# Turnpike-type chaoticity for mean-field Schrödinger problems

Songbo Wang and Wenbin Yan

*CEREMADE, Université Paris-Dauphine, Université PSL, Paris, France*

**Abstract**

We study the long-time behavior of optimal particle systems for mean-field Schrödinger problems. Using first-order functional inequalities, we establish turnpike-type chaoticity: the influence of endpoint correlations decays exponentially away from both endpoints, up to finite-particle errors. These estimates yield uniform-in-time propagation of chaos for chaotic endpoints and generation of chaos in the interior. The framework accommodates flat and displacement convexity, as well as mild flat-semiconvexity. For an equilibrium reference, we use the free energy structure induced by an ergodic corrector to obtain near-optimal relaxation rates.

## Contents

**1 Introduction** **2**

**2 Main results** **6**
2.1 Schrödinger problems . . . . . 6
2.2 Relative formulations . . . . . 7
2.3 Turnpike-type chaoticity . . . . . 9
2.4 Transfer of functional inequalities . . . . . 12

**3 Examples** **15**
3.1 Mildly flat-semiconvex costs on the torus . . . . . 15
3.2 Displacement convex costs on Euclidean space . . . . . 17

**4 Relative formulations** **18**

**5 Entropic turnpike** **20**
5.1 Connection via accelerated flows . . . . . 20
5.2 Action lower bounds . . . . . 25
5.3 Discrete comparison . . . . . 26

*Date*: November 15, 2026.

*Key words and phrases.* Schrödinger problem, mean-field interaction, turnpike property, propagation of chaos, modulated free energy, functional inequality.

**6 Free-energetic turnpike** **28**
6.1 Free energy structure . . . . . 28
6.2 Connection via gradient flows . . . . . 31
6.3 Continuous comparison . . . . . 33

**7 Transfer of functional inequalities** **37**

**A Mean-field estimates on the torus** **42**

**B Estimates for displacement convex costs** **53**

**C Finite-entropy diffusions** **56**
C.1 Brownian tilts . . . . . 57
C.2 Drift perturbations . . . . . 59
C.3 Relative velocities and entropy . . . . . 63

# 1 Introduction

Schrödinger [43, 42] asked how a large system of independent diffusive particles is most likely to evolve between two prescribed distributions at times 0 and $T > 0$. To formulate this problem, let $\mathcal{X} = \mathbb{T}^d$ or $\mathbb{R}^d$, let $m_0$ and $m_T$ be probability measures on $\mathcal{X}$, and let $R$ be the law of Brownian motion on $[0, T]$ with generator $\Delta$ and initial distribution $m_0$. Sanov's theorem characterizes the most likely evolution as a minimizer of the relative entropy

$$\mathcal{H}(P|R) = \int \log \frac{\mathrm{d}P}{\mathrm{d}R} \, \mathrm{d}P$$

over path laws $P$ with marginals $P_0 = m_0$ and $P_T = m_T$; see Léonard [36] for a detailed survey. Adding a running cost $F_t$ on probability measures yields the *mean-field Schrödinger problem*:

$$\inf_{P_0 = m_0,\, P_T = m_T} \left\{ 2\mathcal{H}(P|R) + \int_0^T F_t(P_t) \, \mathrm{d}t \right\}, \tag{1}$$

where $P_t$ denotes the time-$t$ marginal of $P$. When $P_0 = m_0$ and $\mathcal{H}(P|R) < \infty$, Girsanov's theorem expresses $2\mathcal{H}(P|R)$ as the quadratic drift cost $\frac{1}{2}\,\mathbb{E}\int_0^T |b_t|^2 \, \mathrm{d}t$ for the dynamics $\mathrm{d}X_t = b_t \, \mathrm{d}t + \sqrt{2} \, \mathrm{d}B_t$ [19]. This representation connects (1) to mean-field control [13, Chapter 6] through *semi-duality*: the terminal marginal constraint $P_T = m_T$ replaces the usual terminal cost.

