

Entropic turnpike estimates for the kinetic Schrödinger problem*

Alberto Chiarini[†] Giovanni Conforti[‡] Giacomo Greco[§]
Zhenjie Ren[¶]

Abstract

We investigate the *kinetic* Schrödinger problem, obtained considering Langevin dynamics instead of Brownian motion in Schrödinger’s thought experiment. Under a quasilinearity assumption we establish exponential entropic turnpike estimates for the corresponding Schrödinger bridges and exponentially fast convergence of the entropic cost to the sum of the marginal entropies in the long-time regime, which provides as a corollary an entropic Talagrand inequality. In order to do so, we benefit from recent advances in the understanding of classical Schrödinger bridges and adaptations of Bakry–Émery formalism to the kinetic setting. Our quantitative results are complemented by basic structural results such as dual representation of the entropic cost and the existence of Schrödinger potentials.

Keywords: Schrödinger problem; Langevin dynamics; long-time behavior of entropic cost; turnpike estimates; gamma calculus.

MSC2020 subject classifications: 93E20; 47D07; 60E15.

Submitted to EJP on September 7, 2021, final version accepted on September 6, 2022.

Supersedes arXiv:2108.09161.

*Giovanni Conforti acknowledges funding from the grant SPOT (ANR-20-CE40-0014). Giacomo Greco acknowledges support from NWO Research Project 613.009.111 “Analysis meets Stochastics: Scaling limits in complex systems”. This research was also partially funded by Nuffic in the framework of the Van Gogh Programme under the title “The kinetic Schrödinger Problem”.

[†]Università degli Studi di Padova, Department of Mathematics Tullio Levi-Civita, 35121 Padova, Italy.

E-mail: chiarini@math.unipd.it

[‡]École Polytechnique, Département de Mathématiques Appliquées, Palaiseau, France.

E-mail: giovanni.conforti@polytechnique.edu

[§]Eindhoven University of Technology, Department of Mathematics and Computer Science, 5600 MB Eindhoven, The Netherlands.

E-mail: g.greco@tue.nl

[¶]Université Paris-Dauphine, Ceremade, PSL Research University, 75016 Paris, France.

E-mail: ren@ceremade.dauphine.fr

1 Introduction and statement of the main results

In two seminal contributions [43, 44] E. Schrödinger considered the problem of finding the most likely evolution of a cloud of independent Brownian particles conditionally to observations. This problem is nowadays known as *Schrödinger problem* and may be viewed [38, 35] as a more regular and probabilistic proxy for the Optimal transport (Monge-Kantorovich) problem. This observation has motivated recent interest from both the engineering and statistical machine learning communities [16, 41]. Moreover, over the past few years, various kinds of Schrödinger problems have been introduced and studied in the literature with different aims and scopes such as, for example, the multiplicative Schrödinger problem [40] and the mean field Schrödinger problem [1]. In this article we investigate the *Kinetic Schrödinger Problem*, henceforth KSP, with particular emphasis on the long-time and ergodic behaviour of the corresponding Schrödinger bridges. A heuristic formulation of KSP is naturally given in terms of the celebrated Schrödinger's thought experiment. Consider a system of $N \gg 1$ independent stationary particles $(X_t^1, \dots, X_t^N)_{t \in [0, T]}$ evolving according to the Langevin dynamics

$$\begin{cases} dX_t^i = V_t^i dt, \\ dV_t^i = -\nabla U(X_t^i) dt - \gamma V_t^i dt + \sqrt{2\gamma} dB_t^i, \quad i = 1, \dots, N, \end{cases}$$

and assume that two snapshots of the particle system at the initial time $t = 0$ and at the terminal time $t = T$ have been taken. The Schrödinger problem is that of finding the most likely evolution of the particle system conditionally on this information. In order to turn this heuristic description into a sound mathematical problem, we introduce the empirical path measure

$$\mu^N := \frac{1}{N} \sum \delta_{(X^i, V^i)}$$

that is a random probability measure on the space $C([0, T]; \mathbb{R}^{2d}) := \Omega$ of continuous trajectories and consider two probability measures μ, ν on \mathbb{R}^d , representing the observed configuration at initial and final time, that is to say

$$\frac{1}{N} \sum_{i=1}^N \delta_{X_0^i} \approx \mu, \quad \frac{1}{N} \sum_{i=1}^N \delta_{X_T^i} \approx \nu.$$

Then, leveraging Sanov's Theorem [22, Theorem 6.2.10], whose message is that the likelihood of a given evolution ρ is measured through the relative entropy

$$\text{Prob}[\mu^N \approx \rho] \approx \exp(-N\mathcal{H}(\rho|\mathbb{R})),$$

we finally arrive at the variational problem

$$\mathcal{C}_T(\mu, \nu) := \inf \left\{ \mathcal{H}(\mathbb{P}|\mathbb{R}) : \mathbb{P} \in \mathcal{P}(C([0, T]; \mathbb{R}^{2d})), (X_0)_{\#}\mathbb{P} = \mu, (X_T)_{\#}\mathbb{P} = \nu \right\}. \quad (\text{KSPd})$$

In the above, \mathbb{R} is the reference probability measure, that is the law of

$$\begin{cases} dX_t = V_t dt \\ dV_t = -\nabla U(X_t) dt - \gamma V_t dt + \sqrt{2\gamma} dB_t \\ (X_0, V_0) \sim \mathbf{m}, \end{cases} \quad (1.1)$$

where the invariant (probability) measure \mathbf{m} is given by

$$\mathbf{m}(dx, dv) = \frac{1}{Z} e^{-U(x) - \frac{|v|^2}{2}} dx dv,$$

with Z being a normalising constant. Moreover, $(X_t, V_t)_{t \in [0, T]}$ denotes the canonical process on Ω , $\#$ is the push-forward and $\mathcal{H}(\cdot | \mathbb{R})$ is the relative entropy functional defined on $\mathcal{P}(\Omega)$ as

$$\mathcal{H}(\mathbb{P} | \mathbb{R}) := \begin{cases} \mathbb{E}_{\mathbb{P}} \left[\log \frac{d\mathbb{P}}{d\mathbb{R}} \right] & \text{if } \mathbb{P} \ll \mathbb{R}, \\ +\infty & \text{otherwise.} \end{cases}$$

Given that Schrödinger’s thought experiment is motivated by statistical mechanics and the physical relevance of the Langevin dynamics and its various applications, the study of the kinetic Schrödinger problem appears to be quite natural. Nevertheless, to the best of our knowledge, it seems that there has been no dedicated study so far, with the exception of [15]. The objective of this paper is to take some steps forward in this direction, in particular by gaining a quantitative understanding of optimal solutions, called Schrödinger bridges.

Turnpike property for Schrödinger bridges

The turnpike property is a general principle in optimal control theory stipulating that solutions of dynamic control problems are made of three pieces: first a rapid transition from the initial state to the steady state, the *turnpike*, then a long stationary phase localised around the turnpike, and finally another rapid transition to reach the final state. In order to link this concept to KSPd, we need to rephrase it as a stochastic control problem. This task is easily accomplished thanks to classical results on the representation of path measures with finite entropy, see e.g. [26, 36] and we get that KSPd is equivalent to

$$\inf \left\{ \mathcal{H}((X_0, V_0)_{\#} \mathbb{P} | \mathfrak{m}) + \frac{1}{4\gamma} \mathbb{E}_{\mathbb{P}} \left[\int_0^T |\alpha_t^{\mathbb{P}}|^2 dt \right] : \mathbb{P} \in \mathcal{P}(\Omega), \mathbb{P} \text{ admissible} \right\}, \quad (1.2)$$

where a path probability measure \mathbb{P} is admissible if and only if under \mathbb{P} , there exist a Brownian motion $(B_t)_{t \in [0, T]}$ adapted to the canonical filtration and an adapted process $(\alpha_t^{\mathbb{P}})_{t \in [0, T]}$ such that $\mathbb{E}_{\mathbb{P}}[\int_0^T |\alpha^{\mathbb{P}}|^2 dt] < +\infty$ and the canonical process satisfies

$$\begin{cases} dX_t = V_t dt, \\ dV_t = -\nabla U(X_t) dt - \gamma V_t dt + \alpha_t^{\mathbb{P}} dt + \sqrt{2\gamma} dB_t, \\ X_0 \sim \mu, X_T \sim \nu. \end{cases} \quad (1.3)$$

For the control problem (1.2), the turnpike is the invariant measure \mathfrak{m} . Indeed, the natural tendency of the particle system is that of reaching configuration \mathfrak{m} and since Schrödinger bridges aim at approximating as much as possible the unconditional dynamics while matching the observed configurations, they should also favour configurations close to \mathfrak{m} . Obtaining a quantitative rigorous version of this statement is one of the main objectives of this article and, in view of (1.3), it is equivalent to show that Schrödinger bridges satisfy the turnpike property. In the field of deterministic control, the turnpike phenomenon is rather well understood both in a finite and infinite dimensional setting, see either [49, 48] and references therein, or the monographs [51, 52]. The understanding of this phenomenon in stochastic control seems to be much more limited: see [10, 11, 12] for results on mean field games and [17, 1] for results on the classical and mean field Schrödinger problems. The reason why the turnpike property for Schrödinger bridges in the present context cannot be deduced from existing results lies in the hypocoercive [50] nature of the kinetic Fokker-Planck equation

$$\partial_t f_t(x, v) = \gamma \Delta_v f_t(x, v) - \gamma v \cdot \nabla_v f_t(x, v) + \nabla U \cdot \nabla_v f_t(x, v) - v \cdot \nabla_x f_t(x, v), \quad (1.4)$$

describing the probability density of (1.1) with respect to m . It is well known that the problem of quantifying the trend to equilibrium of this PDE is more challenging than for the classical (overdamped) Fokker-Planck equation and this difficulty is of course reflected in the problem of establishing the turnpike property for the corresponding Schrödinger bridges. In this work, we rely on the important progresses made in the study of the long-time behaviour of (1.4) over the last fifteen years using either an analytical approach see e.g. [4, 23, 31, 50] and references therein, or a probabilistic approach, see e.g. [25, 29], as well as on the new developments around the long-time behaviour of Schrödinger bridges, in order to gain some understanding on *controlled* versions of the kinetic Fokker-Planck equation. Leaving a more accurate comparison between our results and the existing literature to the text below, let us first present a very concise summary of our contributions and explain how this article is structured.

Organisation

The document is organised as follows. In the upcoming sections 1.1, 1.2 and 1.3 we state and comment our main results. In particular, Section 1.1 contains additional background material on the Schrödinger problem and structural results such as existence, uniqueness, duality and existence of Schrödinger-Kantorovich potentials for KSPd. Section 1.2 is devoted to the study of the long-time behaviour of the entropic cost, whereas in Section 1.3 we state exponential turnpike estimates for the Fisher information and relative entropy along Schrödinger bridges. Section 2 contains preliminary results on the Langevin dynamics and the associated semigroup that are needed for the proof of the main results, that we carry out in Section 3 working at first under an extra regularity assumption on the marginal measures μ and ν that we eventually remove thanks to the technical results of Section 4.

1.1 The kinetic Schrödinger problem

This article is devoted to the analysis of a stochastic mass transportation problem, that we name *kinetic Schrödinger problem*, owing to the fact that it is obtained from the classical Schrödinger problem by replacing Brownian particles with a system of independent particles following the Langevin dynamics in Schrödinger's thought experiment. The first formulation KSPd, that we proposed on the basis of Sanov's Theorem, is in terms of an entropy minimisation problem over path probability measures. Besides the change of the reference measure, another difference with respect to classical instances of the Schrödinger problem lies in the fact that it is not the full marginal that is constrained at initial and final time, but only its spatial component. Even though KSPd seems to be a more faithful representation of Schrödinger's thought experiment, also the problem with fully constrained marginals

$$\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) := \inf \left\{ \mathcal{H}(P|R) : P \in \mathcal{P}(\Omega), (X_0, V_0)_{\#}P = \bar{\mu}, (X_T, V_T)_{\#}P = \bar{\nu} \right\}. \quad (\text{KFSPd})$$

where $\bar{\mu}, \bar{\nu} \in \mathcal{P}(\mathbb{R}^{2d})$ is worth studying and we shall work on both problems in the sequel. Through a classical argument [26] it is possible to reduce the dynamic formulations (cf. KSPd and KFSPd) to static ones. For example, KSPd is equivalent to solving

$$\inf \left\{ \mathcal{H}(\pi|R_{0,T}) : \pi \in \Pi_X(\mu, \nu) \right\}, \quad (\text{KSP})$$

where $R_{0,T} := ((X_0, V_0), (X_T, V_T))_{\#}R$ is the joint law of R at initial and terminal time and the set $\Pi_X(\mu, \nu)$ is defined as

$$\Pi_X(\mu, \nu) := \left\{ \pi \in \mathcal{P}(\mathbb{R}^{2d} \times \mathbb{R}^{2d}) \mid (\text{proj}_{x_1})_{\#}\pi = \mu, (\text{proj}_{x_2})_{\#}\pi = \nu \right\},$$

with $\text{proj}_{x_i}((x_1, v_1), (x_2, v_2)) := x_i$ for any $i = 1, 2$. In a similar fashion, the static formulation of KFSPd is

$$\inf \left\{ \mathcal{H}(\pi|_{\mathbb{R}_{0,T}}) : \pi \in \Pi(\bar{\mu}, \bar{\nu}) \right\}, \quad (\text{KFSP})$$

where $\Pi(\bar{\mu}, \bar{\nu})$ is the (usual) set of couplings of $\bar{\mu}$ and $\bar{\nu}$. The equivalence between the static and dynamic formulations is obtained mixing optimal static solutions with the bridges of the reference measure, see [37] for details. Finally, one can also derive a fluid dynamic (Benamou-Brenier [5]) formulation as well a stochastic control formulation of both problems. In particular, the latter one, that we sketched at (1.2) is the one that motivated us to investigate the turnpike phenomenon. We now proceed to establish some basic though fundamental structural results on the kinetic Schrödinger problems at hand. But before doing so, let us present the assumptions under which our main results hold.

1.1.1 Assumptions

We state here the assumptions on the potential U and on the constraints $\mu, \nu, \bar{\mu}$ and $\bar{\nu}$ that we use in the sequel. We define $\mathfrak{m}_X, \mathfrak{m}_V \in \mathcal{P}(\mathbb{R}^d)$ to be the respectively the space and velocity marginals of \mathfrak{m} , in particular $\mathfrak{m} = \mathfrak{m}_X \otimes \mathfrak{m}_V$.

(H1) U is a C^∞ strongly convex potential with bounded derivatives of order $k \geq 2$.

(H2) There exist $0 < \alpha < \beta$ such that

$$\sqrt{\beta} - \sqrt{\alpha} \leq \gamma, \quad \text{and} \quad \alpha \text{Id}_d \leq \nabla^2 U(x) \leq \beta \text{Id}_d, \quad \text{for all } x \in \mathbb{R}^d,$$

where $\gamma > 0$ is the friction parameter in (1.1).

(H3) The probability measures μ and ν on \mathbb{R}^d satisfy

$$\mathcal{H}(\mu|\mathfrak{m}_X) < +\infty \quad \text{and} \quad \mathcal{H}(\nu|\mathfrak{m}_X) < +\infty.$$

(H4) $\mu, \nu \ll \mathfrak{m}_X, \frac{d\mu}{d\mathfrak{m}_X}, \frac{d\nu}{d\mathfrak{m}_X} \in L^\infty(\mathfrak{m}_X)$ and are compactly supported on \mathbb{R}^d .

(FH3) The probability measures $\bar{\mu}$ and $\bar{\nu}$ on \mathbb{R}^{2d} satisfy

$$\mathcal{H}(\bar{\mu}|\mathfrak{m}) < +\infty \quad \text{and} \quad \mathcal{H}(\bar{\nu}|\mathfrak{m}) < +\infty.$$

(FH4) $\bar{\mu}, \bar{\nu} \ll \mathfrak{m}, \frac{d\bar{\mu}}{d\mathfrak{m}}, \frac{d\bar{\nu}}{d\mathfrak{m}} \in L^\infty(\mathfrak{m})$ and are compactly supported on \mathbb{R}^{2d} .

Assumption (H2) implies *local* gradient contraction bounds for the semigroup generated by the Langevin dynamics with a certain rate $\kappa > 0$ (see Proposition 2.2 or [4]). The exponential rate κ of Theorems 1.6 and 1.7 below is precisely the one, computed e.g. in [39, 6], at which the synchronous coupling is contractive for the (uncontrolled) Langevin dynamics.

For each of the main results, we will make it explicit which assumptions from the above list are needed.

1.1.2 Duality

We begin with a duality result, analogous to the Monge-Kantorovich duality of optimal transport and the more recent dual representations of the entropic cost for the classical Schrödinger problem [27]. It is worth noticing that, since the stationary Langevin dynamics is not a reversible measure, $\mathcal{C}_T^F(\cdot, \cdot)$ is not symmetric in its arguments. Nevertheless, due to the “physical reversibility” of the dynamics [15], that is, reversibility up to a sign flip in the velocities, it is not hard to show that $\mathcal{C}_T(\cdot, \cdot)$ is symmetric in its arguments.

Proposition 1.1. *Grant (H1) and (H3). Then $\mathcal{C}_T(\mu, \nu) < \infty$ and*

$$\mathcal{C}_T(\mu, \nu) = \sup_{\varphi, \psi \in C_b(\mathbb{R}^d)} \left\{ \int_{\mathbb{R}^d} \varphi \, d\mu + \int_{\mathbb{R}^d} \psi \, d\nu - \log \int_{\mathbb{R}^{4d}} e^{\varphi \oplus \psi} \, dR_{0,T} \right\}. \quad (1.5)$$

Similarly, grant (H1) and (FH3) it holds $\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) < \infty$ and

$$\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) = \sup_{\varphi, \psi \in C_b(\mathbb{R}^{2d})} \left\{ \int_{\mathbb{R}^{2d}} \varphi \, d\bar{\mu} + \int_{\mathbb{R}^{2d}} \psi \, d\bar{\nu} - \log \int_{\mathbb{R}^{4d}} e^{\varphi \oplus \psi} \, dR_{0,T} \right\}. \quad (1.6)$$

1.1.3 The fg-decomposition

Optimal couplings in the Schrödinger problem are characterised by the fact that their density against the reference measure takes a product form, often called fg -decomposition [37]. In KSPd f and g have the additional property of depending only on the first and second space variables respectively.

