \left(e^{-t\lambda_k} - 1\right)^2 \lesssim t^{-1}, \quad (3.90)$$

which is obtained the same way as in (3.88) using additionally the trace formula (3.5).

We now prove (3.89). It follows from the estimate on the probability tails

$$\mathbb{P}\big(|\widehat{\mu^n}(k) - \widehat{\rho}(k)| > \lambda\big) \lesssim \exp\left(-\frac{1}{C} \frac{n^2\lambda^2}{n \log^{\frac{1}{\eta}}(n)\lambda_k^{\frac{1}{\log(n)}} + \lambda_k + n\lambda\lambda_k^{\frac{1}{2}} \log^2(n)}\right)$$
$$\text{for any } \lambda > 0, \quad (3.91)$$

for some $C > 0$, together with a simple application of the layer-cake representation.

To see (3.91), we use (3.83) together with Proposition A.1 to obtain

$$\mathbb{P}\big(|\widehat{\mu^n}(k) - \widehat{\rho}(k)| > \lambda\big) \lesssim \exp\left(-\frac{n^2\lambda^2}{nv^2 + \|\phi_k\|_{\mathrm{L}^\infty}^2 + n\lambda\|\phi_k\|_{\mathrm{L}^\infty} \log^2(n)}\right),$$

with

$$v^2 := \mathbb{E}[|\phi_k(X_1) - \mathbb{E}[\phi_k(X_1)]|^2]$$
$$+ 2\sum_{i<j} \big|\mathbb{E}\big[(\phi_k(X_i) - \mathbb{E}[\phi_k(X_i)])(\phi_k(X_j) - \mathbb{E}[\phi_k(X_j)])\big]\big|. \quad (3.92)$$

The estimate (3.91) is then a consequence of

$$\|\phi_k\|_{\mathrm{L}^\infty} \lesssim \lambda_k^{\frac{1}{2}} \quad \text{and} \quad v^2 \lesssim \log^{\frac{1}{\eta}}(n)\lambda_k^{\frac{1}{\log(n)}}. \quad (3.93)$$

The first item of (3.93) has been treated in (3.6). For the second item of (3.93), we use (3.85) and, combined with (1.14), we obtain

$$\sum_{i<j} \big|\mathbb{E}\big[(\phi_k(X_i) - \mathbb{E}[\phi_k(X_i)])(\phi_k(X_j) - \mathbb{E}[\phi_k(X_j)])\big]\big| \lesssim \|\phi_k\|_{\mathrm{L}^\infty}^{\frac{1}{\log(n)}} \sum_{\ell \geq 0} \exp\left(-b\frac{\ell^\eta}{\log(n)}\right)$$
$$\lesssim \log^{\frac{1}{\eta}}(n)\lambda_k^{\frac{1}{\log(n)}}.$$

$\square$

3.5 Proof of Theorem 1.2: Approximation of the transport plan

We only give the arguments for (1.20), (1.21) is proved the same way using the corresponding results (3.15), (3.73) and (B.2) in the case $\eta > 2$ (where some additional comments are given if necessary along the proof). We split the proof into four steps.

In the first step, we display some preliminary estimates useful all along the proof. In the second step, we deal with the approximation error that occurs in the process of regularizing ρ into ρ_δ. In the third step, we estimate the W_2-distance for the regularized quantity using the quantitative stability result in [4, Theorem 3.2], splitting the estimates in small pieces that we control in the fourth step. We finally comment on the proof of Remark 1.4 and Theorem 1.3, which are obtained with similar techniques.

STEP 1. PRELIMINARY ESTIMATES.

Heat kernel regularization The assumption $\rho \in \mathrm{H}^\varepsilon$ provides

$$\|\rho_s - \rho\|_{\mathrm{L}^2} \lesssim \|\rho\|_{\mathrm{H}^\varepsilon} s^\varepsilon \quad \text{for any } s > 0. \tag{3.94}$$

Indeed, using the definition of the heat-kernel, Minkowski's inequality and the spectral decomposition (3.4), we have

$$
\|\rho_s - \rho\|_{\mathrm{L}^2}^2 = \int_{\mathcal{M}} \mathrm{d}m(x) \left| \int_{\mathcal{M}} \rho(y)\mathrm{d}m(y) \int_0^s \mathrm{d}\tau\, \partial_\tau p_\tau(x, y) \right|^2
$$
$$
\leq \left(\int_0^s \mathrm{d}\tau \left(\int_{\mathcal{M}} \mathrm{d}m(x) \left| \sum_{k \geq 1} \widehat{\rho}(k)\, \lambda_k e^{-\tau\lambda_k} \phi_k(x) \right|^2 \right)^{\frac{1}{2}} \right)^2.
$$

Noticing that

$$
\int_{\mathcal{M}} \mathrm{d}m(x) \left| \sum_{k \geq 1} \widehat{\rho}(k)\, \lambda_k e^{-\tau\lambda_k} \phi_k(x) \right|^2 = \sum_{k \geq 1} \lambda_k^{2\varepsilon} |\widehat{\rho}(k)|^2 \lambda_k^{2(1-\varepsilon)} e^{-2\tau\lambda_k} \lesssim \|\rho\|_{\mathrm{H}^\varepsilon}^2 \tau^{2(\varepsilon - 1)},
$$

we get that

$$
\|\rho_s - \rho\|_{\mathrm{L}^2}^2 \lesssim \|\rho\|_{\mathrm{H}^\varepsilon}^2 \left(\int_0^s \tau^{\varepsilon - 1}\, \mathrm{d}\tau \right)^2 \lesssim_\varepsilon \|\rho\|_{\mathrm{H}^\varepsilon}^2 s^{2\varepsilon}.
$$

L^∞-*estimates* Let $\kappa_1 > 0$ and $\upsilon > \max\{(\kappa + 2)\kappa_1, \frac{1}{\eta}\}$, where κ is given in Proposition 3.4. For the given choices

$$t := \frac{\log^{\kappa_2}(n)}{n} \quad \text{and} \quad \delta := \frac{1}{\log^{\kappa_1}(n)}, \tag{3.95}$$

provided in Proposition 3.4, we define $h_\delta^{n,t} \in \dot{\mathrm{H}}^1$ the weak solution of

$$-\nabla \cdot \rho_\delta \nabla h_\delta^{n,t} = \mu^{n,t} - \nu^{m,t}. \tag{3.96}$$