The mean-field Schrödinger problem (1) has a natural $N$-particle counterpart. Write $[N] = \{1, \dots, N\}$, and for a probability measure $\nu$, denote by $\nu^{\otimes N}$ its $N$-fold product. For $\boldsymbol{x} = (x^1, \dots, x^N) \in \mathcal{X}^N$, its empirical measure is

$$\mu_{\boldsymbol{x}} = \frac{1}{N} \sum_{i \in [N]} \delta_{x^i}.$$

Let $m_0^N$ and $m_T^N$ be two probability measures on $\mathcal{X}^N$. The corresponding *$N$-particle Schrödinger problem* is

$$\inf_{P_0^N = m_0^N,\, P_T^N = m_T^N} \left\{ 2\mathcal{H}(P^N|R^{\otimes N}) + N \int_0^T \mathbb{E}_{\boldsymbol{X} \sim P_t^N}[F_t(\mu_{\boldsymbol{X}})] \, \mathrm{d}t \right\}, \tag{2}$$

where the infimum is taken over path laws $P^N$ on $\mathcal{X}^N$ with time-$t$ marginal $P^N_t$. When the endpoint laws $m_0^N$ and $m_T^N$ are chaotic with limiting marginals $m_0$ and $m_T$ as $N$ tends to infinity, optimizers of (2) are expected to approach independent copies of an optimizer of (1). This convergence is an optimization analogue of propagation of chaos, a notion going back to Kac [34] and Sznitman [45]. Whereas classical propagation of chaos concerns dynamics with a single prescribed datum at $t = 0$, (2) imposes constraints at both $t = 0$ and $t = T$, so chaoticity propagates from both endpoints into the interior of $[0, T]$.

For a fixed time horizon $T$, this propagation of chaos problem is closely related to the convergence of $N$-particle control problems to mean-field control. Indeed, the drift representation above also expresses (2) as a cooperative $N$-particle control problem with a terminal marginal constraint. Using this formulation, Hernández and Tangpi [31] established propagation of chaos for (2) by penalizing the terminal constraint; when $X^i_T \sim m_T$ is imposed on each particle marginal, their recent work [30] proves convergence of the optimal cost in McKean–Vlasov transport. For general mean-field control problems, Lacker [35] established qualitative limit theorems using relaxed controls. Quantitative rates on a fixed horizon $[0, T]$ have been obtained using smoothness of the limiting value function [28, 12, 8], while Daudin, Delarue and Jackson [26] obtained value-function convergence rates without requiring uniqueness or stability of the limiting optimizers. More recently, for smooth terminal costs and independent initial data, Arnese and Lacker [1] established sharp local chaoticity estimates on fixed horizons.

A complementary question concerns the behavior of optimal trajectories as $T \to \infty$. In optimal control problems with a dissipative structure, optimal trajectories typically stay near a steady state over most of $[0, T]$ and depart from it only near $t = 0$ and $t = T$. This phenomenon is termed the *turnpike property*, established for finite-dimensional nonlinear control by Trélat and Zuazua [46]. Turnpike asymptotics for linear and generalized Schrödinger bridges were derived in [23, 25, 22]. For the mean-field Schrödinger problem (1) with pairwise interaction potentials, Backhoff, Conforti, Gentil and Léonard [3] established exponential turnpike convergence in a strongly convex framework. Beyond Schrödinger bridges, closely related long-time results have been obtained in mean-field control and mean-field games, which share the interior optimality equations of (1). Indeed, a smooth optimizer of (1) with marginals $m_t$ and drift $-\nabla u_t$ satisfies, at least formally, the forward-backward Hamilton–Jacobi–Bellman and Fokker–Planck system

$$\begin{aligned} -\partial_t u_t &= \Delta u_t - \frac{1}{2}|\nabla u_t|^2 + \delta F_t(m_t, \cdot), \\ \partial_t m_t &= \Delta m_t + \nabla \cdot (m_t \nabla u_t), \end{aligned}$$

coupled by $m_0$ and $m_T$, where $\delta F_t$ is the flat derivative of $F_t$ in the measure argument; mean-field control problems and potential mean-field games are governed by the same system up to the boundary conditions at $t = 0$ and $t = T$. In that setting, Cardaliaguet, Lasry, Lions and Porretta [9, 10] and Porretta [40] proved long-time convergence and exponential turnpike estimates for monotone mean-field game systems by duality and Riccati stabilization, and Cardaliaguet and Porretta [11] analyzed the ergodic limit through the master equation. Related long-time results use diffusive stabilization under mild non-monotonicity [21] or couplings of controlled processes [24, 15].

However, to the best of our knowledge, the long-time behavior of $N$-particle Schrödinger or mean-field control problems with general interactions has not yet

been studied in the literature. Investigating the former is the aim of this work. Directly combining the two types of estimates above fails to give satisfactory control of the particle system on long horizons: fixed-horizon chaos bounds deteriorate as $T$ grows, whereas continuum turnpike bounds apply only in the limit $N \to \infty$ and do not capture finite-$N$ correlations. Moreover, the joint constraints $P_0^N = m_0^N$ and $P_T^N = m_T^N$ couple the particles through both endpoints, so methods for initial-value McKean–Vlasov problems [18, 48] do not apply directly.