Proposition 1.2. *Grant (H1), (H3). Then, for all $T > 0$, KSP and KSPd admit unique solutions μ^T, P^T with $\mu^T = ((X_0, V_0), (X_T, V_T))_{\#} P^T$ and there exist two non-negative measurable functions f^T, g^T on \mathbb{R}^d such that*

$$\rho^T(x, v, y, w) := \frac{d\mu^T}{dR_{0,T}}(x, v, y, w) = f^T(x)g^T(y), \quad R_{0,T}\text{-a.s.} \quad (1.7)$$

Moreover, f^T, g^T solve the Schrödinger system:

$$\begin{cases} \frac{d\mu}{d\mathfrak{m}_X}(x) = f^T(x) \mathbb{E}_R[g^T(X_T)|X_0 = x], \\ \frac{d\nu}{d\mathfrak{m}_X}(y) = g^T(y) \mathbb{E}_R[f^T(X_0)|X_T = y]. \end{cases} \quad (1.8)$$

For KFSP and KFSPd, the uniqueness of solutions (hereafter $\bar{\mu}^T$ and \bar{P}^T respectively) and the fg -decomposition are a direct consequence of known results, see e.g. [42], whereas the case KSP requires some more work. We remark here that for both dual representation of the cost and the fg -decomposition the strict convexity of U and its smoothness are not really necessary, a bounded Hessian would suffice.

1.2 Long-time behaviour of the entropic costs

Let us now turn the attention to the ergodic properties of KSP and KFSP by investigating the long-time behaviour of the entropic cost. To explain the upcoming results, we remark that (H1) implies ergodicity of the Langevin dynamics [20, Theorem 11.14] and in particular one has the weak convergence

$$R_{0,T} \rightharpoonup \mathfrak{m} \otimes \mathfrak{m}.$$

Intuitively, this implies that the variational problem KSP converges, in a sense to be made precise, to the problem

$$\min_{\pi \in \Pi_X(\mu, \nu)} \mathcal{H}(\pi | \mathfrak{m} \otimes \mathfrak{m}), \quad (1.9)$$

whose optimal solution and optimal value are easily seen to be $(\mu \otimes \mathfrak{m}_V) \otimes (\nu \otimes \mathfrak{m}_V)$ and $\mathcal{H}(\mu | \mathfrak{m}_X) + \mathcal{H}(\nu | \mathfrak{m}_X)$ respectively. From the point of view of the particle system, this means that in the long-time limit, initial and final states of the system become essentially independent of one another. Moreover, the initial and final velocities are well approximated by independent Gaussians, and are independent from the spatial variables. The result below turns this intuition into a solid argument, including a quantitative version of the convergence of the entropic cost towards the sum of the marginal entropies. For the classical Schrödinger problem, an analogous statement can be found in [19].

Theorem 1.3. Grant (H1) and (H3). Then

$$\lim_{T \rightarrow \infty} \mathcal{C}_T(\mu, \nu) = \mathcal{H}(\mu | \mathbf{m}_X) + \mathcal{H}(\nu | \mathbf{m}_X) < \infty. \quad (1.10)$$

Moreover as $T \rightarrow \infty$

$$\mu^T \rightharpoonup (\mu \otimes \mathbf{m}_V) \otimes (\nu \otimes \mathbf{m}_V) \in \Pi_X(\mu, \nu), \quad (1.11)$$

weakly and, granted (H2), there exists a positive constant $C_{d,\alpha,\beta,\gamma}$ (depending only on d, α, β and γ) such that for any $0 < \delta \leq 1$, as soon as $T > (\frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta) \vee \frac{1}{\kappa} \log \frac{C_{d,\alpha,\beta,\gamma}}{\delta^3}$, it holds

$$|\mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu | \mathbf{m}_X) - \mathcal{H}(\nu | \mathbf{m}_X)| \leq C_{d,\alpha,\beta,\gamma} \delta^{-3} e^{-\kappa T} \left[\mathcal{H}(\mu | \mathbf{m}_X) + \mathcal{H}(\nu | \mathbf{m}_X) \right], \quad (1.12)$$

and as a consequence the following entropic Talagrand inequality holds

$$\mathcal{C}_T(\mu, \nu) \leq \left(1 + C_{d,\alpha,\beta,\gamma} \delta^{-3} e^{-\kappa T} \right) \left[\mathcal{H}(\mu | \mathbf{m}_X) + \mathcal{H}(\nu | \mathbf{m}_X) \right]. \quad (1.13)$$

Remark 1.4. Equation (1.11) implies in particular that $\mu_0^T \rightharpoonup \mu \otimes \mathbf{m}_V$ and $\nu_0^T \rightharpoonup \nu \otimes \mathbf{m}_V$. This convergence is also exponential, as we show in Theorem 3.8.

Theorem 1.5. Under the (H1) and (FH3) it holds

$$\lim_{T \rightarrow \infty} \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) = \mathcal{H}(\bar{\mu} | \mathbf{m}) + \mathcal{H}(\bar{\nu} | \mathbf{m}) < \infty. \quad (1.14)$$

Moreover as $T \rightarrow \infty$

$$\bar{\mu}^T \rightharpoonup \bar{\mu} \otimes \bar{\nu} \in \Pi(\bar{\mu}, \bar{\nu}), \quad (1.15)$$

weakly and, granted (H2), there exists a positive constant $C_{d,\alpha,\beta,\gamma}$ such that for any $0 < \delta \leq 1$, as soon as $T > (\frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta) \vee \frac{1}{\kappa} \log \frac{C_{d,\alpha,\beta,\gamma}}{\delta^3}$, it holds

$$|\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) - \mathcal{H}(\bar{\mu} | \mathbf{m}) - \mathcal{H}(\bar{\nu} | \mathbf{m})| \leq C_{d,\alpha,\beta,\gamma} \delta^{-3} e^{-\kappa T} \left[\mathcal{H}(\bar{\mu} | \mathbf{m}) + \mathcal{H}(\bar{\nu} | \mathbf{m}) \right], \quad (1.16)$$

and as a consequence the following entropic Talagrand inequality holds

$$\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) \leq \left(1 + C_{d,\alpha,\beta,\gamma} \delta^{-3} e^{-\kappa T} \right) \left[\mathcal{H}(\bar{\mu} | \mathbf{m}) + \mathcal{H}(\bar{\nu} | \mathbf{m}) \right]. \quad (1.17)$$

The proof of the qualitative statements in the above results rely on Γ -convergence and some simple consequences of the heat kernel estimates in [21]. The key ingredient in the proof of the exponential estimates is a representation formula for the difference

$$\mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu | \mathbf{m}_X) - \mathcal{H}(\nu | \mathbf{m}_X)$$

that we establish at Lemma 3.1 and allows to profit from the turnpike estimates at Theorem 1.6 and 1.7 below.

1.3 Long-time behaviour of Schrödinger bridges

One of the main contributions of this article are the upcoming quantitative results on the long-time behaviour of Schrödinger bridges, which imply in particular exponential convergence to \mathbf{m} when looking at timescales of order T and exponential convergence in T to the Langevin dynamics when looking at the Schrödinger bridge over a fixed time-window $[0, t]$.

1.3.1 Entropic turnpike property

We propose two turnpike results in which distance from equilibrium is measured through the relative entropy $\mathcal{H}(\cdot|\mathfrak{m})$ and the Fisher information $\mathcal{I}(\cdot)$, see (1.18) below. The use of $\mathcal{H}(\cdot|\mathfrak{m})$ is natural in light of the fact that the costs $\mathcal{C}_T(\mu, \nu)$ and $\mathcal{C}_T^F(\bar{\mu}, \bar{\nu})$ are also relative entropies, but computed on different spaces. On the other hand, the bound on $\mathcal{I}(\cdot)$ is reminiscent of the celebrated Bakry-Émery estimates [2]. It is worth noticing that entropic turnpike estimates seem to be very rare in the existing literature and even less so are bounds on the Fisher information: we shall elaborate more on this at Remark 1.9. The key assumption for obtaining (1.20) and (1.23) is (H2), asking U to be strongly convex and such that the difference between the smallest and largest eigenvalues of $\nabla^2 U(x)$ is controlled by the friction parameter γ uniformly in x . This assumption is often encountered in works dealing with the long-time behaviour of the (uncontrolled) kinetic Fokker-Planck equation, see e.g. [6]. Although exponential L^2 estimates are known to hold under considerably weaker assumptions (see e.g. [31] and [9, 32] for singular potentials), and entropic estimates assuming a bounded and positive Hessian have been known for more than a decade [50], it is only recently [29] that entropic estimates have been obtained beyond the bounded Hessian case. In light of this, the question of how to improve our results is quite interesting and deserves to be further investigated. Let us state the announced results, beginning with KSPd. To do so, we need another bit of notation: if P^T is the unique solution of KSPd, we call *entropic interpolation* $(\mu_t^T)_{t \in [0, T]}$ the marginal flow of P^T and denote ρ_t^T its density against \mathfrak{m} , i.e.

$$\forall t \in [0, T], \quad \mu_t^T = (X_t, V_t)_{\#} P^T, \quad \rho_t^T := \frac{d\mu_t^T}{d\mathfrak{m}}.$$

With the obvious small modifications, we also define the entropic interpolation $(\bar{\mu}_t^T)_{t \in [0, T]}$ and their densities $(\bar{\rho}_t^T)_{t \in [0, T]}$ in the framework of KFSPd. Furthermore, we introduce the functional \mathcal{I} to be the Fisher information with respect to \mathfrak{m} , defined for any $q \ll \mathfrak{m} \in \mathcal{P}(\mathbb{R}^{2d})$ as

$$\mathcal{I}(q) := \begin{cases} \int_{\mathbb{R}^{2d}} \left| \nabla \log \frac{dq}{d\mathfrak{m}} \right|^2 dq & \text{if } \nabla \log \frac{dq}{d\mathfrak{m}} \in L^2(q), \\ +\infty, & \text{otherwise.} \end{cases} \quad (1.18)$$

Theorem 1.6 (Entropic turnpike for KSP). *Grant (H1), (H2) and (H3). There exists a positive constant $C_{d, \alpha, \beta, \gamma}$ such that for any $0 < \delta \leq 1$ and $t \in [\delta, T - \delta]$, as soon as $T > \frac{1}{\kappa} \log C_{d, \alpha, \beta, \gamma} + 2\delta$, it holds*

$$\mathcal{I}(\mu_t^T) \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-2\kappa[t \wedge (T-t)]} \mathcal{C}_T(\mu, \nu), \quad (1.19)$$

$$\mathcal{H}(\mu_t^T | \mathfrak{m}) \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-2\kappa[t \wedge (T-t)]} \mathcal{C}_T(\mu, \nu). \quad (1.20)$$

Moreover, as soon as $T > (\frac{1}{\kappa} \log C_{d, \alpha, \beta, \gamma} + 2\delta) \vee \frac{1}{\kappa} \log \frac{C_{d, \alpha, \beta, \gamma}}{\delta^3}$, we have

$$\mathcal{H}(\mu_t^T | \mathfrak{m}) \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-2\kappa[t \wedge (T-t)]} \left[\mathcal{H}(\mu | \mathfrak{m}_X) + \mathcal{H}(\nu | \mathfrak{m}_X) \right]. \quad (1.21)$$

Theorem 1.7 (Entropic turnpike for KFSP). *Grant (H1), (H2) and (FH3). There exists a positive constant $C_{d, \alpha, \beta, \gamma}$ such that for any $0 < \delta \leq 1$ and $t \in [\delta, T - \delta]$, as soon as $T > \frac{1}{\kappa} \log C_{d, \alpha, \beta, \gamma} + 2\delta$, it holds*

$$\mathcal{I}(\bar{\mu}_t^T) \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-2\kappa[t \wedge (T-t)]} \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}), \quad (1.22)$$

$$\mathcal{H}(\bar{\mu}_t^T | \mathfrak{m}) \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-2\kappa[t \wedge (T-t)]} \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}). \quad (1.23)$$

Moreover, as soon as $T > (\frac{1}{\kappa} \log C_{d, \alpha, \beta, \gamma} + 2\delta) \vee \frac{1}{\kappa} \log \frac{C_{d, \alpha, \beta, \gamma}}{\delta^3}$, we have

$$\mathcal{H}(\bar{\mu}_t^T | \mathfrak{m}) \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-2\kappa[t \wedge (T-t)]} \left[\mathcal{H}(\bar{\mu} | \mathfrak{m}) + \mathcal{H}(\bar{\nu} | \mathfrak{m}) \right]. \quad (1.24)$$

Remark 1.8. If we compare our results with what is known in deterministic control we remark that, quite curiously, exponential estimates for the deterministic noiseless version of (1.2), obtained removing the Brownian motion from the controlled state equation, do not seem to be covered from existing results, even in the case when μ and ν are Dirac measures.¹ For linear-quadratic problems though, the result is well known, see e.g. [8] for precise estimates. Theorems 1.6 and 1.7 provide *global* turnpike estimates, that is to say we do not ask μ and ν to be close to m . We do ask $\mathcal{H}(\mu|m), \mathcal{H}(\nu|m) < +\infty$, but this condition is very mild and necessary for the Schrödinger problem to have a finite value. This is in contrast with most exponential turnpike estimates we are aware of in deterministic control (see e.g. [49, Theorem 1]). The passage from local to global estimates seems to be possible [47, 46] under some extra assumptions, such as the existence of a storage function, but this comes at the price of losing quite some information on the multiplicative constants appearing in (1.20). Moreover, the condition $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$ of Theorem 1.6 should be replaced with a condition of the form $T > T_0$ with T_0 depending on the initial conditions and potentially very large.

Remark 1.9. The bound on the Fisher information is our strongest result as it implies immediately an entropic bound thanks to the logarithmic Sobolev inequality (2.4). Moreover, entropic bounds are stronger than bounds expressed by means of a transport distance such as W_1 or W_2 , since m satisfies Talagrand's inequality (2.2).

Remark 1.10. Proving the turnpike property for Schrödinger bridges in this context is harder than in the classical setting, and we need to work under stronger assumptions on the potential U than its strong convexity. This is not a surprise. Indeed, proving the exponential convergence to equilibrium for the kinetic Fokker-Planck equation is a difficult problem that has been, and still is, intensively studied by means of either a probabilistic or an analytic approach, see [13, 25, 45, 29] for some references on the probabilistic approach. Following the terminology introduced by Villani in his monograph [50], this obstruction is a manifestation of the *hypocoercive* nature of the kinetic Fokker-Planck equation. KSP may indeed be regarded as the prototype of an hypocoercive stochastic control problem. For the moment, we have been able to show the turnpike property under a quasilinearity assumption. Assumptions of this type, where the friction parameter has to be in some sense large in comparison with the spectrum of $\nabla^2 U$ are commonly encountered in the literature. In the language of probability, they ensure that the synchronous coupling is contracting for the Langevin dynamics [6, 39]. On the other hand, from an analytical standpoint, Assumption (H2) implies *local* gradient bounds for the semigroup generated by the Langevin dynamics [4]. Finally, we recall that the exponential rate κ of Theorems 1.6 and 1.7 is precisely the one, computed e.g. in [39, 6], at which synchronous coupling is contractive for the (uncontrolled) Langevin dynamics.

Proof strategy A general idea to obtain exponential speed of convergence to equilibrium for hypocoercive equations systematically exploited in [50] is that of modifying the “natural” Lyapunov function of the system by adding some extra terms in such a way that proving exponential dissipation becomes an easier task. For the Langevin dynamics, a suitable modification of the natural Lyapunov functional, that is the relative entropy $\mathcal{H}(\cdot|m)$, is obtained considering

$$\mu \mapsto a\mathcal{H}(\mu|m) + \mathcal{I}(\mu)$$

for a carefully chosen constant $a > 0$. Emulating Bakry-Émery Γ -calculus [4] it is possible to show that the modified Lyapunov functional decays exponentially along solutions

¹For example, if we compare with the reference work [49], the matrix W defined at Eq. (10) therein would not be invertible for the problem under consideration, which thus fails to satisfy the hypothesis of the main turnpike result obtained there.

of the kinetic Fokker-Planck equation. Our proof of the turnpike property consists in implementing this abstract idea on the fg -decomposition of the entropic interpolation, as we now briefly explain. Indeed, in order to bound $\mathcal{I}(\mu_t^T)$ one is naturally led to consider the quantities

$$\int_{\mathbb{R}^{2d}} |\nabla \log f_s^T|^2 f_s^T g_s^T \, \mathrm{d}\mathbf{m}, \tag{1.25a}$$

$$\int_{\mathbb{R}^{2d}} |\nabla \log g_s^T|^2 f_s^T g_s^T \, \mathrm{d}\mathbf{m}. \tag{1.25b}$$

However, it is not clear how to obtain a differential inequality ensuring exponential (forward) dissipation of (1.25a) and exponential (backward) dissipation of (1.25b). But, as we show at Lemma 3.2, it is possible to find two norms $|\cdot|_{M^{-1}}$ and $|\cdot|_{N^{-1}}$, that are equivalent to the Euclidean norm and such that if we define

$$\varphi^T(s) := \int_{\mathbb{R}^{2d}} |\nabla \log f_s^T|_{N^{-1}}^2 f_s^T g_s^T \, \mathrm{d}\mathbf{m} \quad \text{and} \quad \psi^T(s) := \int_{\mathbb{R}^{2d}} |\nabla \log g_s^T|_{M^{-1}}^2 f_s^T g_s^T \, \mathrm{d}\mathbf{m}, \tag{1.26}$$

then $\varphi^T(s)$ and $\psi^T(s)$ satisfy the desired exponential estimates. To complete the proof, one needs to take care of the boundary conditions. This part is non trivial as it demands to prove certain regularity properties of the fg -decomposition and it is accomplished in two steps: we first show at Proposition 3.4 a regularising property of entropic interpolations, namely that if $\mathcal{H}(\mu|\mathfrak{m}_X), \mathcal{H}(\nu|\mathfrak{m}_X)$ are finite, then the Fisher information $\mathcal{I}(\mu_t^T)$ is finite for any $t \in (0, T)$. The proof of this property is based on a gradient bound obtained in [30] and is of independent interest. The second step (Proposition 3.3) consists in showing that for a fixed small δ , $\varphi^T(\delta)$ and $\psi^T(T - \delta)$ can be controlled with by the sum of $\mathcal{I}(\mu_\delta^T)$ and $\mathcal{I}(\mu_{T-\delta}^T)$. We prove this estimate adapting an argument used in [49] in the analysis of deterministic finite dimensional control problems.