Note that by linearity, one can decompose $h_\delta^{n,t} = h_\delta^{(1)n,t} - h_\delta^{(2)n,t}$ with

$$-\nabla \cdot \rho_\delta \nabla h_\delta^{(1)n,t} = \mu^{n,t} - \rho_t \quad \text{and} \quad -\nabla \cdot \rho_\delta \nabla h_\delta^{(2)n,t} = \nu^{m,t} - \rho_t, \tag{3.97}$$

so that, considering $u_\delta^{n,t}$ as in (3.42) and likewise $v_\delta^{n,t}$ with $\mu^{n,t}$ replaced by $\nu^{m,t}$ and defining

$$\mathcal{A}_n := \left\{ \|\mu^{n,t} - \rho_t\|_{L^\infty} + \|\nu^{m,t} - \rho_t\|_{L^\infty} \leq \tfrac{1}{\log^\upsilon(n)} \right\} \tag{3.98}$$

as well as

$$\mathcal{B}_{\delta,n} := \left\{ \left\| \left(\nabla(u_\delta^{n,t}, v_\delta^{n,t}), \nabla^2(u_\delta^{n,t}, v_\delta^{n,t}) \right) \right\|_{L^\infty} \leq \tfrac{1}{\log^\upsilon(n)} \right\}, \tag{3.99}$$

we deduce from Proposition 3.4 and $m = m(n) \underset{n\uparrow\infty}{\sim} qn$ that

$$\left\| (\nabla h_\delta^{n,t}, \nabla^2 h_\delta^{n,t}) \right\|_{L^\infty} \lesssim \frac{1}{\log^{\upsilon - (\kappa+2)\kappa_1}(n)} \quad \text{in } \mathcal{A}_n \cap \mathcal{B}_{\delta,n}. \tag{3.100}$$

L^q-*estimates* A similar decomposition as in (3.97) of (1.18) together with (3.13) and the choice of t in (3.95) yields

$$\left(\int_\mathcal{M} |\nabla h^{n,t}|^{\bar{q}} \, dm \right)^{\frac{2}{\bar{q}}} \leq \mathcal{C}_n \frac{\log(n) + \log^{\frac{1}{\eta}}(n)}{n}, \tag{3.101}$$

where $\mathcal{C}_n$ denotes, all along the proof, a random variable which satisfies (3.14) and may change from line to line.

L^2-*regularization error* Note that from (3.96) and (1.18)

$$-\nabla \cdot \rho_\delta \nabla(h_\delta^{n,t} - h^{n,t}) = \nabla \cdot (\rho_\delta - \rho)\nabla h^{n,t}, \tag{3.102}$$

so that from an energy estimate, Hölder's inequality, (3.101) and (1.6), we obtain

$$\int_\mathcal{M} |\nabla(h_\delta^{n,t} - h^{n,t})|^2 \, dm \leq \mathcal{C}_n \|\rho_\delta - \rho\|_{L^{2(\frac{\bar{q}}{2})'}}^2 \frac{\log(n) + \log^{\frac{1}{\eta}}(n)}{n}, \tag{3.103}$$

where $\bar{q}$ denotes the Meyers' exponent of the operator $-\nabla \cdot \rho\nabla$, see Theorem A.3.

Finally, since $\inf_\pi W_2^2(\pi, \gamma^{n,t}) \leq (\mathrm{diam}(\mathcal{M}))^2$ and (3.46) holds, we can restrict our analysis in $\mathcal{A}_n \cap \mathcal{B}_{\delta,n}$ that we do for the rest of the proof.

STEP 2. REGULARIZATION ERROR. We show that (3.103) survives when measuring the W_2-distance, namely

$$W_2^2(\gamma_\delta^{n,t}, \gamma^{n,t}) \leq \mathcal{C}_n \|\rho_\delta - \rho\|_{L^{2(\frac{\bar{q}}{2})'}}^2 \frac{\log(n) + \log^{\frac{1}{\eta}}(n)}{n} \quad \text{with } \gamma_\delta^{n,t}$$
$$:= \left(\mathrm{Id}, \exp(\nabla h_\delta^{n,t}) \right)_\# \mu^{n,t}. \tag{3.104}$$

Using the coupling $\big((\mathrm{Id}, \exp(\nabla h_\delta^{n,t})), (\mathrm{Id}, \exp(\nabla h^{n,t}))\big)_{\#}\mu^{n,t}$ as a competitor in
(3.1) and the fact that $\|\mu^{n,t}\|_{L^\infty} \lesssim 1$ in $\mathcal{A}_n$, we have

$$
\begin{aligned}
W_2^2(\gamma_\delta^{n,t}, \gamma^{n,t}) &\leq \int_{\mathcal{M}} \mu^{n,t} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} \\
&\lesssim \int_{\mathcal{M}} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm}.
\end{aligned}
\tag{3.105}
$$

We then claim that

$$
\begin{aligned}
\int_{\mathcal{M}} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} &\lesssim \|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2 \\
+ \|\rho_\delta - \rho\|_{L^{(\frac{\bar{q}}{2})'}}^2 \frac{\log(n)}{n} &\frac{\|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2}{\mathbb{E}[\|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2]}.
\end{aligned}
\tag{3.106}
$$

which, combined with (3.105) and (3.103) yields (3.104).

We now justify (3.106). The difficulty arises from the fact that exp is not globally Lipschitz. To overcome this, we define

$$
\begin{aligned}
\mathrm{E}_n &:= \Big\{|\nabla h_\delta^{n,t} - \nabla h^{n,t}| \leq C_n^{-1} \mathbb{E}\big[\|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2\big]^{\frac{1}{2}}\Big\} \\
&\qquad \text{with } C_n := \varsigma^{-1} \|\rho_\delta - \rho\|_{L^{(\frac{\bar{q}}{2})'}} \sqrt{\frac{\log(n) + \log^{\frac{1}{\eta}}(n)}{n}},
\end{aligned}
$$

for a given ς fixed later, and we split

$$
\begin{aligned}
&\int_{\mathcal{M}} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} \\
&= \int_{\mathcal{M}} \mathbb{1}_{\mathrm{E}_n} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} + \int_{\mathcal{M}} \mathbb{1}_{\mathrm{E}_n^c} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} \\
&\leq \int_{\mathcal{M}} \mathbb{1}_{\mathrm{E}_n} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} + (\mathrm{diam}(\mathcal{M}))^2 \mathrm{m}(\mathrm{E}_n^c).
\end{aligned}
$$

For the first right-hand side integral, note that from the choice of C_n and (3.103), we can choose $\varsigma \ll 1$ uniformly in n such that in E_n the quantity $|\nabla h_\delta^{n,t} - \nabla h^{n,t}|$ can be made arbitrary small. Since exp is Lipschitz-continuous in a neighborhood of the null vector, we deduce

$$
\int_{\mathcal{M}} \mathbb{1}_{\mathrm{E}_n} \mathrm{d}^2\big(\exp(\nabla h_\delta^{n,t}), \exp(\nabla h^{n,t})\big) \, \mathrm{dm} \lesssim \|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2.
$$

For the second right-hand side term, we simply apply Markov's inequality in form of

$$
\mathrm{m}(\mathrm{E}_n^c) \leq C_n^2 \frac{\|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2}{\mathbb{E}\big[\|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^2\big]}.
$$

The combination of the two previous estimates gives (3.106).

To prove (1.21), we need to control arbitrary p-moments, according to Lemma A.2. The argument above can be easily adapted in this case by considering

$$
E_n := \left\{ |\nabla h_\delta^{n,t} - \nabla h^{n,t}| \leq C_n^{-1} \mathbb{E}\big[\|\nabla h_\delta^{n,t} - \nabla h^{n,t}\|_{L^2}^{2p} \big]^{\frac{1}{2p}} \right\}
$$

$$
\text{with } C_n := \varsigma \|\rho_\delta - \rho\|_{L^{(\frac{\bar{q}}{2})'}} \sqrt{\frac{\log^{\frac{1}{\eta}}(n) \log(n)}{n}}.
$$

We then follow the same argument, choosing $\varsigma^{-1} = O(\sqrt{p})$.