In this paper, we establish quantitative long-time estimates comparing an optimal $N$-particle bridge directly with a mean-field reference across $[0, T]$. To give a preview of our main results, write $m_t = P_{\mathrm{ref},t}$ for a reference law $P_{\mathrm{ref}}$ satisfying the first-order condition of (1) and $m_t^N = P_{\mathrm{opt},t}^N$ for an optimizer $P_{\mathrm{opt}}^N$ of (2). Under suitable conditions on $m_t$ and $F_t$, we show that, for all sufficiently large $T$ and all $t \in [0, T]$,

$$\mathcal{H}(m_t^N | m_t^{\otimes N}) \leqslant C\big(e^{-ct}\mathcal{H}(m_0^N | m_0^{\otimes N}) + e^{-c(T-t)}\mathcal{H}(m_T^N | m_T^{\otimes N}) + 1\big), \tag{3}$$

where $C$, $c > 0$ are independent of $N$ and $T$. Here, $\mathcal{H}(m_t^N | m_t^{\otimes N})$ compares $m_t^N$ with the moving product law $m_t^{\otimes N}$. We call this two-sided relaxation *turnpike-type chaoticity*. Uniformly bounded endpoint entropies therefore yield uniformly bounded entropy along the bridge and, for symmetric laws, uniform-in-time propagation of chaos. When the endpoint errors are large, for instance if the $N$ particles are strongly correlated at times 0 and $T$, (3) still gives exponential dissipation away from both boundaries and generates chaos in the interior of $[0, T]$.

Our proof of (3) builds on a variational reformulation of (2) around the mean-field reference. Rather than measuring a candidate path law $P^N$ against independent Brownian motions $R^{\otimes N}$, we take $P_{\mathrm{ref}}^{\otimes N}$ as the reference path law. Using the tilt representation (5) to cancel the linear variation of $F_t$ at $m_t$, in the spirit of the modulated free energy functional [44, 7, 6], we rewrite the objective in (2), up to an additive constant independent of $P^N$, as

$$2\mathcal{H}(P^N | P_{\mathrm{ref}}^{\otimes N}) + N \int_0^T \mathbb{E}_{\boldsymbol{X} \sim P_t^N}[F_t(\mu_{\boldsymbol{X}} | m_t)]\, \mathrm{d}t.$$

Here, $F_t(\nu | m_t)$ is the *modulated interaction*, defined by

$$F_t(\nu | m_t) \coloneqq F_t(\nu) - F_t(m_t) - \langle \delta F_t(m_t, \cdot), \nu - m_t \rangle,$$

where $\langle f, \mu \rangle = \int f \, \mathrm{d}\mu$.

We next pass from path laws $P^N$ to their marginal flows $\nu_t^N = P_t^N$, paralleling the action formulation for generalized Schrödinger bridges [27, 22]. If $v_t$ denotes the current velocity transporting the reference flow $m_t$ via $\partial_t m_t + \nabla \cdot (m_t v_t) = 0$, then the product curve $m_t^{\otimes N}$ moves with velocity $v_t^{\oplus N}$, where $b^{\oplus N}(\boldsymbol{x}) = (b(x^1), \dots, b(x^N))$ for a vector field $b$ on $\mathcal{X}$. Any competing marginal flow $\nu_t^N$ can thus be parameterized by a relative velocity $w_t^N$ on top of this background transport through the continuity equation

$$\partial_t \nu_t^N + \nabla \cdot \big(\nu_t^N (v_t^{\oplus N} + w_t^N)\big) = 0.$$

Our first main result reduces (2) to minimizing the *relative action*

$$\int_0^T \left( \frac{1}{2} \int |w_t^N|^2 \, \mathrm{d}\nu_t^N + \frac{1}{2} \mathcal{I}(\nu_t^N | m_t^{\otimes N}) + N\, \mathbb{E}_{\boldsymbol{X} \sim \nu_t^N}[F_t(\mu_{\boldsymbol{X}} | m_t)] \right) \mathrm{d}t \tag{4}$$

over pairs $(\nu_t^N, w_t^N)$ connecting $m_0^N$ to $m_T^N$, where $\mathcal{I}(\nu|\nu') = \int |\nabla \log(\mathrm{d}\nu/\,\mathrm{d}\nu')|^2 \,\mathrm{d}\nu$ denotes the relative Fisher information.