1.3.2 Convergence to the Langevin dynamics over a fixed time-window

We are able to precisely analyse the behaviour of entropic interpolations for a fixed time t , while T grows large. More precisely, we show that the (uncontrolled) Langevin dynamics and the Schrödinger bridge are exponentially close in the long-time regime $T \rightarrow \infty$, for all time-windows $[0, t]$. Note that this result cannot be deduced from the turnpike estimates of the former section.

Theorem 1.11. *Under hypotheses (H1), (H2) and (H3), there exists a positive constant $C_{d,\alpha,\beta,\gamma}$ such that for any $0 < \delta \leq 1$ and $t \in [0, T - \delta]$, as soon as $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$, it holds*

$$\mathcal{W}_2(\mu_t^T, \mu_t^\infty) \leq C_{d,\alpha,\beta,\gamma} \delta^{-\frac{3}{2}} e^{-\kappa(T-t)} \sqrt{\mathcal{C}_T(\mu, \nu)},$$

μ_t^∞ is the law of (X_t, V_t) satisfying

$$\begin{cases} \mathrm{d}X_t = V_t \mathrm{d}t, \\ \mathrm{d}V_t = -\nabla U(X_t) \mathrm{d}t - \gamma V_t \mathrm{d}t + \sqrt{2\gamma} \mathrm{d}B_t, \\ (X_0, V_0) \sim \mu \otimes \mathfrak{m}_V. \end{cases} \tag{1.27}$$

A similar statement holds true for KFSP replacing Assumption (H3) with (FH3) and with initial condition in (1.27) given by $\bar{\mu}$.

2 Preliminaries

In this section we collect useful results about the Markov semigroup associated to the kinetic Fokker-Planck equation. In what follows we write \lesssim to indicate that an inequality holds up to a multiplicative positive constant depending possibly on the dimension d , the bounds on the spectrum of $\nabla^2 U$, α and β , or the friction parameter γ .

2.1 On the assumptions

In this short section we report some straightforward consequences of the various assumptions listed at Section 1.1.1 that we shall repeatedly use from now on. We begin by observing that assumption (H1) guarantees that $m \in \mathcal{P}_2(\mathbb{R}^{2d})$ and that m_X satisfies Talagrand’s inequality because of [3, Corollary 9.3.2], i.e. for any $q \in \mathcal{P}(\mathbb{R}^d)$

$$\mathcal{W}_2(q, m_X)^2 \lesssim \mathcal{H}(q|m_X). \tag{2.1}$$

Since the Talagrand inequality holds also for the Gaussian measure m_V , from [3, Proposition 9.2.4] it follows that for any $q \in \mathcal{P}(\mathbb{R}^{2d})$

$$\mathcal{W}_2(q, m)^2 \lesssim \mathcal{H}(q|m). \tag{2.2}$$

Let us also point out that (H4) implies (H3) and that under (H1) and (H3) it easily follows that $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$. Indeed,

$$\int_{\mathbb{R}^d} |x|^2 d\mu \lesssim \int_{\mathbb{R}^d} |x|^2 dm_X + \mathcal{W}_2(\mu, m_X)^2 \stackrel{(2.1)}{\lesssim} \int_{\mathbb{R}^d} |x|^2 dm_X + \mathcal{H}(\mu|m_X) < +\infty, \tag{2.3}$$

and similarly for the measure ν . We also remark that (FH4) implies (FH3). Moreover, from (H1) and (FH3), by means of (2.2), it follows that $\bar{\mu}, \bar{\nu} \in \mathcal{P}_2(\mathbb{R}^{2d})$.

Finally, let us also notice that (H1) and (H2) guarantee the validity of a log-Sobolev inequality for m_X because of [3, Corollary 5.7.2], and by means of [3, Proposition 5.2.7 and Proposition 5.5.1] it follows that m satisfies a log-Sobolev inequality. Therefore for any $q \ll m$ it holds

$$\mathcal{H}(q|m) \lesssim \mathcal{I}(q). \tag{2.4}$$

2.2 Markov semigroups and heat kernel

The generator L associated to the SDE (1.1) is given by

$$L = \gamma \Delta_v - \gamma v \cdot \nabla_v - \nabla U \cdot \nabla_v + v \cdot \nabla_x$$

while its adjoint in $L^2(m)$ reads as

$$L^* = \gamma \Delta_v - \gamma v \cdot \nabla_v + \nabla U \cdot \nabla_v - v \cdot \nabla_x.$$

Under assumption (H1), it is well known that Hörmander’s Theorem for parabolic hypoellipticity applies [33, Theorem 1.1] to the operator L , and thus the associated semigroup $(P_t)_{t \geq 0}$ admits a probability kernel $p_t((x, y), (y, w))$, which is C^∞ in all of the parameters, with respect to the invariant probability measure

$$dm(x, v) = \frac{1}{Z} e^{-U(x) - \frac{|v|^2}{2}} dx dv,$$

where Z is a normalising constant. Sometimes, with a slight abuse of notation we will write $m(x, v)$ to denote the density of m with respect to the Lebesgue measure. Similarly, we will denote by $(P_t^*)_{t \geq 0}$ the semigroup associated to L^* . Note that the function p_t also represents the density of $R_{0,t}$ (the joint law at time 0 and t of the solution to (1.1)) with respect to $dm \otimes m$. Moreover, according to [21, Theorem 1.1], $p_t(\cdot, \cdot)$ satisfies two-sided Gaussian estimates. Importantly, p_t is locally bounded away from zero and infinity, but with constants that might depend non-trivially on the time horizon T .

For some of our proofs, we need lower bounds that are uniform in T . To this aim, we have the following consequence of the results of [21], whose proof is postponed to the appendix.

Lemma 2.1. *Let $T_0 > 0$ be fixed. Under assumption (H1), there exists a constant $c_{T_0} > 0$ such that for all $T \geq T_0$ and all $(x, v), (y, w) \in \mathbb{R}^{2d}$*

$$\log p_T((x, v), (y, w)) \geq -c_{T_0} \left(1 + |x|^2 + |v|^2 + |y|^2 + |w|^2\right). \tag{2.5}$$

We stress that the Langevin dynamics (1.1) are not reversible, and in particular the probability kernel p_t is not symmetric. However, it is symmetric up to a sign-flip in the velocities,

$$p_t((x, v), (y, w)) = p_t((y, -w), (x, -v)) \quad \forall t \geq 0, \quad \forall (x, v), (y, w) \in \mathbb{R}^{2d}. \tag{2.6}$$

As we said above, this useful property is sometimes called physical reversibility.

2.3 Contraction of the semigroup

In our setup, due to the lack of a curvature condition, the standard Bakry-Emery machinery does not apply to obtain a commutation estimate for the semigroup of the type

$$|\nabla P_t h(z)| \leq e^{-ct} P_t(|\nabla h|)(z), \tag{2.7}$$

for some $c > 0$. It is still possible to obtain a commutation estimate similar to (2.7) by replacing the Euclidean norm $|\cdot|$ by a certain twisted norm $|\xi|_M := \sqrt{\xi \cdot M \xi}$ on \mathbb{R}^{2d} for some well chosen positive definite symmetric matrix $M \in \mathbb{R}^{2d \times 2d}$. This is a common idea in the kinetic setting and it is exploited for example in [4, 29, 39].

For instance, in Theorem 1 of [39] the author studies the contraction properties of the semigroup P_t associated to the SDE on \mathbb{R}^m

$$dZ_t = b(Z_t)dt + \Sigma dB_t, \tag{2.8}$$

with the drift $b : \mathbb{R}^m \rightarrow \mathbb{R}^m$ being globally Lipschitz and Σ a constant positive-semidefinite symmetric matrix. The author shows that the condition on the Jacobian matrix J_b of the drift

$$\xi \cdot (M J_b(z)) \xi \leq -\kappa \xi \cdot M \xi = -\kappa |\xi|_M^2 \quad \forall \xi \in \mathbb{R}^m, \forall z \in \mathbb{R}^m, \tag{2.9}$$

where $\kappa \in \mathbb{R}$ and M is a positive definite symmetric matrix, is equivalent to the commutation estimate

$$|\nabla P_t h(z)|_{M^{-1}} \leq e^{-\kappa t} P_t(|\nabla h|_{M^{-1}})(z).$$

Our setup, which is also discussed in [39, Section 3.3], corresponds to the choice $m = 2d$, and

$$b(x, v) = \begin{pmatrix} v \\ -\nabla U(x) - \gamma v \end{pmatrix} \quad \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2\gamma} \text{Id} \end{pmatrix}$$

and therefore the Jacobian reads as

$$J_b(x, v) = \begin{pmatrix} 0 & \text{Id} \\ -\nabla^2 U(x) & -\gamma \text{Id} \end{pmatrix}.$$

In [39, Proposition 5], the author shows that (2.9) holds with $\kappa > 0$ as long as α and β from assumption (H2) are close enough. By exploiting the symmetry of the heat kernel up to a sign flip, we obtain a similar commutation estimate also for the reversed dynamics.

In view of the above discussion we have the following.

Proposition 2.2. *Assume that (H1) and (H2) hold. Then, there exist a constant $\kappa > 0$ and positive definite symmetric matrices $M, N \in \mathbb{R}^{2d \times 2d}$ such that (2.9) holds and*

(i) *For all $h \in C_c^1(\mathbb{R}^{2d})$, $t \geq 0$ and $z \in \mathbb{R}^{2d}$*

$$|\nabla P_t h(z)|_{M^{-1}} \leq e^{-\kappa t} P_t(|\nabla h|_{M^{-1}})(z). \tag{2.10}$$

(ii) For all $h \in C_c^1(\mathbb{R}^{2d})$, $t \geq 0$ and $z \in \mathbb{R}^{2d}$

$$|\nabla P_t^* h(z)|_{N^{-1}} \leq e^{-\kappa t} P_t^*(|\nabla h|_{N^{-1}})(z). \tag{2.11}$$

Proof. A proof of (2.9) with $\kappa > 0$ under (H1) and (H2) can be obtained by mimicking the computations in Theorem 2.12 of [4], where the case $\gamma = 1$ is discussed. Given (2.9), (i) follows from Theorem 1 in [39].

We now derive (ii) from (i) with the help of (2.6). For any function f on \mathbb{R}^{2d} define the transformation $\mathcal{S}f(x, v) = f(x, -v)$ and set

$$N = \begin{pmatrix} \text{Id} & 0 \\ 0 & -\text{Id} \end{pmatrix} M \begin{pmatrix} \text{Id} & 0 \\ 0 & -\text{Id} \end{pmatrix}.$$

Note that $\mathcal{S}^2 = \text{Id}$, moreover in view of (2.6), for all $h \in C_c^1(\mathbb{R}^{2d})$, $P_t^*(\mathcal{S}h) = \mathcal{S}(P_t h)$ and $\mathcal{S}|\nabla h|_{M^{-1}} = |\nabla(\mathcal{S}h)|_{N^{-1}}$. It is then immediate to derive

$$|\nabla P_t^* h|_{N^{-1}} = \mathcal{S}|\nabla P_t(\mathcal{S}h)|_{M^{-1}} \leq e^{-\kappa t} \mathcal{S}\left(P_t(|\nabla(\mathcal{S}h)|_{M^{-1}})\right) = e^{-\kappa t} P_t^*(|\nabla h|_{N^{-1}}),$$

which is the desired conclusion. □

As a result of Proposition 2.2 and Theorem 1 in [39] we have the equivalent statements, with M, N and $\kappa > 0$ as above, and all $q_1, q_2 \in \mathcal{P}(\mathbb{R}^{2d})$,

$$\mathcal{W}_{M,2}(q_1 P_t, q_2 P_t) \leq e^{-\kappa t} \mathcal{W}_{M,2}(q_1, q_2), \tag{2.12}$$

$$\mathcal{W}_{N,2}(q_1 P_t^*, q_2 P_t^*) \leq e^{-\kappa t} \mathcal{W}_{N,2}(q_1, q_2), \tag{2.13}$$

where $\mathcal{W}_{M,2}(q_1, q_2)$ is the \mathcal{W}_2 -Wasserstein distance on $\mathcal{P}(\mathbb{R}^{2d})$ with the Euclidean metric replaced by $d_M(x, y) = |x - y|_M$ and similarly for $\mathcal{W}_{N,2}(q_1, q_2)$.

3 Proof of the main results

3.1 Duality and fg -decomposition for KSP

Proof of Proposition 1.2. We only sketch the proof as it is rather standard. We consider the measure $\mathbb{R}_{0,T}^X := (\text{proj}_{x_1}, \text{proj}_{x_2})_{\#} \mathbb{R}_{0,T} = (X_0, X_T)_{\#} \mathbb{R}$ and the minimisation problem,

$$\min_{q \in \Pi(\mu, \nu)} \mathcal{H}(q | \mathbb{R}_{0,T}^X), \tag{3.1}$$

where $\Pi(\mu, \nu)$ is the set of couplings of $\mu, \nu \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$. In view of the heat kernel lower bound in Lemma 2.1, we know that for some $C > 0$ and uniformly in x, y it holds

$$\frac{d\mathbb{R}_{0,T}^X}{d(\mathfrak{m}_X \otimes \mathfrak{m}_X)}(x, y) \geq \frac{1}{C} e^{-C(1+|x|^2+|y|^2)}.$$

which in combination with (H3) implies that $\mathcal{H}(\mu \otimes \nu | \mathbb{R}_{0,T}^X) < \infty$. Indeed the bound above implies that for any $T > T_0$, T_0 fixed, there is a constant $C_{d,\alpha,\beta,\gamma,T_0} > 0$ such that

$$\begin{aligned} \mathcal{H}(\mu \otimes \nu | \mathbb{R}_{0,T}^X) &= \mathcal{H}(\mu \otimes \nu | \mathfrak{m}_X \otimes \mathfrak{m}_X) - \int_{\mathbb{R}^{4d}} \log \frac{d\mathbb{R}_{0,T}^X}{d(\mathfrak{m}_X \otimes \mathfrak{m}_X)} d\mu \otimes \nu \\ &\stackrel{(2.3)}{\leq} C_{d,\alpha,\beta,\gamma,T_0} [1 + \mathcal{H}(\mu | \mathfrak{m}_X) + \mathcal{H}(\nu | \mathfrak{m}_X)]. \end{aligned} \tag{3.2}$$

Thus, Proposition 2.5 in [37] applies and the above minimisation problem has indeed a unique solution $\pi \in \Pi(\mu, \nu)$. By applying [28, Proposition 2.1], there exist two non-negative measurable functions f^T, g^T on \mathbb{R}^d such that

$$\frac{d\pi}{d\mathbb{R}_{0,T}^X}(x, y) = f^T(x)g^T(y), \quad \mathbb{R}_{0,T}^X\text{-a.s.}, \tag{3.3}$$

from which (1.8) directly follows. Now, in view of the additive property of the relative entropy, we get for any $P \in \mathcal{P}(\Omega)$

$$\mathcal{H}(P|R) = \mathcal{H}(P_{0,T}^X|R_{0,T}^X) + \int_{\mathbb{R}^{2d}} \mathcal{H}(P^{x,y}|R^{x,y}) dP_{0,T}^X(x, y),$$

with $R^{x,y} = R(\cdot | X_0 = x, X_T = y)$ and similarly for $P^{x,y}$. Therefore, a minimizer to KSPd can be found by defining

$$P^T(\cdot) = \int_{\mathbb{R}^d \times \mathbb{R}^d} R(\cdot | X_0 = x, X_T = y) d\pi(x, y),$$

which satisfies $(X_0, X_T)_{\#} P^T = \pi$ and $\mathcal{C}_T(\mu, \nu) = \mathcal{H}(P^T|R) = \mathcal{H}(\pi|R_{0,T}^X) < \infty$. In particular, in view of (3.2), for all $T > T_0$, T_0 fixed, there is $C_{d,\alpha,\beta,\gamma,T_0} > 0$ such that

$$\mathcal{C}_T(\mu, \nu) \leq C_{d,\alpha,\beta,\gamma,T_0} [1 + \mathcal{H}(\mu|m_X) + \mathcal{H}(\nu|m_X)]. \tag{3.4}$$

Similarly, for any $q \in \Pi_X(\mu, \nu)$, denoting $q^X = (\text{proj}_{x_1}, \text{proj}_{x_2})_{\#} q$, we have $\mathcal{H}(q|R_{0,T}) \geq \mathcal{H}(q^X|R_{0,T}^X) \geq \mathcal{H}(\pi|R_{0,T}^X)$ with equality if and only if $q = \mu^T$ where

$$\mu^T(\cdot) = \int_{\mathbb{R}^d \times \mathbb{R}^d} R_{0,T}(\cdot | X_0 = x, X_T = y) d\pi(x, y). \tag{3.5}$$

Note that $\mu^T = ((X_0, V_0), (X_T, V_T))_{\#} P^T$ and $\mathcal{H}(P^T|R) = \mathcal{H}(\mu^T|R_{0,T}) = \mathcal{H}(\pi|R_{0,T}^X) < \infty$. The solutions are unique by strict convexity of the entropy and the linearity of the constraint. Equation (3.5) implies equality of the conditional distributions of μ^T and $R_{0,T}$ given the space variables. But then, $R_{0,T}$ -a.s.

$$\frac{d\mu^T}{dR_{0,T}}(x, v, y, w) = \frac{d(\text{proj}_{x_1}, \text{proj}_{x_2})_{\#} \mu^T}{d(\text{proj}_{x_1}, \text{proj}_{x_2})_{\#} R_{0,T}}(x, y) = \frac{d\pi}{dR_{0,T}^X}(x, y) = f^T(x)g^T(y). \quad \square$$