STEP 3. QUANTITATIVE STABILITY. We show that

$$
\begin{aligned}
\inf_\pi W_2^2(\pi, \gamma_\delta^{n,t}) \lesssim\ & W_2^2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\big) + W_2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\big) W_2(\mu^n, \nu^m) \\
& + W_2^2(\nu^{m,t}, \nu^m) + W_2^2(\mu^{n,t}, \mu^n) \\
& + \big(W_2(\nu^{m,t}, \nu^m) + W_2(\mu^{n,t}, \mu^n)\big) W_2(\mu^n, \nu^m),
\end{aligned} \tag{3.107}
$$

where we recall that $\gamma_\delta^{n,t}$ is defined in (3.104).

Let π be a coupling between μ^n and ν^m. We introduce a regularization parameter $s < 1$ and, smoothing the measure μ^n into $\mu^{n,s} := P_s \mu^n$, the optimal transport plan $\pi^{n,s}$ from $\mu^{n,s}$ to ν^m is represented by a transport map $T^{n,s}$, according to McCann's theorem [42], that is

$$
\pi^{n,s} = (\text{Id}, T^{n,s})_\# \mu^{n,s}.
$$

We then apply the triangle inequality in form of

$$
\begin{aligned}
W_2(\pi^{n,s}, \gamma_\delta^{n,t}) \leq\ & W_2\big((\text{Id}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,s}, (\text{Id}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\big) \\
& + W_2(\pi^{n,s}, (\text{Id}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,s}).
\end{aligned} \tag{3.108}
$$

First, using (3.100), $\nabla h_\delta^{n,t}$ is Lipschitz-continuous and $\|\nabla h_\delta^{n,t}\|_{L^\infty}$ can be made as small as possible for n large. Since $\exp$ is Lipschitz-continuous in a neighborhood of the null vector, we learn from Lemma 3.1 that

$$
W_2^2\big((\text{Id}, \exp(\nabla h_\delta^{n,t}))_\# \mu^{n,s}, (\text{Id}, \exp(\nabla h_\delta^{n,t}))_\# \mu^{n,t}\big) \lesssim W_2^2(\mu^{n,t}, \mu^{n,s}). \tag{3.109}
$$

Second, we build a competitor for the second right-hand side term of (3.108): Defining

$$
\Gamma := \big((\text{Id}, T^{n,s}), (\text{Id}, \exp(\nabla h^{n,t}))\big)_\# \mu^{n,s},
$$

we have

$$
W_2^2(\pi^{n,s}, (\text{Id}, \exp(\nabla h^{n,t})_\# \mu^{n,s}) \leq \int_{\mathcal{M} \times \mathcal{M} \times \mathcal{M} \times \mathcal{M}} \delta^2\big((x,z), (y,w)\big) d\Gamma\big((x,y), (z,w)\big)
$$

$$= \int_{\mathcal{M}} \delta^2\big((x, T^{n,s}(x)), (x, \exp(\nabla h^{n,t}(x)))\big) \, \mu^{n,s}(x) \, dm(x)$$
$$\overset{(1.15)}{=} \int_{\mathcal{M}} d^2\big(T^{n,s}, \exp(\nabla h_\delta^{n,t})\big)\mu^{n,s} \, dm.$$

Using again (3.100) we can apply, for large n, the quantitative stability result of transport maps, Theorem 2.1, to $\mu_1 = \nu^m$, $\mu_2 = \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,s}$ and $\nu = \mu^{n,s}$ to the effect of

$$\int_{\mathcal{M}} d^2\big(T^{n,s}, \exp(\nabla h_\delta^{n,t})\big)\mu^{n,s} \, dm \lesssim W_2^2\big(\nu^m, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,s}\big)$$
$$+ W_2\big(\nu^m, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,s}\big) W_2\big(\mu^{n,s}, \nu^m\big),$$

which turns into, using the triangle inequality,

$$\int_{\mathcal{M}} d^2\big(T^{n,s}, \exp(\nabla h_\delta^{n,t})\big)\mu^{n,s} \, dm$$
$$\lesssim W_2^2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,t}\big)$$
$$+ W_2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,t}\big) W_2\big(\mu^{n,s}, \nu^m\big) \qquad (3.110)$$
$$+ W_2^2(\nu^{m,t}, \nu^m) + W_2^2\big((\mathrm{Id}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,s}), (\mathrm{Id}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,t})\big)$$
$$+ \big(W_2(\nu^{m,t}, \nu^m) + W_2((\mathrm{Id}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,s}), (\mathrm{Id}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,t}))\big) W_2(\mu^{n,s}, \nu^m).$$

The combination of (3.108), (3.109) and (3.110) yields

$$W_2^2(\pi^{n,s}, \gamma_\delta^{n,t}) \lesssim W_2^2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,t}\big) + W_2\big(\nu^{m,t}, \exp(\nabla h_\delta^{n,t})_{\#}\mu^{n,t}\big) W_2(\mu^{n,s}, \nu^m)$$
$$+ W_2^2(\nu^{m,t}, \nu^m) + W_2^2(\mu^{n,t}, \mu^{n,s}) + \big(W_2(\nu^{m,t}, \nu^m) \qquad (3.111)$$
$$+ W_2(\mu^{n,t}, \mu^{n,s})\big) W_2(\mu^{n,s}, \nu^m).$$

Since $\mu^{n,s} \underset{s\downarrow 0}{\rightharpoonup} \mu^n$, and consequently (up to extracting a subsequence) $\pi^{n,s} \underset{s\downarrow 0}{\rightharpoonup} \pi$, for some optimal transport plan π, according to the qualitative stability result [62, Theorem 5.20], we can pass to the limit as $s \downarrow 0$ in (3.111) which leads to (3.107).

STEP 4. PROOF OF (1.20). We now fix $\kappa_1 = \frac{1}{\eta}(\frac{\bar{q}}{2})\frac{1}{2\varepsilon} + 1$ such that, applying (3.94), the regularization error (3.104) turns into, recalling that δ is given by (3.95),

$$W_2^2(\gamma_\delta^{n,t}, \gamma^{n,t}) \le C_n \delta^{\frac{2\varepsilon}{(\frac{\bar{q}}{2})'}} \frac{\log(n) + \log^{\frac{1}{\eta}}(n)}{n} \le C_n \frac{1}{n}. \qquad (3.112)$$

It remains to show that

$$\inf_\pi W_2(\pi, \gamma_\delta^{n,t})$$
$$\le C_n \frac{\log(n)}{n} \Big(\sqrt{\log^{\kappa_2-1}(n) \|\rho_{t+\frac{1}{n}} - \rho_t\|_{L^1}} + \|\rho_\delta - \rho\|_{L^{2(\frac{\bar{q}}{2})'}}^2 \qquad (3.113)$$
$$+ \|\rho_t - \rho\|_{L^{2(\frac{\bar{q}}{2})'}}^2 + \frac{1}{\log^{\upsilon}(n)} + \sqrt{\frac{\log\log n}{\log n}} \Big),$$

which together with (3.94) and (3.112) leads to (1.20). To show (3.113), we control each terms of (3.107) separately.