As in finite-dimensional optimal control, the quadratic cost of the relative velocity in (4) acts as kinetic energy, while

$$\nu^N \mapsto \frac{1}{2}\mathcal{I}(\nu^N|m_t^{\otimes N}) + N\,\mathbb{E}_{\nu^N}[F_t(\mu_{\boldsymbol{X}}|m_t)]$$

serves as an *effective potential*, with mean-field counterpart

$$\nu \mapsto \frac{1}{2}\mathcal{I}(\nu|m_t) + F_t(\nu|m_t).$$

A main contribution of this paper is a control-theoretic comparison argument showing that, combined with a log-Sobolev inequality for $m_t$, coercivity of this effective potential yields exponential dissipation in the interior of $[0,T]$ and hence (3). Our approach uses only first-order derivatives with respect to the measure, in contrast to the $\Gamma_2$-based approach of Backhoff, Conforti, Gentil and Léonard [3]. The coercive lower energy bound can be verified when $F_t$ is flat convex or mildly flat-semiconvex, as well as in displacement convex settings.

While (3) applies along general reference flows, its decay exponent $c$ is generally not sharp. When $F_t = F$ is time-independent and the reference flow is a stationary *equilibrium* $m_t = m_*$, an *ergodic corrector* $\chi$ defines the modulated free energy $\mathcal{F}_*(\nu) = \mathcal{H}(\nu|m_*) + \chi(\nu|m_*)$, for which the corrector equation rewrites the mean-field effective potential as

$$\frac{1}{2}\mathcal{I}(\nu|m_*) + F(\nu|m_*) = \frac{1}{2}\|D\mathcal{F}_*(\nu,\cdot)\|^2_{L^2(\nu)},$$

where $D\mathcal{F}_*(\nu,x) = \nabla_x \delta\mathcal{F}_*(\nu,x)$. Thus, for an equilibrium reference, the effective potential is half the squared Wasserstein gradient norm of the modulated free energy, revealing a gradient flow structure in (4). Passing to the $N$-particle modulated free energy gives the same gradient representation up to additional error terms, allowing us to establish two-sided free-energetic relaxation along $m_t^N$ under a Polyak–Łojasiewicz-type inequality for the gradient flow.

Both this Polyak–Łojasiewicz condition and the $N$-particle coercivity estimate used above take the form of defective particle functional inequalities as in [48], where they were transferred from their corresponding mean-field limits for pairwise interaction potentials. In mean-field Schrödinger and control problems, however, neither the running cost $F$ nor the ergodic corrector $\chi$ is necessarily pairwise; in fact, $\chi$ need not be a pair functional even when $F$ is. Another main contribution is to extend this transfer principle to general nonlinear functionals. Combined with our free-energetic relaxation estimates, this extension yields *near-optimal* turnpike rates, arbitrarily close to the mean-field free energy dissipation rate as $N \to \infty$.

Finally, through the semi-duality between (1) and mean-field control, the variational approach developed here also applies to mean-field control problems without common noise. Yet this control-theoretic argument does not extend to the case of nonzero common noise, in part because common noise breaks the time-reversal symmetry underlying (4). We treat mean-field control with common noise in a companion work [49] by a different method, a key step of which is to identify a purely one-sided dissipative structure in place of the two-sided relaxation mechanism used here.

**Organization.** The rest of the paper is organized as follows. Section 2 presents our main results:

(i) two relative formulations of the $N$-particle Schrödinger problem;
(ii) the entropic turnpike estimate along a general reference flow;
(iii) the free-energetic turnpike estimate for an equilibrium reference;
(iv) transfer of functional inequalities from the mean-field limit to $N$-particle laws.

Section 3 illustrates these results on two classes of models: mildly flat-semiconvex running costs on $\mathbb{T}^d$ and displacement convex running costs on $\mathbb{R}^d$. Sections 4–7 prove results (i)–(iv), respectively. Appendix A constructs regular stationary correctors and establishes mean-field bridge estimates on $\mathbb{T}^d$. Appendix B proves equilibrium curvature and coercivity estimates for displacement convex costs on $\mathbb{R}^d$, and Appendix C collects lemmas on finite-entropy diffusions and their marginal flows.

**Acknowledgements.** S. W. would like to thank Pierre Cardaliaguet and Zhenjie Ren for stimulating discussions in 2024 on quantitative large deviation estimates for McKean–Vlasov systems, which planted the seeds for the present work.