Proof of Proposition 1.1. We have already seen in the previous proof that $\mathcal{C}_T(\mu, \nu)$ is finite. Now, since KSP is equivalent to the minimisation problem (3.1), from [34, Proposition 6.1] it follows

$$\begin{aligned} \mathcal{C}_T(\mu, \nu) &= \sup_{\varphi, \psi \in C_b(\mathbb{R}^d)} \left\{ \int_{\mathbb{R}^d} \varphi d\mu + \int_{\mathbb{R}^d} \psi d\nu - \int_{\mathbb{R}^{2d}} (e^{\varphi \oplus \psi} - 1) dR_{0,T}^X \right\} \\ &\leq \sup_{\varphi, \psi \in C_b(\mathbb{R}^d)} \left\{ \int_{\mathbb{R}^d} (\varphi \oplus \psi) d\pi - \log \int_{\mathbb{R}^{2d}} e^{\varphi \oplus \psi} dR_{0,T}^X \right\} \\ &\leq \sup_{h \in C_b(\mathbb{R}^{2d})} \left\{ \int_{\mathbb{R}^d} h d\pi - \log \int_{\mathbb{R}^{2d}} e^h dR_{0,T}^X \right\} \stackrel{(\dagger)}{=} \mathcal{H}(\pi|R_{0,T}^X) = \mathcal{C}_T(\mu, \nu), \end{aligned}$$

where π is the unique optimizer in (3.1), while (\dagger) is the Donsker-Varadhan variational formula [24, Lemma 1.4.3a]. This concludes the proof since $R_{0,T}^X := (\text{proj}_{x_1}, \text{proj}_{x_2})_{\#} R_{0,T}$. \square

The Schrödinger system (1.8) is particularly useful when f^T and g^T are regular enough. Under (H1) and (H4) they inherit the regularity (smoothness and integrability) of the densities of μ, ν respectively. This follows from the identities

$$\frac{d\mu}{d\mathbf{m}_X} = f^T \int_{\mathbb{R}^d} P_T g^T d\mathbf{m}_V, \quad \frac{d\nu}{d\mathbf{m}_X} = g^T \int_{\mathbb{R}^d} P_T^* f^T d\mathbf{m}_V,$$

and since $P_T^* f^T$ and $P_T g^T$ are smooth and positive (as a result of the lower bound (2.5)). Moreover, arguing exactly as in Lemma 2.1 in [19], owing to the lower bound in (2.5),

and the continuity of p_T , we have that there is $c_{T_0} > 0$, (possibly depending on μ and ν) such that for all $T \geq T_0$

$$\|f^T\|_{L^\infty(\mathfrak{m})}\|g^T\|_{L^1(\mathfrak{m})} \leq c_{T_0} \left\| \frac{d\mu}{d\mathfrak{m}_X} \right\|_{L^\infty(\mathfrak{m})}, \quad \|f^T\|_{L^1(\mathfrak{m})}\|g^T\|_{L^\infty(\mathfrak{m})} \leq c_{T_0} \left\| \frac{d\nu}{d\mathfrak{m}_X} \right\|_{L^\infty(\mathfrak{m})}.$$

These bounds are pivotal to prove that $f^T \rightarrow d\mu/d\mathfrak{m}_X$ and $g^T \rightarrow d\nu/d\mathfrak{m}_X$ as $T \rightarrow \infty$ in $L^p(\mathfrak{m})$ for all $p \in [1, \infty)$ akin to what is done in Lemma 3.6 of [19].

To ensure that f^T, g^T are in $L^\infty(\mathfrak{m})$ and with compact support, we work under assumption (H1) and (H4) for the rest of the section. With the help of the forward and adjoint semigroup, and (1.7) we can write

$$\begin{cases} \mu_0^T = f^T P_T g^T \mathfrak{m}, \\ \mu_T^T = g^T P_T^* f^T \mathfrak{m}, \end{cases} \tag{3.6}$$

where we recall that $\mu_t^T = (X_t, V_t)_\# P^T$, with P^T being optimal for KSPd. Furthermore, if we set,

$$f_t^T := P_t^* f^T \quad \text{and} \quad g_t^T := P_{T-t} g^T,$$

then $\mu_t^T, t \in [0, T]$, can be represented as

$$d\mu_t^T = f_t^T g_t^T d\mathfrak{m}. \tag{3.7}$$

It is also immediate to check that it holds

$$\begin{cases} \partial_t f_t^T = L^* f_t^T \\ \partial_t g_t^T = -L g_t^T \end{cases} \quad \text{and} \quad \begin{cases} \partial_t \log f_t^T = L^* \log f_t^T + \Gamma(\log f_t^T) \\ \partial_t \log g_t^T = -L \log g_t^T - \Gamma(\log g_t^T), \end{cases} \tag{3.8}$$

where $\Gamma(h) = \gamma |\nabla_v h|^2$ is the *carré du champ* operator associated to the generator L .

The fg -decomposition gives us a nice representation formula for the relative entropy along the entropic interpolation $(\mu_t^T)_{t \in [0, T]}$. Indeed, if we introduce the functions

$$h_f^T(t) := \int_{\mathbb{R}^{2d}} \log f_t^T \rho_t^T d\mathfrak{m} \quad \text{and} \quad h_b^T(t) := \int_{\mathbb{R}^{2d}} \log g_t^T \rho_t^T d\mathfrak{m} \quad \forall t \in [0, T],$$

then it easily follows that

$$\mathcal{H}(\mu_t^T | \mathfrak{m}) = h_f^T(t) + h_b^T(t), \quad \forall t \in [0, T]. \tag{3.9}$$

Moreover, we have

$$\partial_t h_f^T(t) = - \int_{\mathbb{R}^{2d}} \Gamma(\log f_t^T) \rho_t^T d\mathfrak{m} \quad \text{and} \quad \partial_t h_b^T(t) = \int_{\mathbb{R}^{2d}} \Gamma(\log g_t^T) \rho_t^T d\mathfrak{m}. \tag{3.10}$$

For a proof of (3.10) we refer to Lemma 3.8 in [18] where the classical setting is studied, the only difference in the kinetic setting being that the operator that acts on f_t^T should be replaced with L^* , since L is not self-adjoint.

In addition to (3.9), the fg -decomposition gives the following representation for the kinetic entropic cost

$$\begin{aligned} \mathcal{C}_T(\mu, \nu) &= \mathcal{H}(\mu^T | \mathbb{R}_{0, T}) = \mathbb{E}_{\mathbb{R}_{0, T}} [\rho^T \log \rho^T] \\ &= \int_{\mathbb{R}^{2d}} \log f^T \rho_0^T d\mathfrak{m} + \int_{\mathbb{R}^{2d}} \log g^T \rho_T^T d\mathfrak{m} = h_f^T(0) + h_b^T(T). \end{aligned} \tag{3.11}$$

As a byproduct of (3.9), (3.10), and (3.11), we get the identities²

$$\begin{aligned} \mathcal{C}_T(\mu, \nu) &= \mathcal{H}(\mu_0^T | \mathbf{m}) + \int_0^T \int_{\mathbb{R}^{2d}} \Gamma(\log g_t^T) \rho_t^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}t, \\ \mathcal{C}_T(\mu, \nu) &= \mathcal{H}(\mu_T^T | \mathbf{m}) + \int_0^T \int_{\mathbb{R}^{2d}} \Gamma(\log f_t^T) \rho_t^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}t. \end{aligned} \tag{3.12}$$

A straightforward consequence of the previous identities is the following

Lemma 3.1. *Under the assumptions (H1), (H4), for any $t \in [0, T]$ it holds*

$$\begin{aligned} \mathcal{C}_T(\mu, \nu) &= \mathcal{H}(\mu_0^T | \mathbf{m}) + \mathcal{H}(\mu_T^T | \mathbf{m}) + \int_0^t \int_{\mathbb{R}^{2d}} \Gamma(\log g_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s \\ &\quad + \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log f_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s - \mathcal{H}(\mu_t^T | \mathbf{m}). \end{aligned} \tag{3.13}$$

Proof. From (3.12) we can write

$$\mathcal{C}_T(\mu, \nu) = \mathcal{H}(\mu_0^T | \mathbf{m}) + \int_0^t \int_{\mathbb{R}^{2d}} \Gamma(\log g_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s + \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log g_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s.$$

Applying the identities (3.10) we obtain that the last summand equals

$$\begin{aligned} \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log g_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s &= \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log f_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s + \int_t^T \partial_s h_b^T(s) + \partial_s h_f^T(s) \, \mathrm{d}s \\ &= \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log f_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s + \int_t^T \partial_s \mathcal{H}(\mu_s^T | \mathbf{m}) \, \mathrm{d}s \\ &= \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log f_s^T) \rho_s^T \, \mathrm{d}\mathbf{m} \, \mathrm{d}s + \mathcal{H}(\mu_T^T | \mathbf{m}) - \mathcal{H}(\mu_t^T | \mathbf{m}), \end{aligned}$$

and we reach our conclusion. □

3.2 Corrector estimates and proof of Theorem 1.6

Throughout we assume (H1), (H2) and (H4) to be true and we will point out whenever the latter can be relaxed to (H3). Let us start by defining a few key objects whose behaviour will help us in controlling the convergence rates for the turnpike property.

We define the *correctors* as the functions $\varphi^T, \psi^T : [0, T] \rightarrow \mathbb{R}$ given by

$$\varphi^T(s) := \int_{\mathbb{R}^{2d}} |\nabla \log f_s^T|_{N-1}^2 \rho_s^T \, \mathrm{d}\mathbf{m} \quad \text{and} \quad \psi^T(s) := \int_{\mathbb{R}^{2d}} |\nabla \log g_s^T|_{M-1}^2 \rho_s^T \, \mathrm{d}\mathbf{m}, \tag{3.14}$$

where $M, N \in \mathbb{R}^{2d \times 2d}$ are the matrices appearing in Proposition 2.2. Let us also note that by the fg -decomposition it follows $\mathcal{I}(\mu_s^T) \lesssim \varphi^T(s) + \psi^T(s)$.

With the next lemma we show that the contraction properties introduced in the previous section translate into an exponentially fast contraction for φ^T and ψ^T .

Lemma 3.2. *Under (H1), (H2) and (H4), for any $0 < t \leq s \leq T$ it holds*

$$\varphi^T(s) \leq \varphi^T(t) e^{-2\kappa(s-t)} \quad \text{and} \quad \psi^T(T-s) \leq \psi^T(T-t) e^{-2\kappa(s-t)}. \tag{3.15}$$

Proof. By definition $f_s^T = P_{s-t}^* f_t^T$ and thus

$$\varphi^T(s) = \int_{\mathbb{R}^{2d}} |\nabla \log f_s^T|_{N-1}^2 \rho_s^T \, \mathrm{d}\mathbf{m} = \int_{\mathbb{R}^{2d}} |\nabla P_{s-t}^* f_t^T|_{N-1}^2 (P_{s-t}^* f_t^T)^{-1} g_s^T \, \mathrm{d}\mathbf{m} \tag{3.16}$$

²Let us point out that when considering the optimal solution $P^T \in \mathcal{P}(\Omega)$ in (1.2), the optimal control is given by $\alpha_t^P = 2\gamma \nabla_v \log g_t^T(X_t)$ and the stochastic control formulation (1.2) reads as the first identity in (3.12).

An application of the gradient estimate (2.11) and Cauchy-Schwartz inequality yields

$$\begin{aligned} \varphi^T(s) &\stackrel{(2.11)}{\leq} e^{-2\kappa(s-t)} \int_{\mathbb{R}^{2d}} (P_{s-t}^* |\nabla f_t^T|_{N-1})^2 (P_{s-t}^* f_t^T)^{-1} P_{T-s} g^T \, dm \\ &\leq e^{-2\kappa(s-t)} \int_{\mathbb{R}^{2d}} P_{s-t}^* \left(\frac{|\nabla f_t^T|_{N-1}^2}{f_t^T} \right) P_{T-s} g^T \, dm \\ &= e^{-2\kappa(s-t)} \int_{\mathbb{R}^{2d}} |\nabla \log f_t^T|_{N-1}^2 \rho_t^T \, dm \leq e^{-2\kappa(s-t)} \varphi^T(t), \end{aligned} \tag{3.17}$$

which concludes the proof for the first inequality. The analogous inequality for ψ^T runs as above by using inequality (2.10) for the semigroup $(P_t)_{t \in [0, T]}$. \square

Proposition 3.3. *Grant (H1), (H2) and (H4). There exists $C_{d,\alpha,\beta,\gamma} > 0$ such that for any $0 < \delta \leq 1$ and for any $t \in [\delta, T]$, as soon as $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$, it holds*

$$\varphi^T(t) \lesssim e^{-2\kappa t} [\mathcal{I}(\mu_\delta^T) + \mathcal{I}(\mu_{T-\delta}^T)] \quad \text{and} \quad \psi^T(T-t) \lesssim e^{-2\kappa t} [\mathcal{I}(\mu_\delta^T) + \mathcal{I}(\mu_{T-\delta}^T)]. \tag{3.18}$$

Proof. Without loss of generalities we may assume $\mathcal{I}(\mu_\delta^T)$ and $\mathcal{I}(\mu_{T-\delta}^T)$ to be finite, otherwise the above bounds are trivial. From Lemma 3.2 and the fg -decomposition of $\rho_t^T = f_t^T g_t^T$ we know that

$$\begin{aligned} \varphi^T(T-\delta) &\leq e^{-2\kappa T + 4\kappa\delta} \varphi^T(\delta) \\ &= e^{-2\kappa T + 4\kappa\delta} \int |\nabla \log \rho_\delta^T - \nabla \log g_\delta^T|_{N-1}^2 \, d\mu_\delta^T \\ &\lesssim e^{-2\kappa T + 4\kappa\delta} \mathcal{I}(\mu_\delta^T) + e^{-2\kappa T + 4\kappa\delta} \psi^T(\delta) \\ &\lesssim e^{-2\kappa T + 4\kappa\delta} \mathcal{I}(\mu_\delta^T) + e^{-4\kappa T + 8\kappa\delta} \psi^T(T-\delta). \end{aligned}$$

Using the basic inequality $|a - b|^2 \geq a^2/2 - b^2$ we obtain

$$\varphi^T(T-\delta) = \int_{\mathbb{R}^{2d}} |\nabla \log g_{T-\delta}^T - \nabla \log \rho_{T-\delta}^T|_{N-1}^2 \, d\mu_{T-\delta}^T \gtrsim \psi^T(T-\delta) - 2\mathcal{I}(\mu_{T-\delta}^T).$$

As a result, we get

$$\psi^T(T-\delta) - 2\mathcal{I}(\mu_{T-\delta}^T) \lesssim e^{-2\kappa T + 4\kappa\delta} \mathcal{I}(\mu_\delta^T) + e^{-4\kappa T + 8\kappa\delta} \psi^T(T-\delta).$$

Therefore, as soon as $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$ for some constant $C_{d,\alpha,\beta,\gamma} > 0$, we find

$$\psi^T(T-\delta) \lesssim \mathcal{I}(\mu_\delta^T) + \mathcal{I}(\mu_{T-\delta}^T).$$

Plugging this bound into the contraction estimate (3.15) gives the second inequality in (3.18) for any $t \in [\delta, T]$. The first inequality is obtained by exchanging the roles of φ and ψ in the above discussion. \square

Proposition 3.4. *Assume (H1), (H2) and (H3). Let $0 < \delta \leq 1$ be fixed. Then, for all $t \in [\delta, T - \delta]$*

$$\mathcal{I}(\mu_t^T) \lesssim \delta^{-3} \left(\mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu_t^T \mid \mathbf{m}) \right). \tag{3.19}$$

Proof. Let us first work under (H4). We claim that for any $t \in [\delta, T - \delta]$ it holds

$$|\nabla P_{T-t} g^T|^2 \lesssim \delta^{-3} \left[P_{T-t}(g^T \log g^T) - (P_{T-t} g^T) \log(P_{T-t} g^T) \right] P_{T-t} g^T. \tag{3.20}$$

Indeed by applying Corollary 3.2 in [30] to any directional derivative we have

$$\begin{aligned} |\partial_{x_i} P_{T-t} g^T|^2 &\leq 4 \inf_{s \in (0, T-t]} \Psi_s(1, 0) \left[P_{T-t}(g^T \log g^T) - (P_{T-t} g^T) \log(P_{T-t} g^T) \right] P_{T-t} g^T, \\ |\partial_{v_i} P_{T-t} g^T|^2 &\leq 4 \inf_{s \in (0, T-t]} \Psi_s(0, 1) \left[P_{T-t}(g^T \log g^T) - (P_{T-t} g^T) \log(P_{T-t} g^T) \right] P_{T-t} g^T, \end{aligned}$$

where $\Psi_s(a, b)$ is defined for any $a, b > 0$ as the quantity

$$\Psi_s(a, b) := \frac{1}{2\gamma} s \left[a \left(\frac{6}{s^2} + \beta + \frac{3\gamma}{2s} \right) + b \left(\frac{4}{s} + \frac{4\beta}{27} s + \gamma \right) \right]^2.$$

By considering $s = \delta \in (0, 1]$ we can bound the above RHS with δ^{-3} , up to a multiplicative constant. Particularly this yields (3.20). Similarly one can prove that it holds

$$|\nabla P_t^* f^T|^2 \lesssim \delta^{-3} \left[P_t^*(f^T \log f^T) - (P_t^* f^T) \log(P_t^* f^T) \right] P_t^* f^T.$$

Therefore because of the fg -decomposition (3.7) we obtain

$$\begin{aligned} \mathcal{I}(\mu_t^T) &\leq 2 \int_{\mathbb{R}^{2d}} \left[\frac{|\nabla P_{T-t} g^T|^2}{P_{T-t} g^T} P_t^* f^T + \frac{|\nabla P_t^* f^T|^2}{P_t^* f^T} P_{T-t} g^T \right] dm \\ &\lesssim \delta^{-3} \int_{\mathbb{R}^{2d}} \left\{ \left[P_{T-t}(g^T \log g^T) - (P_{T-t} g^T) \log(P_{T-t} g^T) \right] P_t^* f^T \right. \\ &\quad \left. + \left[P_t^*(f^T \log f^T) - (P_t^* f^T) \log(P_t^* f^T) \right] P_{T-t} g^T \right\} dm \end{aligned}$$

By integration by parts and (3.7) this last displacement equals

$$\delta^{-3} \left(\int_{\mathbb{R}^d} \log g^T d\nu + \int_{\mathbb{R}^d} \log f^T d\mu - \int_{\mathbb{R}^{2d}} \log \rho_t^T \rho_t^T dm \right),$$

and the thesis follows in view of (3.11).