The three last terms are controlled using the contractivity estimate (3.72) and (B.1) which gives

$$W_2^2(v^{m,t}, v^m) + W_2^2(\mu^{n,t}, \mu^n) + \left(W_2(v^{m,t}, v^m) + W_2(\mu^{n,t}, \mu^n)\right) W_2(\mu^n, v^m)$$

$$\leq C_n \frac{\log(n)}{n} \left(\sqrt{\frac{\log\log(n)}{\log(n)}} + \sqrt{\log^{\kappa-1}(n) \|\rho_{t+\frac{1}{n}} - \rho_t\|_{L^1}} \right).$$

For the first two terms, we argue that

$$W_2^2\left(v^{m,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\right) \leq C_n \left(\|\rho_\delta - \rho\|^2_{L^2(\frac{\bar q}{2})'} + \|\rho_t - \rho\|^2_{L^2(\frac{\bar q}{2})'} + \frac{1}{\log^\nu(n)} \right) \frac{\log(n)}{n}, \tag{3.114}$$

which combined with (B.1) leads to (3.113).

Let us define the curve $\eta : s \in [0,1] \mapsto \eta_s := s\mu^{n,t} + (1-s)v^{m,t}$ and note that from (1.18) we have

$$\frac{d}{ds}\eta_s + \nabla \cdot \left(\eta_s \frac{\rho_\delta \nabla h_\delta^{n,t}}{\eta_s} \right) = 0.$$

Applying Benamou–Brenier' theorem [10], we learn that

$$v^{m,t} = \phi(1, \cdot)_\# \mu^{n,t} \quad \text{with} \phi \text{ is the flow induced by} s \mapsto \frac{\rho_\delta \nabla h_\delta^{n,t}}{\eta_s}.$$

Next, using that

$$\left| \frac{\rho_\delta \nabla h_\delta^{n,t}}{\eta_s} - \nabla h_\delta^{n,t} \right| \lesssim \left(|\rho_\delta - \rho| + |\rho_t - \rho| + |\mu^{n,t} - \rho_t| + |v^{m,t} - \rho_t| \right) |\nabla h_\delta^{n,t}|$$

$$\lesssim \left(|\rho_\delta - \rho| + |\rho_t - \rho| + \frac{1}{\log^\nu(n)} \right) |\nabla h_\delta^{n,t}|,$$

and applying [3, Proposition A.1] together with Hölder's inequality yields

$$W_2^2\left(v^{m,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\right) = W_2^2\left(\phi(1,\cdot)_\# \mu^{n,t}, \exp(\nabla h_\delta^{n,t})_\# \mu^{n,t}\right)$$

$$\lesssim \int_{\mathcal{M}} \left(|\rho_\delta - \rho| + |\rho_t - \rho| + \frac{1}{\log^\nu(n)} \right)^2 |\nabla h_\delta^{n,t}|^2$$

$$\leq \left(\|\rho_\delta - \rho\|^2_{L^2(\frac{\bar q}{2})'} + \|\rho_t - \rho\|^2_{L^2(\frac{\bar q}{2})'} + \frac{1}{\log^\nu(n)} \right) \left(\int_{\mathcal{M}} |\nabla h_\delta^{n,t}|^{\bar q} \right)^{\frac{2}{\bar q}}. \tag{3.115}$$

Using Meyers' estimate of Proposition A.3 to (3.102) together with (1.6) and (3.101) provides

$$\left(\int_{\mathcal{M}} |\nabla h_\delta^{n,t}|^{\bar{q}}\right)^{\frac{2}{\bar{q}}} \lesssim \left(\int_{\mathcal{M}} |\nabla h^{n,t}|^{\bar{q}}\right)^{\frac{2}{\bar{q}}} \overset{(3.101)}{\leq} C_n \frac{\log(n) + \log^{\frac{1}{\eta}}(n)}{n},$$

which, combined with (3.115), yields (3.114).

We finally point out that, in the case $\eta > 2$, we use (3.73) and (B.2) and the same computations lead to (1.21).

STEP 5. PROOF OF THEOREM 1.3 AND REMARK 1.4. The proof of Theorem 1.3 follows the same strategy with the main difference that Step 3 is now dropped and Theorem 2.1 is directly applied with $\mu_1 = \rho$, $\nu = \mu^n$ and $\mu_2 = \exp(\nabla f_\delta^{n,t})_\# \mu^n$ where $f_\delta^{n,t}$ solves

$$-\nabla \cdot \rho_\delta \nabla f^{n,t} = \mu^{n,t} - \rho_t.$$

The improvement of Remark 1.4 follows from the improved contractivity estimate (3.73): Under the assumption 1.25, we have (keeping the notations as in Proposition 3.5)

$$W_2^2(\mu^{n,t}, \mu^n) \leq \mathcal{D}_n \frac{\log\log(n)}{n}, \tag{3.116}$$

i.e. we do not have the loss $\log^{\frac{1}{\eta}}(n)$ in (3.73). Inspecting the proof of (3.73), the loss $\log^{\frac{1}{\eta}}(n)$ comes from estimating v^2 defined in (3.92). We obtain (3.116) by simply using (1.25) and (1.11) to upgrade the second item of (3.93) into

$$v^2 \lesssim 1 + \sum_{\ell \geq 0} \exp(-b\ell^\eta) \lesssim 1.$$

Declarations

Appendix A: Probabilistic and PDE tools

This section is devoted to recall some probabilistic and analytical tools needed in the proofs. We first recall some concentration inequalities for sequences of random variable satisfying Assumption 1.1. Originally proved for i.i.d. samples, see for instance [20, Theorem 3.6 & 3.7], the proofs in the correlated case can be found in [45, Theorem 1] and [44, Theorem 2].

Proposition A.1 *Let $n \in \mathbb{N}$, $M > 0$, $\{X_i\}_i$ be a family of centred random variables such that $\sup_{i \geq 1} |X_i| \leq M$ for which (1.14) holds.*

For any $\lambda > 0$, it holds for some constants $(C_i)_{i \in \{1, \cdot, 5\}}$ depending on a, b:

(i) If $\eta < 1$,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right| > \lambda\right) \leq n \exp\left(-\frac{1}{C_1}\left(\frac{n\lambda}{M}\right)^{\eta}\right) + \exp\left(-\frac{1}{C_2}\frac{n^2\lambda^2}{M^2 + nv^2}\right)$$
$$+ \exp\left(-\frac{1}{C_3}\frac{n\lambda}{M^2}\exp\left(\frac{1}{C_4}\left(\frac{n\lambda}{M}\right)^{\eta(1-\eta)}\log^{-1}\left(\frac{n\lambda}{M}\right)\right)\right),$$

with

$$v^2 := \sup_{i \geq 1}\left(\mathbb{E}[X_i^2] + 2\sum_{j > i}|\mathbb{E}[X_i X_j]|\right).$$

(ii) If $\eta = 1$,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i\right| > \lambda\right) \leq \exp\left(-\frac{1}{C_5}\frac{n^2\lambda^2}{nv^2 + M^2 + n\lambda M(\log(n))^2}\right).$$

We then recall the link between algebraic moments and exponential moments. The proof is a direct consequence of the Taylor expansion of the exponential function.