**AI use.** The main ideas of this work were developed by the authors in March–April 2026 without AI assistance. The authors directed AI to generate an initial draft, which they then substantially revised. Finally, they used AI to assist with correctness checks and language polishing. They reviewed the manuscript in its entirety and take full responsibility for all mathematical content.

**Funding.** Both authors acknowledge funding provided by the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (ELISA project, grant agreement No. 101054746). This funding supported W. Y.'s visit to Nice at the invitation of S. W. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

S. W. also acknowledges funding provided by the French State and administered by the French National Research Agency (ANR) under the France 2030 programme (reference ANR-23-IACL-0008).

# 2 Main results

After setting out our conventions and standing assumptions for (1) and (2), we give equivalent path and marginal-flow formulations of the particle problem relative to a mean-field reference. Particle functional inequalities then yield two turnpike estimates, one for relative entropy along a general reference flow and one for modulated free energy with an equilibrium reference. The final subsection derives the particle inequalities from their mean-field counterparts.

## 2.1 Schrödinger problems

For a metric space $Y$, let $\mathcal{P}_2(Y)$ be the space of probability measures with finite second moment, equipped with the Wasserstein distance $W_2$. The continuous path space $\mathcal{C}([0,T];\mathcal{X})$ is equipped with the uniform distance. For probabilities $\nu$, $\nu'$ on

the same space, we set $\mathcal{H}(\nu|\nu') = \mathcal{I}(\nu|\nu') = +\infty$ if $\nu \not\ll \nu'$, and understand relative Fisher information as $\mathcal{I}(\nu|\nu') = 4\int |\nabla\sqrt{\mathrm{d}\nu/\,\mathrm{d}\nu'}|^2\,\mathrm{d}\nu'$ when $\nu \ll \nu'$. We identify measures with their densities when these exist.

For a trajectory $x_{[0,T]} \in \mathcal{C}([0,T];\mathcal{X})$ and a path law $P$, write $x_t$ for the value of $x_{[0,T]}$ at time $t$, and write $x_{[s,t]}$ and $P_{[s,t]}$ for their restrictions to $[s,t] \subset [0,T]$. Here and below, bold symbols denote $N$-tuples, such as $\boldsymbol{x} = (x^1,\dots,x^N) \in \mathcal{X}^N$ and $\boldsymbol{X} = (X^1,\dots,X^N)$.

Given endpoint laws $m_0$, $m_T \in \mathcal{P}_2(\mathcal{X})$ and $m_0^N$, $m_T^N \in \mathcal{P}_2(\mathcal{X}^N)$ and running costs $F_t\colon \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$, the infima in (1) and (2) are taken over $P \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}))$ and $P^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$, respectively. In each problem, the infimum is taken over the effective domain of the objective, where the integrals and their sum are well defined and the value is less than $+\infty$; in particular, $+\infty + (-\infty)$ is excluded. We use this convention for all variational problems, set $\inf\varnothing = +\infty$, and call a competitor finite-cost if its objective is real-valued.

For a mean-field functional $A\colon \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$, flat derivatives are understood through integration along linear mixtures:

$$A(\nu') - A(\nu) = \int_0^1 \big\langle \delta A\big((1-t)\nu + t\nu'\big), \nu' - \nu\big\rangle\,\mathrm{d}t,$$

with the analogous formula for $\delta A$ with derivative $\delta^2 A$. We write $DA(\nu,x) = \nabla_x \delta A(\nu,x)$ and $D^2A(\nu,x,y) = \nabla_x\nabla_y\delta^2 A(\nu,x,y)$ for the corresponding intrinsic derivatives. We use the same convention for functionals on $\mathcal{P}_2(\mathcal{X}^N)$.

For the growth conditions below, put $\langle x\rangle_{\mathcal{X}} = 1$ on $\mathbb{T}^d$ and $\langle x\rangle_{\mathcal{X}} = (1+|x|^2)^{1/2}$ on $\mathbb{R}^d$, and set $M_2(\nu) = \int \langle x\rangle_{\mathcal{X}}^2\,\mathrm{d}\nu$.