Now let us just assume (H3). Firstly, define the probability measure q_n^T as the measure whose $R_{0,T}$ -density is given by

$$\frac{dq_n^T}{dR_{0,T}} := (\rho^T \wedge n) \frac{\mathbb{1}_{K_n}}{C_n}, \tag{3.21}$$

where $(K_n)_{n \in \mathbb{N}}$ is an increasing sequence of compact sets in \mathbb{R}^{4d} and C_n is the normalising constant. Then, by applying Lemma 4.1 we know that the marginals $\mu^n := (\text{proj}_{x_1})_{\#} q_n^T$ and $\nu^n := (\text{proj}_{x_2})_{\#} q_n^T$ satisfy (H4) and by means of Proposition 4.2 and Corollary 4.3 it follows that there exists a unique minimizer $\mu^{n,T} \in \mathcal{P}(\Omega)$ for KSP with marginals μ^n, ν^n , and as soon as n diverges it holds

$$\mu_t^{n,T} \rightharpoonup \mu_t^T \quad \text{and} \quad \mathcal{C}_T(\mu^n, \nu^n) \rightarrow \mathcal{C}_T(\mu, \nu). \tag{3.22}$$

Then, the thesis in the general case follows from the one under (H4) and the lower semicontinuity of $\mathcal{I}(\cdot)$ and $\mathcal{H}(\cdot|\mathbf{m})$. \square

As a byproduct of Proposition 3.3 and Proposition 3.4 we get

Corollary 3.5. *Under (H1), (H2) and (H4), there exists $C_{d,\alpha,\beta,\gamma} > 0$ such that for any $0 < \delta \leq 1$ and $t \in [\delta, T]$, as soon as $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$, it holds*

$$\varphi^T(t) \lesssim \delta^{-3} e^{-2\kappa t} \mathcal{C}_T(\mu, \nu) \quad \text{and} \quad \psi^T(T-t) \lesssim \delta^{-3} e^{-2\kappa t} \mathcal{C}_T(\mu, \nu). \tag{3.23}$$

Proof of Theorem 1.6. We start proving the result under (H4). Since $\mathcal{I}(\mu_t^T) \lesssim \varphi^T(t) + \psi^T(t)$, the first inequality (cf. (1.19)) is an immediate consequence of Corollary 3.5. The relative entropy bound (cf. (1.20)) follows from the first one by means of (2.4). In order to extend (1.19) and (1.20) to (H3), it is enough to consider the approximation of the optimizer (cf. (3.21) and (3.22)) together with the lower semicontinuity of $\mathcal{I}(\cdot)$ and $\mathcal{H}(\cdot|\mathfrak{m})$. Finally, (1.21) follows from (1.20) by means of Lemma 3.9 below. \square

3.3 Long-time behaviour of the kinetic entropic cost

Throughout the whole section we will always assume (H1) and (H3) to be true. Let us remark that (H1) implies that $\mathfrak{m} \in \mathcal{P}_2(\mathbb{R}^{2d})$. In what follows we are going to prove Theorem 1.3, but first we need some preparation. The first two claims in Theorem 1.3 will be proved via a Γ -convergence approach (cf. Proposition 3.6 and Lemma 3.7) similar to the one used in [19] for the classical Schrödinger problem. The main difference with [19] is the lack of compactness for the set $\Pi_X(\mu, \nu)$. On the other hand, the proof of the latter two bounds in Theorem 1.3 will rely on the corrector estimates given in Section 3.2.

Proposition 3.6 (A Γ -convergence result). *Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of positive real numbers converging to ∞ , and for each $n \in \mathbb{N}$ consider the functional $\mathcal{H}(\cdot | R_{0,T_n})$ defined on $\Pi_X(\mu, \nu)$ endowed with the weak topology. Then*

$$\Gamma - \lim_{n \rightarrow \infty} \mathcal{H}(\cdot | R_{0,T_n}) = \mathcal{H}(\cdot | \mathfrak{m} \otimes \mathfrak{m}).$$

Proof. (Γ -convergence lower bound inequality) We prove that for any sequence $(q_n)_{n \in \mathbb{N}} \subset \Pi_X(\mu, \nu)$ that converges weakly to some $q \in \Pi_X(\mu, \nu)$

$$\liminf_{n \rightarrow \infty} \mathcal{H}(q_n | R_{0,T_n}) \geq \mathcal{H}(q | \mathfrak{m} \otimes \mathfrak{m}). \tag{3.24}$$

Note that since P_t is strongly mixing [20, Theorem 11.14] for any $\psi, \phi \in C_b(\mathbb{R}^{2d})$ it holds

$$\int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \psi(x, v) \phi(y, w) dR_{0,T_n} = \int_{\mathbb{R}^{2d}} \psi P_{T_n} \phi d\mathfrak{m} \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \psi(x, v) \phi(y, w) d\mathfrak{m} \otimes d\mathfrak{m}.$$

From the Portmanteau Theorem, it follows that $R_{0,T_n} \rightharpoonup \mathfrak{m} \otimes \mathfrak{m}$. Then (3.24) follows from the lower semicontinuity of the relative entropy.

(Γ -convergence upper bound inequality) We prove that for any $q \in \Pi_X(\mu, \nu)$ it holds

$$\limsup_{n \rightarrow \infty} \mathcal{H}(q | R_{0,T_n}) \leq \mathcal{H}(q | \mathfrak{m} \otimes \mathfrak{m}). \tag{3.25}$$

We may assume $\mathcal{H}(q | \mathfrak{m} \otimes \mathfrak{m}) < \infty$ otherwise the above inequality is trivial. Note that this implies $q \in \mathcal{P}_2(\mathbb{R}^{4d})$ since $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ (cf. (2.3)) while

$$\begin{aligned} \int_{\mathbb{R}^d} |v|^2 d(\text{proj}_{v_1})_{\#} q &\leq 2 \int_{\mathbb{R}^d} |v|^2 d\mathfrak{m}_V + 2 \mathcal{W}_2((\text{proj}_{v_1})_{\#} q, \mathfrak{m}_V)^2 \\ &\stackrel{(2.2)}{\lesssim} 1 + \mathcal{H}((\text{proj}_{v_1})_{\#} q | \mathfrak{m}_V) \leq 1 + \mathcal{H}(q | \mathfrak{m} \otimes \mathfrak{m}) < \infty, \end{aligned}$$

and similarly for the measure $(\text{proj}_{v_2})_{\#} q$. Then we have

$$\mathcal{H}(q | R_{0,T_n}) = \mathcal{H}(q | \mathfrak{m} \otimes \mathfrak{m}) - \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \log p_{T_n}((x, v), (y, w)) dq,$$

Thanks to the lower bound given in Lemma 2.1 and the fact that $q \in \mathcal{P}_2(\mathbb{R}^{4d})$, we can apply Fatou's Lemma and get

$$\limsup_{n \rightarrow \infty} \mathcal{H}(q | R_{0,T_n}) \leq \mathcal{H}(q | \mathfrak{m} \otimes \mathfrak{m}) - \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \liminf_{n \rightarrow \infty} \log p_{T_n}((x, v), (y, w)) dq. \tag{3.26}$$

Now, for all $t > 0$ and for all $(x, v), (y, w) \in \mathbb{R}^{2d}$

$$p_{T_n}((x, v), (y, w)) = P_{T_n-t}(p_t(\cdot, (y, w)))(x, v).$$

For any $M > 0$, we introduce the function $p_t^M(\cdot, (y, w)) := p_t(\cdot, (y, w)) \wedge M \in C_b(\mathbb{R}^{2d})$. Then, since P_T is strongly mixing (cf. [20, Theorem 11.14]), we get

$$p_{T_n}((x, v), (y, w)) \geq P_{T_n-t}(p_t^M(\cdot, (y, w)))(x, v) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^{2d}} p_t^M((x, v), (y, w)) \, d\mathbf{m}(x, v).$$

Taking the limit as $M \rightarrow \infty$, by dominated convergence we get that

$$\int_{\mathbb{R}^{2d}} p_t^M((x, v), (y, w)) \, d\mathbf{m}(x, v) \xrightarrow{M \rightarrow \infty} \int_{\mathbb{R}^{2d}} p_t((x, v), (y, w)) \, d\mathbf{m}(x, v) = 1,$$

where the last equality follows from (2.6). Therefore it holds

$$\liminf_{n \rightarrow \infty} \log p_{T_n}((x, v), (y, w)) \geq 0, \quad \mathbf{m} \otimes \mathbf{m} - a.s.$$

which, together with $q \ll \mathbf{m} \otimes \mathbf{m}$ (since $\mathcal{H}(q | \mathbf{m} \otimes \mathbf{m}) < \infty$), leads to

$$\liminf_{n \rightarrow \infty} \log p_{T_n}((x, v), (y, w)) \geq 0, \quad q\text{-a.s.}$$

Therefore, from (3.26) we get inequality (3.25). The desired Γ -convergence follows as a byproduct of (3.24) and (3.25). \square

Even though here we are just interested in the Γ -convergence (as introduced by De Giorgi) on $\Pi_X(\mu, \nu)$ equipped with the weak topology, the previous result is actually stronger: indeed we have actually proven the Mosco convergence of the functional $\mathcal{H}(\cdot | R_{0, T_n})$ since we have considered a constant sequence $q_n = q$ for the upper bound inequality.

Lemma 3.7 (Equicoerciveness). *The family of functionals $\mathcal{H}(\cdot | R_{0, T_n}) : \Pi_X(\mu, \nu) \rightarrow [0, \infty]$, $n \in \mathbb{N}$ is equicoercive, i.e. for any $h \in \mathbb{R}$ there exists a (weakly) compact subset $K_h \subset \Pi_X(\mu, \nu)$ such that*

$$\{q \in \Pi_X(\mu, \nu) \text{ s.t. } \mathcal{H}(q | R_{0, T_n}) \leq h\} \subseteq K_h \quad \forall n \in \mathbb{N}.$$

Proof. Since $(R_{0, T_n})_{n \in \mathbb{N}}$ is tight, a proof of this result is obtained by following the same argument given in [24, Lemma 1.4.3c]. \square

The next two results are a consequence of the corrector estimates of Section 3.2. In the first one we consider the long-time behaviour of the marginals of the solution to KSP at times $t = 0, T$; in the second one we give a bound for the entropic cost, uniformly in time.

Theorem 3.8. *Under assumptions (H1), (H2) and (H4) there exists a positive constant $C_{d, \alpha, \beta, \gamma}$ such that for any $0 < \delta \leq 1$ and $T > \frac{1}{\kappa} \log C_{d, \alpha, \beta, \gamma} + 2\delta$ it holds*

$$\begin{aligned} |\mathcal{H}(\mu_0^T | \mathbf{m}) - \mathcal{H}(\mu | \mathbf{m}_X)| &\leq C_{d, \alpha, \beta, \gamma} \delta^{-3} \mathcal{C}_T(\mu, \nu) e^{-2\kappa T}, \\ |\mathcal{H}(\mu_T^T | \mathbf{m}) - \mathcal{H}(\nu | \mathbf{m}_X)| &\leq C_{d, \alpha, \beta, \gamma} \delta^{-3} \mathcal{C}_T(\mu, \nu) e^{-2\kappa T}. \end{aligned} \tag{3.27}$$

Proof. We will prove only the first bound since the second one can be proved similarly. Since $\frac{d\mu}{d\mathbf{m}_X}(\cdot) = \int_{\mathbb{R}^d} \rho_0^T(\cdot, v) \, d\mathbf{m}_V(v)$, the log-Sobolev inequality for the Gaussian measure

m_V gives

$$\begin{aligned} \mathcal{H}(\mu_0^T | m) - \mathcal{H}(\mu | m_X) &= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} \rho_0^T \log \rho_0^T - \left(\int_{\mathbb{R}^d} \rho_0^T dm_V \right) \log \left(\int_{\mathbb{R}^d} \rho_0^T dm_V \right) dm_V \right] dm_X \\ &\leq \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} \left| \nabla_v \sqrt{\rho_0^T} \right|^2 dm_V \right] dm_X = \int_{\mathbb{R}^{2d}} \frac{1}{2} \left| \frac{\nabla_v \rho_0^T}{\rho_0^T} \right|^2 \rho_0^T dm \\ &= \frac{1}{2} \int_{\mathbb{R}^{2d}} \left| \nabla_v \log(f^T P_T g^T) \right|^2 \rho_0^T dm = \frac{1}{2} \int_{\mathbb{R}^{2d}} \left| \nabla_v \log f^T + \nabla_v \log P_T g^T \right|^2 \rho_0^T dm \\ &= \frac{1}{2} \int_{\mathbb{R}^{2d}} \left| \nabla_v \log g_0^T \right|^2 \rho_0^T dm \lesssim \psi^T(0) \stackrel{(3.23)}{\lesssim} \delta^{-3} \mathcal{C}_T(\mu, \nu) e^{-2\kappa T}, \end{aligned}$$

where the equality in the last line follows from the fact that $f^T = f^T(x)$ does not depend on the velocity variable. Finally, since $(\text{proj}_x)_\# \mu_0^T = \mu$ we know that the left hand side term above is positive. \square

Since it holds $\mathcal{H}(\mu | m_X) = \mathcal{H}((X_0)_\# \mu^T | (X_0)_\# \mathbb{R}_{0,T}) \leq \mathcal{H}(\mu^T | \mathbb{R}_{0,T}) = \mathcal{C}_T(\mu, \nu)$, and similarly $\mathcal{H}(\nu | m_X) \leq \mathcal{C}_T(\mu, \nu)$, the following lower bound is always true

$$\mathcal{C}_T(\mu, \nu) \geq \frac{\mathcal{H}(\mu | m_X) + \mathcal{H}(\nu | m_X)}{2}. \tag{3.28}$$

We now give a corresponding upper bound for sufficiently large times.

Lemma 3.9. *Under (H1) and (H2) there exists a constant $C_{d,\alpha,\beta,\gamma} > 0$ such that for any $0 < \delta \leq 1$ and $T > (\frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta) \vee (\frac{1}{\kappa} \log \frac{C_{d,\alpha,\beta,\gamma}}{\delta^3})$ it holds*

$$\mathcal{C}_T(\mu, \nu) \leq C_{d,\alpha,\beta,\gamma} \left[\mathcal{H}(\mu | m_X) + \mathcal{H}(\nu | m_X) \right]. \tag{3.29}$$

Proof. Firstly, let us assume (H4) to hold. Owing to $|\nabla_v \log g_s^T|^2 \lesssim |\nabla \log g_s^T|_{M^{-1}}^2$ and $|\nabla_v \log f_s^T|^2 \lesssim |\nabla \log f_s^T|_{N^{-1}}^2$, from (3.13) it follows

$$\begin{aligned} \mathcal{C}_T(\mu, \nu) &\leq \mathcal{H}(\mu_0^T | m) + \mathcal{H}(\mu_T^T | m) + \int_0^{\frac{T}{2}} \psi^T(s) ds + \int_{\frac{T}{2}}^T \varphi^T(s) ds \\ &\stackrel{(3.27)}{\lesssim} \mathcal{H}(\mu | m_X) + \mathcal{H}(\nu | m_X) + 2\delta^{-3} \mathcal{C}_T(\mu, \nu) e^{-2\kappa T} + \int_0^{\frac{T}{2}} \psi^T(s) ds + \int_{\frac{T}{2}}^T \varphi^T(s) ds. \end{aligned}$$

We first consider $\int_{T/2}^T \varphi^T(s) ds$. For any $s \in [T/2, T]$ from Corollary 3.5 we have

$$\int_{\frac{T}{2}}^T \varphi^T(s) ds \lesssim \delta^{-3} \mathcal{C}_T(\mu, \nu) \int_{\frac{T}{2}}^T e^{-2\kappa s} ds \lesssim \delta^{-3} \mathcal{C}_T(\mu, \nu) (e^{-\kappa T} - e^{-2\kappa T}). \tag{3.30}$$

By reasoning in the same way, this time by using the fact that $s \in [0, T/2]$, we get

$$\int_0^{\frac{T}{2}} \psi^T(s) ds \lesssim \delta^{-3} \mathcal{C}_T(\mu, \nu) \int_0^{\frac{T}{2}} e^{-2\kappa(T-s)} ds \lesssim \delta^{-3} \mathcal{C}_T(\mu, \nu) (e^{-\kappa T} - e^{-2\kappa T}). \tag{3.31}$$

Therefore there exists a positive constant $C_{d,\alpha,\beta,\gamma}$ such that

$$\mathcal{C}_T(\mu, \nu) \leq C_{d,\alpha,\beta,\gamma} \left[\mathcal{H}(\mu | m_X) + \mathcal{H}(\nu | m_X) + \delta^{-3} \mathcal{C}_T(\mu, \nu) (e^{-\kappa T} - e^{-2\kappa T}) \right],$$

which yields our thesis as soon as $T > \frac{1}{\kappa} \log \frac{C_{d,\alpha,\beta,\gamma}}{\delta^3}$ for a well chosen $C_{d,\alpha,\beta,\gamma} > 0$.