Lemma A.2 *Let X be a non-negative random variable. The following two statements are equivalent:*

(i) There exists $C_1 > 0$ such that

$$\mathbb{E}\left[\exp\left(\frac{1}{C_1}X\right)\right] \leq 2.$$

(ii) There exists $C_2 > 0$ such that

$$\mathbb{E}[X^p]^{\frac{1}{p}} \leq p\,C_2 \quad \text{for any } p < \infty.$$

We conclude this section by recalling the standard Meyers' estimate for elliptic equations in divergence form, see for instance the original paper [46].

Theorem A.3 (Meyers estimate) *Let* $a : \mathcal{M} \to \mathbb{R}^{2 \times 2}$ *be measurable and uniformly elliptic. Consider* $u \in H^1$ *the solution of the Neumann boundary problem*

$$\begin{cases} -\nabla \cdot a\nabla u = \nabla \cdot g \text{ in } \mathcal{M}, \\ a\nabla u \cdot n_{\mathcal{M}} = 0 \quad \text{ on } \partial \mathcal{M}, \end{cases}$$

for some $g \in L^q$ *with* $q > 2$. *There exists* $2 < \bar{q} < q$ *such that*

$$\nabla u \in L^{\bar{q}} \quad \text{and} \quad \|\nabla u\|_{L^{\bar{q}}} \lesssim \|g\|_{L^{\bar{q}}}.$$

Appendix B: Matching cost for point clouds

This section is devoted to recall the upper bounds on the matching cost, results which can be found in [12, Theorem 2] under mild β-mixing conditions. The case of Markov chains have been studied in [24, 51] where sharp upper bounds are obtained. We include a short proof for convenience.

Proposition B.1 (Matching cost) *Let* ρ *satisfying* (1.6) *and* $\{\mu^n\}_n$ *be defined in* (1.2) *with point clouds satisfying the Assumption 1.1 or in the class of Markov chains satisfying the Assumption 1.5. There exists a constant* $C > 0$ *such that*

$$W_2^2(\mu^n, \rho \, dm) \leq C_n \frac{\log(n)}{n} \quad \text{with} \quad \sup_{n \geq 1} \mathbb{E}\big[\tfrac{1}{C} C_n\big] \leq 1. \tag{B.1}$$

Furthermore, if (1.14) *holds with* $\eta \geq 1$ *then the assumption* (1.13) *can be dropped and the stochastic integrability can be improved up to losing a* $\log(n)$ *factor, namely*

$$W_2^2(\mu^n, \rho \, dm) \leq D_n \frac{\log^{\frac{1}{\eta}}(n) \log(n)}{n} \quad \text{with} \quad \sup_{n \geq 1} \mathbb{E}\big[\exp(\tfrac{1}{C} D_n)\big] \leq 1. \tag{B.2}$$

Proof Note that the proof of (B.1) can be found in [12, Theorem 2] when the point cloud satisfies the Assumption 1.1. We first show how (B.1) can be extended to point clouds which are sampled from a Markov chain satisfying the Assumption 1.5. Second, we show how the stochastic integrability can be improved to (B.2) when (1.14) holds with $\eta \geq 1$.

STEP 1. MARKOV CHAINS CASE. Recall that a Markov chain satisfying the Assumption 1.5 admits an absolutely continuous invariant measure of the form $\mu_\infty = \rho \, dm$ with ρ satisfying (1.6), that is $\lambda \leq \rho \leq \Lambda$. Recalling that we denote by $\{\lambda_n, \phi_n\}_n$ the set of eigenvalues and normalized eigenfunctions of the Laplace-Beltrami operator $-\Delta$ on $\mathcal{M}$, we have by definition (1.2) of μ^n, for any $k \geq 1$

$$\widehat{\mu^n}(k) - \widehat{\rho}(k) = \frac{1}{n} \sum_{\ell=1}^n (\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)]) + \frac{1}{n} \sum_{\ell=1}^n (\mathbb{E}[\phi_k(X_\ell)] - \mu_\infty(\phi_k)), \tag{B.3}$$

where we use interchangeably the notation $\widehat{\rho}(k) = \int \phi_k \rho \, dm = \mu_\infty(\phi_k)$. Using the Berry–Esseen smoothing inequality [12, Theorem 5] together with (B.3), we get

$$
\mathbb{E}[W_2^2(\mu^n, \mu_\infty)] \lesssim \frac{1}{n} + \sum_{k \geq 1} \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k} \mathbb{E}\left[\left(\frac{1}{n}\sum_{\ell=1}^n (\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)])\right)^2\right]
$$
$$
+ \sum_{k \geq 1} \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k} \mathbb{E}\left[\left(\frac{1}{n}\sum_{\ell=1}^n (\mathbb{E}[\phi_k(X_\ell)] - \mu_\infty(\phi_k))\right)^2\right]. \tag{B.4}
$$

We now estimate the last two terms of (B.4) separately and we start with the third one. Using (1.33) and (3.93), we have

$$
|\mathbb{E}[\phi_k(X_\ell)] - \mu_\infty(\phi_k)| \overset{(1.33)}{\lesssim} \exp(-b\ell^\eta)\|\phi_k\|_{L^\infty} \overset{(3.93)}{\lesssim} \exp(-b\ell^\eta)\lambda_k^{\frac{1}{2}},
$$

Thus, using in addition (3.5), we get

$$
\sum_{k \geq 1} \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k} \mathbb{E}\left[\left(\frac{1}{n}\sum_{\ell=1}^n (\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)])\right)^2\right]
$$
$$
\lesssim \frac{1}{n^2}\sum_{k \geq 1} e^{-\frac{1}{n}\lambda_k} \overset{(3.5)}{=} \frac{1}{n^2}\int_\mathcal{M} p_{\frac{1}{n}}(x, x)\, dm(x) \lesssim \frac{1}{n}. \tag{B.5}
$$

We now turn to the second term of (B.4). Expanding the square provides

$$
\sum_{k \geq 1} \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}\left(\frac{1}{n}\sum_{\ell=1}^n (\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)])\right)^2
$$
$$
= \frac{1}{n^2}\sum_{k \geq 1}\sum_{\ell=1}^n \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}|\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)]|^2 \tag{B.6}
$$
$$
+ \frac{2}{n^2}\sum_{k \geq 1}\sum_{1 \leq \ell < \ell' \leq n} \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}(\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)])(\phi_k(X_{\ell'}) - \mathbb{E}[\phi_k(X_{\ell'})]).
$$

We now estimate the two terms on the right hand side of (B.6). For the first term, an easy induction argument combining (1.29) and (1.28) show that for any $n \geq 1$, $\mathbb{P}_{X_n} \ll m$ with $\lambda \leq \frac{d\mathbb{P}_{X_n}}{dm} \leq \Lambda$. Therefore, we have

$$
\mathbb{E}\left[|\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)]|^2\right] \leq \Lambda,
$$

and we deduce

$$
\frac{1}{n^2}\sum_{k \geq 1}\sum_{\ell=1}^n \frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}\mathbb{E}\left[|\phi_k(X_\ell) - \mathbb{E}[\phi_k(X_\ell)]|^2\right] \lesssim \frac{1}{n}\sum_{k \geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}. \tag{B.7}
$$

For the second term, we use (1.32) to obtain

$$\frac{2}{n^2}\mathbb{E}\left[\sum_{k\geq 1}\sum_{1\leq \ell<\ell'\leq n}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}(\phi_k(X_\ell)-\mathbb{E}[\phi_k(X_\ell)])(\phi_k(X_{\ell'})-\mathbb{E}[\phi_k(X_{\ell'})])\right]$$

$$\lesssim \frac{1}{n^2}\sum_{1\leq \ell<\ell'\leq n}\beta_{\ell\ell'}\sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}\phi_k(x)\phi_k(y)\right|$$

$$\lesssim \frac{1}{n}\sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}\phi_k(x)\phi_k(y)\right|.$$

Combining the latter with (B.4), (B.5), (B.6) and (B.7) yields

$$\mathbb{E}[W_2^2(\mu^n,\mu_\infty)]\lesssim \frac{1}{n}+\frac{1}{n}\sum_{k\geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}+\frac{1}{n}\sup_{x,y\in\mathcal{M}}\left|\sum_{k\geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}\phi_k(x)\phi_k(y)\right|. \quad \text{(B.8)}$$

We finally conclude similarly as for (3.88).

STEP 2. HIGHER STOCHASTIC INTEGRABILITY. We now prove (B.2). We argue using the moment estimate (3.89) which, together with Minkowski's inequality and $\lambda_k^{\frac{1}{\log(n)}}e^{-\frac{1}{n}\lambda_k}\lesssim e^{-\frac{1}{2n}\lambda_k}$ implies

$$\mathbb{E}\left[\left(\sum_{k\geq 1}\frac{e^{-\frac{1}{n}\lambda_k}}{\lambda_k}|\widehat{\mu^n}(k)-\widehat{\rho}(k)|^2\right)^p\right]^{\frac{1}{p}}\lesssim p^2\frac{1}{n}\sum_{k\geq 1}\frac{e^{-\frac{1}{2n}\lambda_k}}{\lambda_k}\left(\log^{\frac{1}{\eta}}(n)+\frac{\lambda_k(1+\log^2(n))}{n}\right).$$

Finally, combining the latter with the Berry–Esseen smoothing inequality [12, Theorem 5] and arguing similarly as for (3.88) yields (B.2) thanks to Proposition A.2.
□

Appendix C. Proof for the class of Markov chains

We provide in this Section the arguments for extending Theorems 1.2 and 1.3 to the class of Markov chains introduced in Sect. 1.4. The proof follows the lines of the proof of Theorem 1.2, where the main difference is that we drop the assumption that the point clouds is identically distributed. That affects the proofs of the main ingredients (we recall that the scaling of the cost has already be proven in Proposition B.1), namely the L^q estimates in Proposition 3.3, the fluctuation estimates in Proposition 3.4 and the contractivity estimates in Proposition 3.5. We show in the following how to adapt the proofs for a given Markov chain $\{X_n\}_{n\geq 1}$ satisfying Assumption 1.5. In the following, we recall that $\mu_\infty=\rho\, dm$ denotes the unique invariant measure of the chain. We split the proof into three steps.

STEP 1. L^q ESTIMATES. We have to understand the extra error term coming from the deviation of $\mathbb{E}[\mu^n]$ from μ_∞. In view of (3.21), it is

$$\left(\int_{\mathcal{M}} dm \left(\int_0^\infty ds \left((-s\Delta)^{\frac{1}{2}} P_{s+t} (\mathbb{E}[\mu^n] - \mu_\infty) \right)^2 \right)^2 \right)^{\frac{1}{2}}.$$

Using the definition (1.2) of μ^n, the convergence to equilibrium (1.33) applied to $f = (-s\Delta)^{\frac{1}{2}} p_{s+t}(x, \cdot)$ and the heat-kernel estimates (3.3), we have for any $s \geq 0$ and $x \in \mathcal{M}$

$$\begin{aligned}
&(-s\Delta)^{\frac{1}{2}} P_{s+t} \big(\mathbb{E}[\mu^n] - \rho \big)(x) \\
&= \frac{1}{n} \sum_{\ell=1}^n (\mathbb{E}[(-s\Delta)^{\frac{1}{2}} p_{s+t}(x, X_\ell)] - \mu_\infty((-s\Delta)^{\frac{1}{2}} p_{s+t}(x, \cdot)) \\
&\overset{(1.33)}{\lesssim} \frac{\|(-s\Delta)^{\frac{1}{2}} p_{s+t}(x, \cdot)\|_{L^\infty}}{n} \overset{(3.3)}{\lesssim} \frac{s^{\frac{1}{2}}(s+t)^{-\frac{3}{2}}}{n},
\end{aligned} \tag{C.1}$$

so that, recalling $t = \frac{\log^\kappa(n)}{n}$,

$$\begin{aligned}
&\left(\int_{\mathcal{M}} dm \left(\int_0^\infty ds \left((-s\Delta)^{\frac{1}{2}} P_{s+t} (\mathbb{E}[\mu^n] - \mu_\infty) \right)^2 \right)^2 \right)^{\frac{1}{2}} \lesssim \frac{1}{n^2} \int_0^\infty s(s+t)^{-3} \\
&\lesssim \frac{1}{n \log^\kappa(n)} \ll \frac{\log(n)}{n}.
\end{aligned}$$

STEP 2. FLUCTUATION ESTIMATES. Here, the distribution of the Markov chain affects the concentration estimate (3.52). We show that, defining

$$\bar{u}_\delta^{n,t} := \int_0^\infty e^{-s} P_s \left(\frac{1}{\rho_\delta} (\mu^{n,t} - \mathbb{E}[\mu^{n,t}]) \right) ds, \tag{C.2}$$

we have

$$\mathbb{P}\left(|\partial_{ij}^2 u_\delta^{n,t}(x)| \geq \frac{1}{2\log^\nu(n)} \right) \leq \mathbb{P}\left(|\partial_{ij}^2 \bar{u}_\delta^{n,t}(x)| \geq \frac{1}{4\log^\nu(n)} \right) \quad \text{for any } x \in \mathcal{M}, \tag{C.3}$$

where the r.h.s can be estimated following the lines of the proof of (3.52). As before, we investigate the extra term coming from the deviation of $\mathbb{E}[\mu^n]$ from μ_∞. The estimate (C.3) follows from

$$\|\partial_{ij}^2 (u_\delta^{n,t} - \bar{u}_\delta^{n,t})\|_{L^\infty} \ll \frac{1}{\log^\nu(n)}. \tag{C.4}$$

We argue as in (3.60), decomposing $u_\delta^{n,t} - \bar{u}_\delta^{n,t}$ into a regular-part and a singular part: for any $x \in \mathcal{M}$

$$
\begin{aligned}
(u_\delta^{n,t} - \bar{u}_\delta^{n,t})(x) =& \frac{1}{\rho_\delta(x)} \int_0^\infty e^{-s} \left(\mathbb{E}[\mu^{n,t+s}] - \rho_{t+s}\right)(x) \, ds \\
&+ \int_0^\infty e^{-s} \, \mathrm{P}_s\left(\left(\tfrac{1}{\rho_\delta} - \tfrac{1}{\rho_\delta(x)}\right)\left(\mathbb{E}[\mu^{n,t}] - \rho_t\right)\right)(x) \, ds.
\end{aligned}
\tag{C.5}
$$