Now, let us prove the result under (H2). Firstly, notice that we may assume that μ and ν satisfy (H3), otherwise the bound is trivial. The main idea is defining the probability measures μ_n^M and ν_n^M on \mathbb{R}^d , approximating μ and ν , as the measures whose \mathfrak{m}_X -densities are given by

$$\frac{d\mu_n^M}{d\mathfrak{m}_X} := \left(\frac{d\mu}{d\mathfrak{m}_X} \wedge n \right) \frac{\mathbb{1}_{K_n}}{C_n^\mu} \quad \text{and} \quad \frac{d\nu_n^M}{d\mathfrak{m}_X} := \left(\frac{d\nu}{d\mathfrak{m}_X} \wedge n \right) \frac{\mathbb{1}_{K_n}}{C_n^\nu}, \quad (3.32)$$

where $(K_n)_{n \in \mathbb{N}}$ is an increasing sequence of compact sets in \mathbb{R}^d and C_n^μ, C_n^ν are the normalising constants. Then, μ_n^M and ν_n^M satisfy (H4) and by means of (4.2) it follows

$$\mathcal{H}(\mu_n^M | \mathfrak{m}_X) \xrightarrow{n \rightarrow \infty} \mathcal{H}(\mu | \mathfrak{m}_X) \quad \text{and} \quad \mathcal{H}(\nu_n^M | \mathfrak{m}_X) \xrightarrow{n \rightarrow \infty} \mathcal{H}(\nu | \mathfrak{m}_X). \quad (3.33)$$

Owing to Lemma 4.4 and (3.29) for the approximated μ_n^M, ν_n^M , we conclude our proof. \square

The results given in Theorem 3.8 and Lemma 3.9 will come at hand while proving the exponential convergence in (1.12).

Proof of Theorem 1.3. We start with the proof of (1.10) and (1.11). Firstly note that the unique minimizer in (1.9) is given the probability measure $\mu^\infty := (\mu \otimes \mathfrak{m}_V) \otimes (\nu \otimes \mathfrak{m}_V)$. Now let us consider $(T_n)_{n \in \mathbb{N}}$ to be any diverging sequence of positive real times. Then, from the optimality of μ^{T_n} it follows

$$\limsup_{n \rightarrow \infty} \mathcal{H}(\mu^{T_n} | \mathbb{R}_{0, T_n}) \leq \limsup_{n \rightarrow \infty} \mathcal{H}(\mu^\infty | \mathbb{R}_{0, T_n}) \stackrel{(3.25)}{\leq} \mathcal{H}(\mu^\infty | \mathfrak{m} \otimes \mathfrak{m}) = \mathcal{H}(\mu | \mathfrak{m}_X) + \mathcal{H}(\nu | \mathfrak{m}_X),$$

which is finite by our assumptions. Then Lemma 3.7 implies that the subsequence $(\mu^{T_n})_{n \in \mathbb{N}}$ is weakly relatively compact. Then, from Proposition 3.6, the Fundamental theorem of Γ -convergence [7, Theorem 2.10], the uniqueness of the minimizer in (1.9) and from the metrizable of the weak convergence on $\mathcal{P}(\mathbb{R}^{4d})$ we deduce (1.10) and (1.11).

We continue with the proof of (1.12) and of the entropic Talagrand inequality (1.13). We firstly assume (H4) to hold. By (3.13) and owing to $|\nabla_v \log g_s^T|^2 \lesssim |\nabla \log g_s^T|_{M^{-1}}^2$ and $|\nabla_v \log f_s^T|^2 \lesssim |\nabla \log f_s^T|_{N^{-1}}^2$, we know that

$$|\mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu_0^T | \mathfrak{m}) - \mathcal{H}(\mu_T^T | \mathfrak{m})| \lesssim \mathcal{H}(\mu_{\frac{T}{2}}^T | \mathfrak{m}) + \int_0^{\frac{T}{2}} \psi^T(s) ds + \int_{\frac{T}{2}}^T \varphi^T(s) ds,$$

and from (3.30), (3.31) and the entropic turnpike (1.20) it follows

$$|\mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu_0^T | \mathfrak{m}) - \mathcal{H}(\mu_T^T | \mathfrak{m})| \lesssim \delta^{-3} e^{-\kappa T} \mathcal{C}_T(\mu, \nu) + \delta^{-3} \mathcal{C}_T(\mu, \nu) e^{-\kappa T}.$$

As a byproduct of the above inequality and Theorem 3.8 we get

$$|\mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu | \mathfrak{m}_X) - \mathcal{H}(\nu | \mathfrak{m}_X)| \leq C_{d, \alpha, \beta, \gamma} \delta^{-3} e^{-\kappa T} \mathcal{C}_T(\mu, \nu). \quad (3.34)$$

Let us now assume that μ and ν satisfy (H3) only. Firstly, consider the approximating sequence $(\mu_n^T)_{n \in \mathbb{N}}$ of the optimizer (cf. (3.21) and (3.22)). Then, by means of (3.34) under (H4) and the lower semicontinuity of the relative entropy we have

$$\begin{aligned} \mathcal{H}(\mu | \mathfrak{m}_X) + \mathcal{H}(\nu | \mathfrak{m}_X) - \mathcal{C}_T(\mu, \nu) &\stackrel{(3.22)}{\leq} \liminf_{n \rightarrow \infty} \left[\mathcal{H}(\mu^n | \mathfrak{m}_X) + \mathcal{H}(\nu^n | \mathfrak{m}_X) - \mathcal{C}_T(\mu^n, \nu^n) \right] \\ &\stackrel{(3.34)}{\lesssim} \delta^{-3} e^{-\kappa T} \liminf_{n \rightarrow \infty} \mathcal{C}_T(\mu^n, \nu^n) = \delta^{-3} e^{-\kappa T} \mathcal{C}_T(\mu, \nu) \stackrel{(3.29)}{\lesssim} \delta^{-3} e^{-\kappa T} \left[\mathcal{H}(\mu | \mathfrak{m}_X) + \mathcal{H}(\nu | \mathfrak{m}_X) \right]. \end{aligned}$$

For the other bound we are going to use the approximation on the marginals (cf. (3.32)). Therefore, let us consider μ_n^M, ν_n^M such that (H4) holds. Then, from Lemma 4.4 and the convergence of the relative entropies in (3.33), we get

$$\begin{aligned} \mathcal{C}_T(\mu, \nu) - \mathcal{H}(\mu | \mathbf{m}_X) - \mathcal{H}(\nu | \mathbf{m}_X) &\leq \liminf_{n \rightarrow \infty} \left[\mathcal{C}_T(\mu_n^M, \nu_n^M) - \mathcal{H}(\mu_n^M | \mathbf{m}_X) - \mathcal{H}(\nu_n^M | \mathbf{m}) \right] \\ &\stackrel{(3.34)}{\lesssim} \delta^{-3} e^{-\kappa T} \liminf_{n \rightarrow \infty} \mathcal{C}_T(\mu_n^M, \nu_n^M) \stackrel{(3.29)}{\lesssim} \delta^{-3} e^{-\kappa T} \liminf_{n \rightarrow \infty} \left[\mathcal{H}(\mu_n^M | \mathbf{m}_X) + \mathcal{H}(\nu_n^M | \mathbf{m}_x) \right] \\ &= \delta^{-3} e^{-\kappa T} \left[\mathcal{H}(\mu | \mathbf{m}_X) + \mathcal{H}(\nu | \mathbf{m}_X) \right]. \quad \square \end{aligned}$$

3.4 Corrector estimates for KFSP and proof of Theorems 1.5 and 1.7

In this section we collect results in the kinetic-full setting analogous to the ones already presented for KSP. We omit the proofs since the arguments are very similar and do not present any new difficulty with respect to the KSP case.

Let us start by mentioning that also in this case KFSP and KFSPd admit unique solutions $\bar{\mu}^T, \bar{P}^T$ with $\bar{\mu}^T = ((X_0, V_0), (X_T, V_T))_{\#} \bar{P}^T$ which can be decomposed as

$$\bar{\rho}^T(x, v, y, w) := \frac{d\bar{\mu}^T}{d\mathbb{R}_{0,T}}(x, v, y, w) = \bar{f}^T(x, v) \bar{g}^T(y, w) \quad \mathbb{R}_{0,T}\text{-a.s.} \quad (3.35)$$

where \bar{f}^T, \bar{g}^T are two non-negative measurable functions on \mathbb{R}^{2d} solving the Schrödinger system

$$\begin{cases} \frac{d\bar{\mu}}{dm}(x, v) = \bar{f}^T(x, v) \mathbb{E}_{\mathbb{R}}[\bar{g}^T(X_T, V_T) | X_0 = x, V_0 = v], \\ \frac{d\bar{\nu}}{dm}(y, w) = \bar{g}^T(y, w) \mathbb{E}_{\mathbb{R}}[\bar{f}^T(X_0, V_0) | X_T = y, V_T = w]. \end{cases} \quad (3.36)$$

Note that in this case f and g are function of both space and velocity. Moreover, if we define for any $t \in [0, T]$

$$\bar{f}_t^T := P_t^* \bar{f}^T \quad \text{and} \quad \bar{g}_t^T := P_{T-t} \bar{g}^T,$$

then $\bar{\mu}_t^T = (X_t, V_t)_{\#} \bar{P}^T$ can be written as $\bar{\mu}_t^T = \bar{f}_t^T \bar{g}_t^T \mathbf{m}$ and, similarly to (3.13), under (H1) and (FH4) it holds that, for any $t \in [0, T]$

$$\begin{aligned} \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) &= \mathcal{H}(\bar{\mu} | \mathbf{m}) + \mathcal{H}(\bar{\nu} | \mathbf{m}) - \mathcal{H}(\bar{\mu}_t^T | \mathbf{m}) \\ &\quad + \int_0^t \int_{\mathbb{R}^{2d}} \Gamma(\log \bar{g}_s^T) \bar{\rho}_s^T \, dm \, ds + \int_t^T \int_{\mathbb{R}^{2d}} \Gamma(\log \bar{f}_s^T) \bar{\rho}_s^T \, dm \, ds. \end{aligned} \quad (3.37)$$

We can therefore define the *correctors* as the functions $\bar{\varphi}^T, \bar{\psi}^T: [0, T] \rightarrow \mathbb{R}$ given by

$$\bar{\varphi}^T(s) := \int_{\mathbb{R}^{2d}} |\nabla \log \bar{f}_s^T|_{N^{-1}}^2 \bar{\rho}_s^T \, dm \quad \text{and} \quad \bar{\psi}^T(s) := \int_{\mathbb{R}^{2d}} |\nabla \log \bar{g}_s^T|_{M^{-1}}^2 \bar{\rho}_s^T \, dm, \quad (3.38)$$

where $M, N \in \mathbb{R}^{2d \times 2d}$ are positive definite symmetric matrices as appearing in Proposition 2.2. In the next result we collect all the contraction properties satisfied by the above correctors, which correspond to the ones proven for KSP in Lemma 3.2, Proposition 3.3, Proposition 3.4 and Corollary 3.5.

Lemma 3.10. *Grant (H1), (H2), (FH4) and fix $\delta \in (0, 1]$. For any $0 < t \leq s \leq T$ it holds*

$$\bar{\varphi}^T(s) \leq \bar{\varphi}^T(t) e^{-2\kappa(s-t)} \quad \text{and} \quad \bar{\psi}^T(T-s) \leq \bar{\psi}^T(T-t) e^{-2\kappa(s-t)} \quad \forall 0 < t \leq s \leq T.$$

Moreover, for any fixed $\delta \in (0, 1]$ as soon as $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$ the followings hold true

$$\begin{aligned} \bar{\varphi}^T(t) &\lesssim e^{-2\kappa t} \left[\mathcal{I}(\bar{\mu}_\delta^T) + \mathcal{I}(\bar{\mu}_{T-\delta}^T) \right], \quad \forall t \in [\delta, T], \\ \bar{\psi}^T(T-t) &\lesssim e^{-2\kappa t} \left[\mathcal{I}(\bar{\mu}_\delta^T) + \mathcal{I}(\bar{\mu}_{T-\delta}^T) \right], \quad \forall t \in [\delta, T], \end{aligned} \quad (3.39a)$$

$$\mathcal{I}(\bar{\mu}_t^T) \lesssim \delta^{-3} \left(\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) - \mathcal{H}(\bar{\mu}_t^T | \mathbf{m}) \right), \quad \forall t \in [\delta, T - \delta], \quad (3.39b)$$

$$\bar{\varphi}^T(t) \lesssim \delta^{-3} e^{-2\kappa t} \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) \quad \text{and} \quad \bar{\psi}^T(T-t) \lesssim \delta^{-3} e^{-2\kappa t} \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}), \quad \forall t \in [\delta, T]. \quad (3.39c)$$

Proof of Theorem 1.7. The proof of the result under (FH4) follows the same reasoning presented in the first part of the proof of Theorem 1.6 and for this reason is omitted. An approximating argument akin to the one in the proof of Theorem 1.6, this time considering the full marginals $\bar{\mu}^n, \bar{\nu}^n$ and the corresponding KFSP, gives

$$\bar{\mu}_t^{n,T} \rightharpoonup \bar{\mu}_t^T \quad \text{and} \quad \mathcal{C}_T^F(\bar{\mu}^n, \bar{\nu}^n) \rightarrow \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}).$$

Therefore the first two bounds follow from the lower semicontinuity of $\mathcal{I}(\cdot)$ and $\mathcal{H}(\cdot | \mathbf{m})$. Finally, (1.24) follows from (1.23) by means of (3.40) presented below. \square

Proof of Theorem 1.5. The proof of (1.14) and (1.15) runs similarly to the one given above in the kinetic setting and for this reason it is omitted. The main difference is that in this case the equicoerciveness is not needed since we have the weak compactness of $\Pi(\bar{\mu}, \bar{\nu})$.

We now discuss (1.16) and (1.17). With similar arguments as for KSP one can show that under (H1) and (H2), there exists a constant $C_{d,\alpha,\beta,\gamma} > 0$ such that for any $0 < \delta \leq 1$ and $T > (\frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta) \vee (\frac{1}{\kappa} \log \frac{C_{d,\alpha,\beta,\gamma}}{\delta^3})$ it holds

$$\mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) \leq C_{d,\alpha,\beta,\gamma} \left[\mathcal{H}(\bar{\mu} | \mathbf{m}) + \mathcal{H}(\bar{\nu} | \mathbf{m}) \right]. \quad (3.40)$$

Further, by means of (3.37), the corrector estimates (3.39c) and the turnpike estimate (1.23), at least under (FH4), it follows that for any $0 < \delta \leq 1$, as soon as $T > (\frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta) \vee \frac{1}{\kappa} \log \frac{C_{d,\alpha,\beta,\gamma}}{\delta^3}$,

$$\left| \mathcal{C}_T^F(\bar{\mu}, \bar{\nu}) - \mathcal{H}(\bar{\mu} | \mathbf{m}) - \mathcal{H}(\bar{\nu} | \mathbf{m}) \right| \leq C_{d,\alpha,\beta,\gamma} \delta^{-3} e^{-\kappa T} \left[\mathcal{H}(\bar{\mu} | \mathbf{m}) + \mathcal{H}(\bar{\nu} | \mathbf{m}) \right],$$

and from this immediately deduce (1.17). The extension to (FH3) is a consequence of a standard approximation argument. \square

3.5 Convergence over a fixed time-window

In this section we show in Theorem 3.11 that the entropic interpolations for KSP and KFSP enjoy a turnpike property with respect to the Wasserstein distance.

Notice that a turnpike property in the Wasserstein distance could be deduced from the entropic one (cf. Theorem 1.6 and 1.7) by means of the Talagrand inequality (2.2). However, below we provide a different proof that is of independent interest for two reasons. Firstly, the inequality below holds for any $t \in [0, T]$, while the entropic turnpike is restricted to the sub-interval $[\delta, T - \delta]$. Secondly, the argument in the proof, which uses the optimal control formulation of the Schrödinger problem, will be instrumental for the study of the short-time behaviour of the Schrödinger bridge.

Theorem 3.11 (Wasserstein turnpike). *Under hypotheses (H1), (H2) and (H3), there exists a positive constant $C_{d,\alpha,\beta,\gamma}$ such that for any $0 < \delta \leq 1$, as soon as $T > \frac{1}{\kappa} \log C_{d,\alpha,\beta,\gamma} + 2\delta$, for any $t \in [0, T]$ it holds*

$$\mathcal{W}_2(\mu_t^T, \mathbf{m}) \leq C_{d,\alpha,\beta,\gamma} \delta^{-\frac{3}{2}} e^{-\kappa [t \wedge (T-t)]} \sqrt{\mathcal{C}_T(\mu, \nu)}.$$

Proof. Let us firstly assume (H4). We will prove our result for the distorted Wasserstein distance $\mathcal{W}_{M,2}$ induced by the metrics $|\cdot|_M$. Fix $\delta \in (0, 1)$ and assume $t \in [0, T - \delta]$. Define $\tilde{\mu}_t^T$ as the marginal flow generated by the uncontrolled process

$Z_s^{0,T} := (X_s^{0,T}, V_s^{0,T})_{s \in [0,T]}$ solution of (1.1) started at the initial distribution $\mu_0^T \in \mathcal{P}(\mathbb{R}^{2d})$. Then, since $\tilde{\mu}_0^T = \mu_0^T$, it holds

$$\mathcal{W}_{M,2}(\mu_t^T, \mathbf{m}) \leq \mathcal{W}_{M,2}(\mu_t^T, \tilde{\mu}_t^T) + \mathcal{W}_{M,2}(\tilde{\mu}_t^T, \mathbf{m}) \stackrel{(2.12)}{\leq} \mathcal{W}_{M,2}(\mu_t^T, \tilde{\mu}_t^T) + e^{-\kappa t} \mathcal{W}_{M,2}(\mu_0^T, \mathbf{m}), \tag{3.41}$$

The second term in the right hand side can be handled with the Talagrand inequality:

$$\mathcal{W}_{M,2}(\mu_0^T, \mathbf{m}) \lesssim \mathcal{W}_2(\mu_0^T, \mathbf{m}) \stackrel{(2.2)}{\lesssim} \sqrt{\mathcal{H}(\mu_0^T | \mathbf{m})} \leq \sqrt{\mathcal{C}_T(\mu, \nu)},$$

where the last step holds since $\mathcal{C}_T(\mu, \nu) \geq \mathcal{H}((\text{proj}_{x_1})_{\#} \mu^T | (\text{proj}_{x_1})_{\#} \mathbf{R}_{0,T})$.