To estimate the second r.h.s integral of (C.5), we use (3.68). For the first r.h.s integral, that we denote by $\mathcal{J}$, we use the definition (1.2) of μ^n, the convergence to equilibrium (1.33) and the heat-kernel bounds (3.3) to obtain

$$
\begin{aligned}
|\partial_{ij}^2 \mathcal{J}(x)| =& \left| \frac{1}{n} \sum_{k=1}^n \partial_{ij}^2 \left(\frac{1}{\rho_\delta(\cdot)} \int_0^\infty e^{-s} \left(\mathbb{E}[p_{t+s}(\cdot, X_k)] - \mu_\infty(p_{t+s}(x, \cdot))\right) ds \right)(x) \right| \\
\lesssim& \frac{1}{n} \sum_{k=1}^n \left(\|\nabla^2 \tfrac{1}{\rho_\delta}\|_{\mathrm{L}^\infty} \int_0^\infty |\mathbb{E}[p_{s+t}(\cdot, X_k)] - \mu_\infty(p_{t+s}(x, \cdot))| \right. \\
&+ \int_0^\infty |\mathbb{E}[\nabla^2 p_{s+t}(\cdot, X_k)] - \mu_\infty(\nabla^2 p_{t+s}(x, \cdot))| \\
&+ \left. \|\nabla \tfrac{1}{\rho_\delta}\|_{\mathrm{L}^\infty} \int_0^\infty |\mathbb{E}[\nabla p_{s+t}(\cdot, X_k)] - \mu_\infty(\nabla p_{t+s}(x, \cdot))| \right) \\
\overset{(1.33),(3.3)(3.50)}{\lesssim}& \frac{1}{n} \left(\delta^{-2} \int_0^\infty \min\{(s+t)^{-1}, (s+t)^{-\frac{3}{2}}\} \, ds \right. \\
&+ \left. \int_0^\infty (s+t)^{-2} \, ds + \delta^{-1} \int_0^\infty (s+t)^{-\frac{3}{2}} \, ds \right) \\
\lesssim& \frac{1}{n}(\delta^{-2} t^{-\frac{1}{2}} + t^{-1} + \delta^{-1} t^{-\frac{1}{2}}) \ll \frac{1}{\log^\nu(n)}.
\end{aligned}
$$

STEP 3. CONTRACTIVITY ESTIMATE. Here, the law of the Markov chain affects the estimate (3.82). The extra error term coming from the deviation of $\mathbb{E}[\mu^n]$ from μ_∞ reads

$$
\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} (e^{-t\lambda_k} - 1)^2 |\mathbb{E}[\widehat{\mu^n}(k)] - \mu_\infty(\phi_k)|^2.
$$

Using the definition (1.2) of μ^n and the convergence to equilibrium (1.33) applied with $f = \phi_k$ and the bound on the eigenfunctions (3.6), we have for any $k \leq n$

$$
|\mathbb{E}[\widehat{\mu^n}(k)] - \mu_\infty(\phi_k)| = \frac{1}{n} \left| \sum_{\ell=1}^n (\mathbb{E}[\phi_k(X_\ell)] - \mu_\infty(\phi_k)) \right|^2 \overset{(1.33)}{\lesssim} \frac{1}{n} \|\phi_k\|_{\mathrm{L}^\infty} \overset{(3.6)}{\lesssim} \frac{1}{n} \lambda_k^{\frac{1}{2}},
$$

so that, using the trace formula (3.5) and the heat-kernel estimates (3.3), we deduce

$$\sum_{k \geq 1} \frac{e^{-\frac{2}{n}\lambda_k}}{\lambda_k} \left(e^{-t\lambda_k} - 1\right)^2 |\mathbb{E}[\widehat{\mu^n}(k)] - \mu_\infty(\phi_k)|^2$$

$$\lesssim \frac{1}{n^2} \sum_{k \geq 1} e^{-\frac{2}{n}\lambda_k} (e^{-t\lambda_k} - 1)^2 \overset{(3.5),(3.3)}{\lesssim} \frac{1}{n} \ll \frac{\log\log(n)}{n}.$$

References

1. Ajtai, M., Komlós, J., Tusnády, G.: On optimal matchings. Combinatorica **4**, 259–264 (1984)
2. Alsmeyer, G.: On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. J. Theor. Probab. **16**(1), 217–247 (2003)
3. Ambrosio, L., Glaudo, F.: Finer estimates on the 2-dimensional matching problem. J. l'École Polytechnique-Mathématiques **6**, 737–765 (2019)
4. Ambrosio, L., Glaudo, F., Trevisan, D.: On the optimal map in the 2-dimensional random matching problem. Discrete Contin. Dyn. Syst. **39**(12), 7291–7308 (2019)
5. Ambrosio, L., Goldman, M., Trevisan, D.: On the quadratic random matching problem in two-dimensional domains. Electron. J. Probab. **27**, 1–35 (2022)
6. Ambrosio, L., Stra, F., Trevisan, D.: A PDE approach to a 2-dimensional matching problem. Probab. Theory Relat. Fields **173**(1), 433–477 (2019)
7. Aubin, T.: Some Nonlinear Problems in Riemannian Geometry. Springer Science & Business Media, Berlin (1998)
8. Bapat, R.B., Raghavan, T.E.S.: Doubly Stochastic Matrices. Encyclopedia of Mathematics and its Applications, pp. 59–114. Cambridge University Press (1997)
9. Barthe, F., Bordenave, C.: Combinatorial Optimization Over Two Random Point Sets, pp. 483–535. Springer International Publishing, Heidelberg (2013)
10. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math. **84**(3), 375–393 (2000)
11. Benedetto, D., Caglioti, E.: Euclidean random matching in 2D for non-constant densities. J. Stat. Phys. **181**(3), 854–869 (2020)
12. Borda, B.: Empirical measures and random walks on compact spaces in the quadratic Wasserstein metric. Ann. l'Institut Henri Poincare (B) Probab. stat. **59**(4), 2017–2035 (2023)
13. Boutet de Monvel, J.H., Martin, O.C.: Almost sure convergence of the minimum bipartite matching functional in Euclidean space. Combinatorica **22**(4), 523–530 (2002)
14. Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions. Commun. Pure Appl. Math. **44**, 375–417 (1991)
15. Caglioti, E., Pieroni, F.: Random matching in 2d with exponent 2 for densities defined on unbounded sets. arXiv preprint arXiv:2302.02602 (2023)
16. Caracciolo, S., Lucibello, C., Parisi, G., Sicuro, G.: Scaling hypothesis for the Euclidean bipartite matching problem. Phys. Rev. E **90**(1), 012118 (2014)
17. Caracciolo, S., Sicuro, G.: Scaling hypothesis for the Euclidean bipartite matching problem. II. correlation functions. Phys. Rev. E **91**(6), 062125 (2015)
18. Chafaï, D., Hardy, A., Maïda, M.: Concentration for Coulomb gases and Coulomb transport inequalities. J. Funct. Anal. **275**(6), 1447–1483 (2018)
19. Chavel, I.: Eigenvalues in Riemannian Geometry. Academic press, Cambridge (1984)
20. Chung, F., Lu, L.: Concentration inequalities and martingale inequalities: a survey. Internet Math. **3**(1), 79–127 (2006)
21. Crawford, V.P., Knoer, E.M.: Job matching with heterogeneous firms and workers. Econometrica **49**(2), 437–450 (1981)
22. Douc, Randal, Fort, Gersende, Moulines, Eric, Soulier, Philippe: Practical drift conditions for subgeometric rates of convergence. Ann. Appl. Probab. **14**(3), 1353–1377 (2004)

23. Erbar, M., Kuwada, K., Sturm, K.-T.: On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. **201**(3), 993–1071 (2015)
24. Fournier, N., Guillin, A.: On the rate of convergence in Wasserstein distance of the empirical measure. Probab. Theory Relat. Fields **162**(3), 707–738 (2015)
25. Gale, D., Shapley, L.S.: College admissions and the stability of marriage. Am. Math. Mon. **69**(1), 9–15 (1962)
26. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**(2), 113–161 (1996)
27. García-Zelada, D.: Concentration for Coulomb gases on compact manifolds. Electron. Commun. Probab. **24**, 1–18 (2019)
28. Giaquinta, M., Martinazzi, L.: An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs. Springer Science & Business Media, Berlin (2013)
29. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order, vol. 224. Springer, Berlin (2015)
30. Goldman, M., Huesmann, M.: A fluctuation result for the displacement in the optimal matching problem. Ann. Probab. **50**(4), 1446–1477 (2022)
31. Goldman, M., Huesmann, M., Otto, F.: A large-scale regularity theory for the Monge-Ampere equation with rough data and application to the optimal matching problem. arXiv:1808.09250 (2018)
32. Goldman, M., Trevisan, D.: Convergence of asymptotic costs for random Euclidean matching problems. Probab. Math. Phys. **2**, 121–142 (2021)
33. Goldman, M., Trevisan, D.: Optimal transport methods for combinatorial optimization over two random point sets. Probab. Theory Relat. Fields (2023). https://doi.org/10.1007/s00440-023-01245-1
34. Huesmann, M., Mattesini, F., Trevisan, D.: Wasserstein asymptotics for the empirical measure of fractional Brownian motion on a flat torus. Stoch. Process. Appl. **155**, 1–26 (2023)
35. Jalowy, J.: The Wasserstein distance to the circular law. Ann. Inst. Henri Poincaré Probab. Stat. **59**(4), 2285–2307 (2023)
36. Koch, L.: Geometric linearisation for optimal transport with strongly p-convex cost. arXiv preprint arXiv:2303.10760 (2023)
37. Ledoux, M.: On optimal matching of Gaussian samples. J. Math. Sci. **238**(4), 495–522 (2019)
38. Ledoux, M.: On optimal matching of Gaussian samples II (2019)
39. Ledoux, M., Zhu, J.-X.: On optimal matching of gaussian samples. Probab. Math. Stat. **41**, 237–265 (2021)
40. Liebscher, E.: Towards a unified approach for proving geometric ergodicity and mixing properties of nonlinear autoregressive processes. J. Time Ser. Anal. **26**(5), 669–689 (2005)
41. Lovász, L., Plummer, M.D.: Matching Theory, vol. 367. American Mathematical Society, Providence (2009)
42. McCann, R.J.: Polar factorization of maps on Riemannian manifolds. Geom. Funct. Anal. GAFA **11**(3), 589–608 (2001)
43. Mehta, A.: Online matching and Ad allocation. Found. Trends® Theor. Comput. Sci. **8**(4), 265–368 (2013)
44. Merlevède, F., Peligrad, M., Rio, E.: Bernstein inequality and moderate deviations under strong mixing conditions. In: High Dimensional Probability V: the Luminy Volume. Institute of Mathematical Statistics, pp. 273–292 (2009)
45. Merlevède, F., Peligrad, M., Rio, E.: A Bernstein type inequality and moderate deviations for weakly dependent sequences. Probab. Theory Relat. Fields **151**(3), 435–474 (2011)
46. Meyers, N.G.: An L^p-estimate for the gradient of solutions of second order elliptic divergence equations. Annali della Scuola Normale Superiore di Pisa-Classe di Scienze **17**(3), 189–206 (1963)
47. Mézard, M., Parisi, G.: The Euclidean matching problem. J. Phys. France **49**(12), 2019–2025 (1988)
48. Nicolaescu, L.I.: Lectures on the Geometry of Manifolds. World Scientific, Singapore (2020)
49. Peyre, R.: Comparison between W2 distance and $\dot{H}$-1 norm, and localization of Wasserstein distance. ESAIM Control Optim. Calc. Var. **24**(4), 1489–1501 (2018)
50. Prod'Homme, M.: Contributions to the optimal transport problem and its regularity. Université Paul Sabatier—Toulouse III, Theses (2021)
51. Riekert, A.: Convergence rates for empirical measures of Markov chains in dual and Wasserstein distances. Stat. Probab. Lett. **189**, 109605 (2022)

52. Santambrogio, F.: Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling. Progress in Nonlinear Differential Equations and Their Applications, Springer International Publishing, New York (2015)
53. Serfaty, S.: Coulomb gases and Ginzburg–Landau vortices. arXiv preprint arXiv:1403.6860 (2014)
54. Sicuro, G.: Euclidean Matching Problems, pp. 59–118. Springer, Cham (2017)
55. Steele, J.M.: Probability Theory and Combinatorial Optimization. SIAM, New Delhi (1997)
56. Stein, E.M.: Topics in Harmonic Analysis Related to the Littlewood-Paley Theory. (AM-63). Princeton University Press, Princeton (2016)
57. Stroock, D.W., Turetsky, J.: Upper bounds on derivatives of the logarithm of the heat kernel. Commun. Anal. Geom. **6**(4), 669–685 (1998)
58. Talagrand, M.: The Ajtai-Komlós-Tusnády matching theorem for general measures. In: Probability in Banach Spaces, 8: Proceedings of the Eighth International Conference. Springer, pp. 39–54 (1992)
59. Tao, T., Vu, V., Krishnapur, M.: Random matrices: Universality of ESDs and the circular law. Ann. Probab. **38**(5), 2023–2065 (2010)
60. Toth, J.A., Zelditch, S.: Riemannian manifolds with uniformly bounded eigenfunctions. Duke Math. J. **111**(1), 97–132 (2002)
61. Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence (2003)
62. Villani, C.: Optimal Transport: Old and New. Grundlehren der mathematischen Wissenschaften, Springer, Berlin Heidelberg (2008)
63. Wang, F.-Y.: Analysis for Diffusion Processes on Riemannian Manifolds, vol. 18. World Scientific, Singapore (2014)
64. Wang, F.-Y.: Convergence in Wasserstein distance for empirical measures of semilinear SPDEs. Ann. Appl. Probab. **33**(1), 70–84 (2023)
65. Wang, F.-Y., Wu, B.: Wasserstein convergence for empirical measures of subordinated diffusions on Riemannian manifolds. Potential Anal. **59**(3), 933–954 (2023)
66. Wang, F.-Y., Zhu, J.-X.: Limit theorems in Wasserstein distance for empirical measures of diffusion processes on Riemannian manifolds. Ann. Inst. Henri Poincaré Probab. Stat. **59**(1), 437–475 (2023)