Let us now focus on $\mathcal{W}_{M,2}(\mu_t^T, \tilde{\mu}_t^T)$. We will use a synchronous coupling between these two measures. Therefore we introduce the process $Z_s^{\mathbf{u},T} := (X_s^{\mathbf{u},T}, V_s^{\mathbf{u},T}) \sim \mu_s^T$, i.e. the solution of (1.3) (driven by the same Brownian motion for $Z_s^{0,T}$) when considering the control $\mathbf{u}_s = 2\gamma \nabla_v \log g_s^T(X_s^{\mathbf{u},T}, V_s^{\mathbf{u},T})$. Particularly, from (3.8) it follows that \mathbf{u} is the optimal control and $Z_s^{\mathbf{u},T} \sim \mu_s^T$. For notation's sake set $Z_s^{\Delta,T} := Z_s^{\mathbf{u},T} - Z_s^{0,T}$. Then it holds

$$dZ_s^{\Delta,T} = \left[b(Z_s^{\mathbf{u},T}) - b(Z_s^{0,T}) \right] ds + \begin{pmatrix} 0 \\ \mathbf{u}_s \end{pmatrix} ds,$$

where $b(z)$ denotes the drift of the Langevin dynamics (1.1). By Itô's Formula we obtain

$$\begin{aligned} d|Z_s^{\Delta,T}|_M^2 &= 2MZ_s^{\Delta} \cdot (b(Z_s^{\mathbf{u}}) - b(Z_s^0)) ds + 2MZ_s^{\Delta} \cdot \begin{pmatrix} 0 \\ \mathbf{u}_s \end{pmatrix} ds \\ &= 2 \int_0^1 Z_s^{\Delta,T} \cdot MJ_b(rZ_s^{\mathbf{u},T} + (1-r)Z_s^{0,T}) Z_s^{\Delta,T} dr ds + 2MZ_s^{\Delta,T} \cdot \begin{pmatrix} 0 \\ \mathbf{u}_s \end{pmatrix} ds \\ &\leq -2\kappa |Z_s^{\Delta,T}|_M^2 ds + 2MZ_s^{\Delta,T} \cdot \begin{pmatrix} 0 \\ \mathbf{u}_s \end{pmatrix} ds, \end{aligned}$$

where the last inequality follows from (2.9). By taking the expectation, and applying Hölder's inequality we get

$$\frac{d}{ds} \mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right] \leq -2\kappa \mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right] + 2\mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right]^{\frac{1}{2}} \mathbb{E}_{\mathbb{R}} \left[|(0, \mathbf{u}_s)^T|_M^2 \right]^{\frac{1}{2}}.$$

Therefore it holds

$$\frac{d}{ds} \sqrt{\mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right]} \leq -\kappa \sqrt{\mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right]} + \mathbb{E}_{\mathbb{R}} \left[|(0, \mathbf{u}_s)^T|_M^2 \right]^{\frac{1}{2}}.$$

Recalling that the optimal control is given by $\mathbf{u}_s = 2\gamma \nabla_v \log g_s^T(X_s^{\mathbf{u},T}, V_s^{\mathbf{u},T})$ we obtain that

$$\frac{d}{ds} \sqrt{\mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right]} \lesssim \left(\int_{\mathbb{R}^{2d}} |\nabla_v \log g_s^T|^2 \rho_s^T dm \right)^{\frac{1}{2}} \lesssim \psi^T(s)^{\frac{1}{2}}.$$

Therefore, by integrating over $s \in [0, t]$ we get

$$\sqrt{\mathbb{E}_{\mathbb{R}} \left[|Z_t^{\Delta,T}|_M^2 \right]} = \int_0^t \frac{d}{ds} \sqrt{\mathbb{E}_{\mathbb{R}} \left[|Z_s^{\Delta,T}|_M^2 \right]} ds \stackrel{(3.23)}{\lesssim} \delta^{-\frac{3}{2}} e^{-\kappa(T-t)} \sqrt{\mathcal{C}_T(\mu, \nu)},$$

and hence it holds

$$\mathcal{W}_{M,2}(\mu_t^T, \tilde{\mu}_t^T) \lesssim \delta^{-\frac{3}{2}} e^{-\kappa(T-t)} \sqrt{\mathcal{C}_T(\mu, \nu)}. \tag{3.42}$$

Then, from (3.41) we deduce that for any $t \in [0, T - \delta]$ it holds

$$\begin{aligned} \mathcal{W}_{M,2}(\mu_t^T, \mathbf{m}) &\lesssim \delta^{-\frac{3}{2}} e^{-\kappa(T-t)} \sqrt{\mathcal{C}_T(\mu, \nu)} + e^{-\kappa t} \sqrt{\mathcal{C}_T(\mu, \nu)} \\ &\lesssim \delta^{-\frac{3}{2}} e^{-\kappa[t \wedge (T-t)]} \sqrt{\mathcal{C}_T(\mu, \nu)}. \end{aligned}$$

By considering the contraction along P^* , the same argument gives us the same bound for $t \in [\delta, T]$ and therefore on the whole domain $[0, T]$.

In order to relax the assumption to (H3), it is enough to consider once again the approximation of the optimizer (as in the proof of Theorem 1.6) together with the lower semicontinuity of the Wasserstein distance. \square

The previous argument can also be applied in order to prove Theorem 1.11.

Proof of Theorem 1.11. At first, let us assume (H4). We have

$$\begin{aligned} \mathcal{W}_2(\mu_t^T, \mu_t^\infty) &\leq \mathcal{W}_2(\mu_t^T, \tilde{\mu}_t^T) + \mathcal{W}_2(\tilde{\mu}_t^T, \mu_t^\infty) \\ &\stackrel{(3.42), (2.12)}{\lesssim} \delta^{-\frac{3}{2}} e^{-\kappa(T-t)} \sqrt{\mathcal{C}_T(\mu, \nu)} + e^{-\kappa t} \mathcal{W}_2(\mu_0^T, \mu \otimes \mathbf{m}_V), \end{aligned} \tag{3.43}$$

where $\tilde{\mu}^T$ is the marginal flow defined in the previous proof, i.e. the flow generated by the uncontrolled process $(X_s^{0,T}, V_s^{0,T})_{s \in [0, T]}$ started at the initial distribution $\mu_0^T \in \mathcal{P}(\mathbb{R}^{2d})$. Using the inequality

$$\mathcal{W}_2(\mu_0^T, \mu \otimes \mathbf{m}_V)^2 \leq \int_{\mathbb{R}^d} \mathcal{W}_2(\mu_0^T(\cdot|x), \mathbf{m}_V)^2 d\mu(x),$$

applying Talagrand’s inequality for \mathbf{m}_V and using the additivity of relative entropy [37, Formula A.8] we obtain

$$\begin{aligned} \mathcal{W}_2(\mu_0^T, \mu \otimes \mathbf{m}_V)^2 &\leq 2 \int_{\mathbb{R}^d} \mathcal{H}(\mu_0^T(\cdot|x)|\mathbf{m}_V) d\mu(x) = 2 \mathcal{H}(\mu_0^T|\mu \otimes \mathbf{m}_V) \\ &= 2 \mathcal{H}(\mu_0^T|\mathbf{m}) - 2 \int_{\mathbb{R}^{2d}} \log \frac{d(\mu \otimes \mathbf{m}_V)}{d\mathbf{m}} d\mu_0^T \\ &= 2 \mathcal{H}(\mu_0^T|\mathbf{m}) - 2 \int_{\mathbb{R}^d} \log \frac{d\mu}{d\mathbf{m}_X} d\mu \\ &= 2 \mathcal{H}(\mu_0^T|\mathbf{m}) - 2 \mathcal{H}(\mu|\mathbf{m}_X) \stackrel{(3.27)}{\lesssim} \delta^{-3} \mathcal{C}_T(\mu, \nu) e^{-2\kappa T}. \end{aligned}$$

By combining the above inequalities with (3.43) we get our result.

The extension of the result to the weaker (H3) follows from the same approximating argument discussed in the previous proof. \square

With a similar reasoning one can prove that the Wasserstein turnpike holds also for KFSP under (FH3). Notice that since in this setting we fix the whole marginals at time 0 and T , it holds $\bar{\mu}_0^T = \bar{\mu}$ and $\bar{\mu}_T^T = \bar{\nu}$ and therefore in this case we do not need a result similar to Theorem 3.8. Therefore we have the following

Theorem 3.12 (Wasserstein turnpike). *Under hypotheses (H1), (H2) and (FH3), there exists a positive constant $C_{d, \alpha, \beta, \gamma}$ such that for any $0 < \delta \leq 1$, as soon as $T > \frac{1}{\kappa} \log C_{d, \alpha, \beta, \gamma} + 2\delta$, for any $t \in [0, T]$ it holds*

$$\mathcal{W}_2(\bar{\mu}_t^T, \mathbf{m}) \leq C_{d, \alpha, \beta, \gamma} \delta^{-\frac{3}{2}} e^{-\kappa[t \wedge (T-t)]} \sqrt{\mathcal{C}_T^F(\bar{\mu}, \bar{\nu})}.$$

4 From compact support to finite entropy

In this section we discuss two types of approximating sequences that we have used in order to extend our main results from (H4) to (H3).

In Section 4.1 we deal with the *Approximation of the optimizer* where we are able to prove the convergence of the entropic cost of the approximated problem to the original entropic cost but not the convergence of the associated entropies of the marginals at time $t = 0, T$. On the other hand in Section 4.2, we investigate the *Approximation of*

the marginals, by approximating directly the marginals and consider the associated Schrödinger problems. In this case, we get the convergence of the marginals' relative entropies, but not the one of the entropic cost. The two aforementioned strategies produce complementary bounds which can be applied together in order to relax the assumptions from (H4) to (H3).

We will deal exclusively with the approximations and proofs for KSP and omit those for KFSP, since the latter can be treated in the same way.

4.1 Approximating the optimizer

Fix a couple of marginals $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ satisfying (H3). We already know that there exists a unique minimizer $\mu^T \in \Pi_X(\mu, \nu)$ for KSP, with $R_{0,T}$ -density given by ρ^T . Now consider an increasing sequence of rectangular compact sets $(K_n)_{n \in \mathbb{N}}$ in \mathbb{R}^{4d} whose union gives the whole space. For each $n \in \mathbb{N}$ define the probability measure q_n^T as the measure whose $R_{0,T}$ -density is given by

$$\hat{\rho}_n^T = \frac{dq_n^T}{dR_{0,T}} := (\rho^T \wedge n) \frac{\mathbb{1}_{K_n}}{C_n},$$

where $C_n := \int_{K_n} (\rho^T \wedge n) dR_{0,T}$ is the normalising constant. Notice that $C_n \uparrow 1$ by monotone convergence and $\hat{\rho}_n^T \rightarrow \rho^T$. For convenience, we fix in this section some $\bar{n} \in \mathbb{N}$ such that $C_n \geq 1/2$ for any $n \geq \bar{n}$.

Lemma 4.1. *The following properties hold true.*

- (i) The marginals $\mu^n := (\text{proj}_{x_1})_{\#} q_n^T$ and $\nu^n := (\text{proj}_{x_2})_{\#} q_n^T$ satisfy (H4).
- (ii) $q_n^T \rightharpoonup \mu^T$.
- (iii) $\mathcal{H}(q_n^T | R_{0,T}) \rightarrow \mathcal{H}(\mu^T | R_{0,T}) = \mathcal{C}_T(\mu, \nu)$.

Proof. We start with i). Since q_n^T has compact support, so do its marginals μ^n, ν^n . Moreover if $B \subseteq \mathbb{R}^d$ is a Borel set, then

$$\mu^n(B) = q_n^T(B \times \mathbb{R}^{3d}) \leq \frac{n}{C_n} \int_{B \times \mathbb{R}^{3d}} dR_{0,T} = \frac{n}{C_n} R_{0,T}(B \times \mathbb{R}^{3d}) = \frac{n}{C_n} m_X(B)$$

and therefore $\|\text{d}\mu^n / \text{d}m_X\|_{L^\infty(m_X)} \leq \frac{n}{C_n}$. The same reasoning applies also to ν^n .

The weak convergence in (ii) follows from dominated convergence.

Let us prove point (iii). Notice that for each $n \geq \bar{n}$ it holds

$$|\hat{\rho}_n^T \log \hat{\rho}_n^T| \leq \max \{e^{-1}, (\rho^T C_n^{-1}) \log(\rho^T C_n^{-1})\},$$

and the above RHS is $R_{0,T}$ -integrable since it holds

$$\int_{\mathbb{R}^{4d}} (\rho^T C_n^{-1}) \log(\rho^T C_n^{-1}) dR_{0,T} = \frac{1}{C_n} \mathcal{C}_T(\mu, \nu) + \frac{1}{C_n} \log \left(\frac{1}{C_n} \right) < \infty,$$

which is finite under (H3). From the Dominated Convergence Theorem we get (iii). \square

Proposition 4.2. *Assume (H1) and (H3) to be true for $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. Let μ^T be the unique minimizer in KSP with marginals μ, ν . Suppose we are given a sequence $(q_n^T)_{n \in \mathbb{N}} \subset \mathcal{P}(\mathbb{R}^{4d})$ such that*

- (i) $q_n^T \rightharpoonup \mu^T$,
- (ii) $\mathcal{H}(q_n^T | R_{0,T}) \rightarrow \mathcal{H}(\mu^T | R_{0,T})$.

Moreover for each $n \in \mathbb{N}$ define the marginals $\mu^n := (\text{proj}_{x_1})_{\#} q_n^T$ and $\nu^n := (\text{proj}_{x_2})_{\#} q_n^T$. Then, for each $n \in \mathbb{N}$, there exists a unique minimizer $\mu_n^T \in \Pi_X(\mu^n, \nu^n)$ in KSP with marginals μ^n, ν^n . Moreover it holds

$$\mu_n^T \rightharpoonup \mu^T \quad \text{and} \quad \mathcal{C}_T(\mu^n, \nu^n) \rightharpoonup \mathcal{C}_T(\mu, \nu).$$

Proof. Firstly, (H3) and the convergence of the entropies in the assumptions imply that $\mathcal{H}(\mu^n | \mathfrak{m}_X), \mathcal{H}(\nu^n | \mathfrak{m}_X) \leq \mathcal{H}(q_n^T | \mathbb{R}_{0,T}) < C$ for some positive constant C , uniformly in $n \in \mathbb{N}$. Hence (H3) holds also for μ^n, ν^n . This gives the existence and uniqueness of the minimizer in KSP with marginals μ^n and ν^n for each $n \in \mathbb{N}$.

Then, from (3.4) we deduce

$$\sup_{n \in \mathbb{N}} \mathcal{H}(\mu_n^T | \mathbb{R}_{0,T}) = \sup_{n \in \mathbb{N}} \mathcal{C}_T(\mu^n, \nu^n) \lesssim 1 + \sup_{n \in \mathbb{N}} [\mathcal{H}(\mu^n | \mathfrak{m}_X) + \mathcal{H}(\nu^n | \mathfrak{m}_X)] \lesssim 1 + 2C.$$

Since the relative entropy $\mathcal{H}(\cdot | \mathbb{R}_{0,T})$ has compact level sets [24, Lemma 1.4.3], there is a subsequence $(\mu_{n_k}^T)_{k \in \mathbb{N}}$ and a probability measure $\tilde{\mu}^T \in \mathcal{P}(\mathbb{R}^{4d})$ such that $\mu_{n_k}^T \rightharpoonup \tilde{\mu}^T$ weakly. Moreover, from the lower semicontinuity of $\mathcal{H}(\cdot | \mathbb{R}_{0,T})$ and the optimality of $\mu_{n_k}^T$ we get

$$\mathcal{H}(\tilde{\mu}^T | \mathbb{R}_{0,T}) \leq \liminf_{k \rightarrow \infty} \mathcal{H}(\mu_{n_k}^T | \mathbb{R}_{0,T}) \leq \liminf_{k \rightarrow \infty} \mathcal{H}(q_{n_k}^T | \mathbb{R}_{0,T}) = \mathcal{H}(\mu^T | \mathbb{R}_{0,T}). \quad (4.1)$$

Now, we claim that $\tilde{\mu}^T \in \Pi_X(\mu, \nu)$. Indeed we have for any $i = 1, 2$

$$(\text{proj}_{x_i})_{\#} \tilde{\mu}^T = \lim_{k \rightarrow \infty} (\text{proj}_{x_i})_{\#} \mu_{n_k}^T = \lim_{k \rightarrow \infty} (\text{proj}_{x_i})_{\#} q_{n_k}^T = (\text{proj}_{x_i})_{\#} \mu^T = \begin{cases} \mu & i = 1 \\ \nu & i = 2, \end{cases}$$

where the second equality holds because $\mu_{n_k}^T$ and $q_{n_k}^T$ share the same marginals, while the third follows from our hypotheses. Therefore, from the bound (4.1) and the optimality of μ^T as unique minimizer in $\Pi_X(\mu, \nu)$ for KSP, it follows $\tilde{\mu}^T = \mu^T$.

Hence, as $k \rightarrow \infty$, it holds $\mu_{n_k}^T \rightharpoonup \mu^T$ and

$$\exists \lim_{k \rightarrow \infty} \mathcal{C}_T(\mu^{n_k}, \nu^{n_k}) = \lim_{k \rightarrow \infty} \mathcal{H}(\mu_{n_k}^T | \mathbb{R}_{0,T}) = \mathcal{H}(\mu^T | \mathbb{R}_{0,T}) = \mathcal{C}_T(\mu, \nu).$$

Since in both the limits above the limit objects do not depend on the subsequence and since the weak convergence is metrizable, we get the desired thesis. \square

Corollary 4.3. *Under the same setting of the previous proposition, if $\mu^{n,T} \in \mathcal{P}(\Omega)$ denotes the minimizer in KSPd, then for each $t \in [0, T]$*

$$\mu^{n,T} \rightharpoonup \mu^T \quad \text{and} \quad \mu_t^{n,T} \rightharpoonup \mu_t^T.$$

Proof. From the relation between KSPd and KSP, for any $\phi \in C_b(\Omega)$ we have

$$\int_{\Omega} \phi d\mu^{n,T} = \int_{\mathbb{R}^{4d}} \left(\int_{\Omega} \phi dR^{x,v,y,w} \right) d\mu_n^T \rightarrow \int_{\mathbb{R}^{4d}} \left(\int_{\Omega} \phi dR^{x,v,y,w} \right) d\mu^T = \int_{\Omega} \phi d\mu^T,$$

where $R^{x,v,y,w}$ denotes the bridge of the reference measure. Let us just justify the middle step. Since ϕ is bounded, so does $\int_{\Omega} \phi dR^{x,v,y,w}$. Moreover since the bridge $R^{x,v,y,w}$ is weakly continuous with respect to its extremes [14, Corollary 1], from the continuity of ϕ , it follows the continuity of the function $(x, v, y, w) \mapsto \int_{\Omega} \phi dR^{x,v,y,w}$. Hence the above function is bounded and continuous on \mathbb{R}^{4d} and from the weak convergence $\mu_n^T \rightharpoonup \mu^T$ it follows $\mu^{n,T} \rightharpoonup \mu^T$. The other limit follows by taking the time marginals of $\mu^{n,T}$. \square

4.2 Approximating the marginals

In this section we are going to perform the approximating arguments directly on the fixed marginals. This will not lead to the convergence of the respective kinetic entropic costs, nevertheless it will be useful in proving the bounds where the previous approximating argument fails. The idea is similar to the one performed previously: consider an increasing sequence of compact sets $(K_n)_{n \in \mathbb{N}}$ in \mathbb{R}^d whose union gives the whole space and for any $q \in \mathcal{P}(\mathbb{R}^d)$, satisfying (H3) and $n \in \mathbb{N}$ large enough so that $q(K_n) > 0$, define the probability measure q_n^M as the measure whose m_X -density is given by

$$\frac{dq_n^M}{dm_X} := \left(\frac{dq}{dm_X} \wedge n \right) \frac{\mathbb{1}_{K_n}}{C_n^q},$$

where $C_n^q := \int_{K_n} (\frac{dq}{dm_X} \wedge n) dm_X \geq 0$ is the normalising constant. Note that monotone convergence yields $C_n^q \uparrow 1$. Then it follows that q_n^M satisfies (H4), $q_n^M \rightharpoonup q$ and by mimicking the argument performed in the Lemma 4.1 it follows that

$$\mathcal{H}(q_n^M | m_X) \xrightarrow{n \rightarrow \infty} \mathcal{H}(q | m_X). \tag{4.2}$$

Lemma 4.4. Fix $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ satisfying (H3). Then, up to restricting ourselves to a subsequence, it holds

$$\mathcal{C}_T(\mu, \nu) \leq \liminf_{n \rightarrow \infty} \mathcal{C}_T(\mu_n^M, \nu_n^M).$$

Proof. Let $\mu_{M,n}^T$ denotes the optimizer for $\mathcal{C}_T(\mu_n^M, \nu_n^M)$. Then we have

$$\begin{aligned} \mathcal{H}(\mu_{M,n}^T | R_{0,T}) &= \mathcal{C}_T(\mu_n^M, \nu_n^M) \\ &\stackrel{(3.4)}{\lesssim} 1 + \mathcal{H}(\mu_n^M | m_X) + \mathcal{H}(\nu_n^M | m_X) \xrightarrow{n \rightarrow \infty} 1 + \mathcal{H}(\mu | m_X) + \mathcal{H}(\nu | m_X), \end{aligned}$$

which is finite because of (H3). Since $\mathcal{H}(\cdot | R_{0,T})$ has compact level set, we know that there exists $\mu^* \in \mathcal{P}(\mathbb{R}^{4d})$ such that $\mu_{M,n}^T \rightharpoonup \mu^*$, up to considering a subsequence. We claim that $\mu^* \in \Pi_X(\mu, \nu)$. Indeed we have $(X_0)_\# \mu_{M,n}^T \rightharpoonup (X_0)_\# \mu^*$ but $(X_T)_\# \mu_{M,n}^T = \mu_n^M \rightharpoonup \mu$ and hence $(X_0)_\# \mu^* = \mu$. Similarly it holds $(X_T)_\# \mu^* = \nu$. Therefore we have $\mathcal{C}_T(\mu, \nu) \leq \mathcal{H}(\mu^* | R_{0,T})$ and from the lower semicontinuity of $\mathcal{H}(\cdot | R_{0,T})$ we deduce our thesis. \square

Appendix

Proof of Lemma 2.1

Let $T_0 > 0$ be fixed. From Jensen’s inequality we know that

$$\begin{aligned} \log p_T((x, v), (y, w)) &= \log \int_{\mathbb{R}^{2d}} p_{T-T_0/2}((x, v), (z, u)) p_{T_0/2}((z, u), (y, w)) dm(z, u) \\ &\geq \int_{\mathbb{R}^{2d}} \log p_{T-T_0/2}((x, v), (z, u)) dm(z, u) + \int_{\mathbb{R}^{2d}} \log p_{T_0/2}((z, u), (y, w)) dm(z, u). \end{aligned} \tag{.3}$$

By Theorem 1.1 in [21], there exists $C \geq 1$ depending on T_0 such that

$$p_{T_0/2}((z, u), (y, w)) \gtrsim p_{T_0/2}((z, u), (y, w)) m(y, w) \geq C^{-1} e^{-C|\theta_{T_0/2}(z,u)-(y,w)^T|^2}$$

where $\theta_t(x_0, v_0) = (\theta_t^x, \theta_t^v)^T$ denotes the solution of the denoised Langevin ODE system

$$\begin{cases} \frac{d}{dt} \theta_t^x = \theta_t^v \\ \frac{d}{dt} \theta_t^v = -\theta_t^v - \nabla U(\theta_t^x) \end{cases} \quad \text{with } \theta_0 = (x_0, v_0)^T.$$

Since under (H1) there exists a large enough positive $r \in \mathbb{R}$ such that (2.9) holds for $(\text{Id}, -r)$, from [39, Theorem 1] (with $\Sigma = 0$) it follows

$$|\theta_t(y, w) - \theta_t(0, 0)| \leq e^{rt} |(y, w)^T|, \quad \forall (y, w)^T \in \mathbb{R}^{2d}, \forall t \geq 0. \quad (.4)$$

Therefore, up to changing the constants C from line to line, we have

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \log p_{T_0/2}((z, u), (y, w)) \, \text{d}\mathbf{m}(z, u) &\geq \log C^{-1} - C \int_{\mathbb{R}^{2d}} |\theta_{T_0/2}(z, u) - (y, w)^T|^2 \, \text{d}\mathbf{m}(z, u) \\ &\geq -C \left(1 + |y|^2 + |w|^2 + \int_{\mathbb{R}^{2d}} |\theta_{T_0/2}(z, u)|^2 \, \text{d}\mathbf{m}(z, u) \right) \geq -C \left(1 + |y|^2 + |w|^2 \right), \end{aligned} \quad (.5)$$

where the last step holds since $\mathbf{m} \in \mathcal{P}_2(\mathbb{R}^{2d})$, and therefore

$$\begin{aligned} \int_{\mathbb{R}^{2d}} |\theta_{T_0/2}(z, u)|^2 \, \text{d}\mathbf{m}(z, u) &\leq 2 |\theta_{T_0/2}(0, 0)|^2 + 2 \int_{\mathbb{R}^{2d}} |\theta_{T_0/2}(z, u) - \theta_{T_0/2}(0, 0)|^2 \, \text{d}\mathbf{m}(z, u) \\ &\stackrel{(.4)}{\leq} 2 |\theta_{T_0/2}(0, 0)|^2 + 2 e^{2r} \int_{\mathbb{R}^{2d}} (|z|^2 + |u|^2) \, \text{d}\mathbf{m}(z, u) \leq C. \end{aligned}$$

Now, notice that we can rewrite the first integral of the RHS in (.3) as

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \log p_{T-T_0/2}((x, v), (z, u)) \, \text{d}\mathbf{m}(z, u) \\ = \int_{\mathbb{R}^{2d}} \log \left[\int_{\mathbb{R}^{2d}} p_{T_0/2}((x, v), (q, r)) p_{T-T_0}((q, r), (z, u)) \, \text{d}\mathbf{m}(q, r) \right] \, \text{d}\mathbf{m}(z, u). \end{aligned}$$

Because of (2.6), we know that $p_{T-T_0/2}((q, r), (z, u)) \, \text{d}\mathbf{m}(q, r)$ is a probability measure over \mathbb{R}^{2d} and therefore by Jensen's inequality and Fubini the above displacement can be lower bounded by

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \log [p_{T_0/2}((x, v), (q, r))] p_{T-T_0}((q, r), (z, u)) \, \text{d}\mathbf{m}(q, r) \, \text{d}\mathbf{m}(z, u) \\ = \int_{\mathbb{R}^{2d}} \log p_{T_0/2}((x, v), (q, r)) \, \text{d}\mathbf{m}(q, r) \stackrel{(.6)}{=} \int_{\mathbb{R}^{2d}} \log p_{T_0/2}((q, -r), (x, -v)) \, \text{d}\mathbf{m}(q, r) \\ = \int_{\mathbb{R}^{2d}} \log p_{T_0/2}((q, r), (x, -v)) \, \text{d}\mathbf{m}(q, r) \stackrel{(.5)}{\geq} -C \left(1 + |x|^2 + |v|^2 \right). \end{aligned}$$

Putting the above lower bound and (.5) into inequality (.3), we get

$$\log p_T((x, v), (y, w)) \geq -c_{T_0} \left(1 + |x|^2 + |v|^2 + |y|^2 + |w|^2 \right).$$

References

- [1] J. Backhoff, G. Conforti, I. Gentil, and C. Léonard, *The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities*, Probab. Theory Related Fields **178** (2020), no. 1-2, 475–530. MR4146543
- [2] D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206. MR889476
- [3] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR3155209
- [4] F. Baudoin, *Bakry-Émery meet Villani*, J. Funct. Anal. **273** (2017), no. 7, 2275–2291. MR3677826
- [5] Y. Benamou, J.-D. and Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393. MR1738163

- [6] F. Bolley, A. Guillin, and F. Malrieu, *Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation*, M2AN Math. Model. Numer. Anal. **44** (2010), no. 5, 867–884. MR2731396
- [7] A. Braides, *A handbook on Γ -convergence*, Handbook of Differential Equations: Stationary Partial Differential Equations, vol. 3, Elsevier, 2006, pp. 101–213.
- [8] T. Breiten and L. Pfeiffer, *On the turnpike property and the receding-horizon method for linear-quadratic optimal control problems*, SIAM J. Control Optim. **58** (2020), no. 2, 1077–1102. MR4082481
- [9] E. Camrud, D. P. Herzog, G. Stoltz, and M. Gordina, *Weighted L^2 -contractivity of Langevin dynamics with singular potentials*, Nonlinearity **35** (2022), no. 2, 998–1035. MR4373993
- [10] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta, *Long time average of mean field games*, Netw. Heterog. Media **7** (2012), no. 2, 279–301. MR2928380
- [11] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta, *Long time average of mean field games with a nonlocal coupling*, SIAM J. Control Optim. **51** (2013), no. 5, 3558–3591. MR3103242
- [12] P. Cardaliaguet and A. Porretta, *Long time behavior of the master equation in mean field game theory*, Anal. PDE **12** (2019), no. 6, 1397–1453. MR3921309
- [13] P. Cattiaux, A. Guillin, P. Monmarché, and C. Zhang, *Entropic multipliers method for Langevin diffusion and weighted log Sobolev inequalities*, J. Funct. Anal. **277** (2019), no. 11, 108288, 24. MR4013832
- [14] L. Chaumont and G. Uribe Bravo, *Markovian bridges: weak continuity and pathwise constructions*, Ann. Probab. **39** (2011), no. 2, 609–647. MR2789508
- [15] Y. Chen, T. T. Georgiou, and M. Pavon, *Fast cooling for a system of stochastic oscillators*, J. Math. Phys. **56** (2015), no. 11, 113302, 17. MR3422383
- [16] Y. Chen, T. T. Georgiou, and M. Pavon, *Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge*, SIAM Rev. **63** (2021), no. 2, 249–313. MR4253788
- [17] G. Clerc, G. Conforti, and I. Gentil, *Long-time behaviour of entropic interpolations*, arXiv:2007.07594
- [18] G. Conforti, *A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost*, Probab. Theory Related Fields **174** (2019), no. 1-2, 1–47. MR3947319
- [19] G. Conforti and L. Tamanini, *A formula for the time derivative of the entropic cost and applications*, J. Funct. Anal. **280** (2021), no. 11, Paper No. 108964, 48. MR4232667
- [20] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, second ed., Encyclopedia of Mathematics and its Applications, vol. 152, Cambridge University Press, Cambridge, 2014. MR3236753
- [21] F. Delarue and S. Menozzi, *Density estimates for a random noise propagating through a chain of differential equations*, J. Funct. Anal. **259** (2010), no. 6, 1577–1630. MR2659772
- [22] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010, Corrected reprint of the second (1998) edition. MR2571413
- [23] J. Dolbeault, C. Mouhot, and C. Schmeiser, *Hypocoercivity for linear kinetic equations conserving mass*, Trans. Amer. Math. Soc. **367** (2015), no. 6, 3807–3828. MR3324910
- [24] P. Dupuis and R. S. Ellis, *A weak convergence approach to the theory of large deviations*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1997, A Wiley-Interscience Publication. MR1431744
- [25] A. Eberle, A. Guillin, and R. Zimmer, *Couplings and quantitative contraction rates for Langevin dynamics*, Ann. Probab. **47** (2019), no. 4, 1982–2010. MR3980913
- [26] H. Föllmer, *Random fields and diffusion processes*, École d’Été de Probabilités de Saint-Flour XV–XVII, 1985–87, Lecture Notes in Math., vol. 1362, Springer, Berlin, 1988, pp. 101–203. MR983373
- [27] N. Gigli and L. Tamanini, *Benamou-Brenier and duality formulas for the entropic cost on $RCD^*(K, N)$ spaces*, Probab. Theory Related Fields **176** (2020), no. 1-2, 1–34. MR4055185

- [28] N. Gigli and L. Tamanini, *Second order differentiation formula on $RCD^*(K, N)$ spaces*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 5, 1727–1795. MR4244516
- [29] A. Guillin, W. Liu, L. Wu, and C. Zhang, *The kinetic Fokker-Planck equation with mean field interaction*, J. Math. Pures Appl. (9) **150** (2021), 1–23. MR4248461
- [30] A. Guillin and F.-Y. Wang, *Degenerate Fokker-Planck equations: Bismut formula, gradient estimate and Harnack inequality*, J. Differential Equations **253** (2012), no. 1, 20–40. MR2917400
- [31] F. Hérau and F. Nier, *Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential*, Arch. Ration. Mech. Anal. **171** (2004), no. 2, 151–218. MR2034753
- [32] D. P. Herzog and J. C. Mattingly, *Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials*, Comm. Pure Appl. Math. **72** (2019), no. 10, 2231–2255. MR3998640
- [33] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR0222474
- [34] C. Léonard, *Minimization of energy functionals applied to some inverse problems*, Appl. Math. Optim. **44** (2001), no. 3, 273–297. MR1851741
- [35] C. Léonard, *From the Schrödinger problem to the Monge-Kantorovich problem*, J. Funct. Anal. **262** (2012), no. 4, 1879–1920. MR2873864
- [36] C. Léonard, *Girsanov theory under a finite entropy condition*, Séminaire de Probabilités XLIV, Lecture Notes in Math., vol. 2046, Springer, Heidelberg, 2012, pp. 429–465. MR2953359
- [37] C. Léonard, *A survey of the Schrödinger problem and some of its connections with optimal transport*, Discrete Contin. Dyn. Syst. **34** (2014), no. 4, 1533–1574. MR3121631
- [38] T. Mikami, *Monge’s problem with a quadratic cost by the zero-noise limit of h -path processes*, Probab. Theory Related Fields **129** (2004), no. 2, 245–260. MR2063377
- [39] P. Monmarché, *Almost sure contraction for diffusions on \mathbb{R}^d . Application to generalised Langevin diffusions*, arXiv:2009.10828
- [40] S. Pal and T.-K. L. Wong, *Multiplicative Schrödinger problem and the Dirichlet transport*, Probab. Theory Related Fields **178** (2020), no. 1-2, 613–654. MR4146546
- [41] G. Peyré, M. Cuturi, et al., *Computational optimal transport: With applications to data science*, Foundations and Trends® in Machine Learning **11** (2019), no. 5-6, 355–607, arXiv:1803.00567
- [42] L. Rüschemdorf and W. Thomsen, *Note on the Schrödinger equation and I -projections*, Statist. Probab. Lett. **17** (1993), no. 5, 369–375. MR1237783
- [43] E. Schrödinger, *Über die Umkehrung der Naturgesetze.*, Sitzungsberichte Preuss. Akad. Wiss. Berlin. Phys. Math. **144** (1931), 144–153.
- [44] E. Schrödinger, *Sur la théorie relativiste de l’électron et l’interprétation de la mécanique quantique*, Ann. Inst. H. Poincaré **2** (1932), no. 4, 269–310. MR1508000
- [45] D. Talay, *Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme*, Markov Process. Related Fields **8** (2002), no. 2, 163–198, Inhomogeneous random systems (Cergy-Pontoise, 2001). MR1924934
- [46] E. Trélat, *Linear turnpike theorem*, arXiv:2010.13605
- [47] E. Trélat and C. Zhang, *Integral and measure-turnpike properties for infinite-dimensional optimal control systems*, Math. Control Signals Systems **30** (2018), no. 1, Art. 3, 34. MR3780457
- [48] E. Trélat, C. Zhang, and E. Zuazua, *Steady-state and periodic exponential turnpike property for optimal control problems in Hilbert spaces*, SIAM J. Control Optim. **56** (2018), no. 2, 1222–1252. MR3780737
- [49] E. Trélat and E. Zuazua, *The turnpike property in finite-dimensional nonlinear optimal control*, J. Differential Equations **258** (2015), no. 1, 81–114. MR3271298
- [50] C. Villani, *Hypocoercivity*, Mem. Amer. Math. Soc. **202** (2009), no. 950, iv+141. MR2562709
- [51] A. J. Zaslavski, *Turnpike properties in the calculus of variations and optimal control*, Nonconvex Optimization and its Applications, vol. 80, Springer, New York, 2006. MR2164615
- [52] A. J. Zaslavski, *Turnpike conditions in infinite dimensional optimal control*, Springer Optimization and Its Applications, vol. 148, Springer, Cham, 2019. MR3966463