**Assumption Cost** (Running costs)**.** The running costs $F_t$ are measurable in time and twice flat differentiable on $\mathcal{P}_2(\mathcal{X})$. These derivatives are jointly continuous in measure and space and measurable in time. For some fixed $\bar\nu \in \mathcal{P}_2(\mathcal{X})$, $t \mapsto F_t(\bar\nu)$ belongs to $L^1(0,T)$. There is $C_T \geqslant 0$ such that

$$F_t(\nu) \leqslant C_T\big(1 + M_2(\nu)\big).$$

Moreover, for all $A \geqslant 0$ there is $C_{T,A} \geqslant 0$ such that for all $\nu \in \mathcal{P}_2(\mathcal{X})$ with $M_2(\nu) \leqslant A$ and all $x$, $y \in \mathcal{X}$,

$$|\delta F_t(\nu,x)| \leqslant C_{T,A}\langle x\rangle_{\mathcal{X}}^2, \quad |\delta^2 F_t(\nu,x,y)| \leqslant C_{T,A}\langle x\rangle_{\mathcal{X}}^2\langle y\rangle_{\mathcal{X}}^2.$$

## 2.2 Relative formulations

We give two formulations of (2): a path problem relative to independent copies of a mean-field reference law, and a control problem for marginal flows. Extending the modulated interaction $F_t(\nu|m_t)$ from Section 1, for any flat differentiable functional $A\colon \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$ and $m \in \mathcal{P}_2(\mathcal{X})$ we write

$$A(\nu|m) = A(\nu) - A(m) - \langle \delta A(m,\cdot), \nu - m\rangle$$

for its Bregman divergence.

To carry out this modulation, we assume that the mean-field reference satisfies the following first-order condition.

**Assumption Ref** (Mean-field reference)**.** There is a finite-cost competitor $P_{\mathrm{ref}}$ in (1), with marginals $m_t = P_{\mathrm{ref},t}$. There are nonnegative Borel endpoint factors $a_0$, $a_T \colon \mathcal{X} \to \mathbb{R}$ such that for $R$-almost every $x_{[0,T]} \in \mathcal{C}([0,T];\mathcal{X})$,

$$\frac{\mathrm{d}P_{\mathrm{ref}}}{\mathrm{d}R}(x_{[0,T]}) = a_0(x_0)a_T(x_T)\exp\!\left(-\frac{1}{2}\int_0^T \delta F_t(m_t, x_t)\,\mathrm{d}t\right). \tag{5}$$

The factors are finite and positive $m_e$-almost everywhere, and $\varphi_e \coloneqq 2\log a_e \in L^1(m_e)$ for $e = 0,\, T$.

By Lemma 24 in Appendix C, the factorization (5) makes $P_{\mathrm{ref}}$ Markov, and finiteness of the path entropy $\mathcal{H}(P_{\mathrm{ref}}|R)$ gives a Borel forward drift $b_{\mathrm{ref}} \colon [0,T]\times\mathcal{X} \to \mathbb{R}^d$ such that the canonical process under $P_{\mathrm{ref}}$ weakly solves

$$\mathrm{d}X_t = b_{\mathrm{ref},t}(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t.$$

The marginal flow $m_t$ satisfies the continuity equation with current velocity $v_t = b_{\mathrm{ref},t} - \nabla \log m_t$ on $(0,T)$.

We also assume that the particle endpoint laws are compatible with the chosen mean-field reference.

**Assumption End** (Particle endpoints)**.** The endpoint laws $m_0^N$, $m_T^N$ in (2) belong to $\mathcal{P}_2(\mathcal{X}^N)$ and satisfy $\mathcal{H}(m_e^N|m_e^{\otimes N}) < \infty$, $\int \sum_{i\in[N]} |\varphi_e(x^i)|\, m_e^N(\mathrm{d}\boldsymbol{x}) < \infty$ for $e = 0,\, T$.

Modulating (2) around $P_{\mathrm{ref}}$ gives the path problem

$$\inf_{P_0^N = m_0^N,\, P_T^N = m_T^N} \left\{ 2\mathcal{H}(P^N|P_{\mathrm{ref}}^{\otimes N}) + N\int_0^T \mathbb{E}_{P_r^N}[F_r(\mu_{\boldsymbol{X}}|m_r)]\,\mathrm{d}r \right\}, \tag{6}$$

taken over $P^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$.

We next introduce a control problem for a marginal flow $\nu^N$ and a relative velocity $w^N$ with respect to $v^{\oplus N}$. Fix $0 \leqslant s < t \leqslant T$. Let $\nu^N \in \mathcal{C}([s,t];\mathcal{P}_2(\mathcal{X}^N))$ satisfy $\mathcal{H}(\nu_s^N|m_s^{\otimes N}) < \infty$, and let $w^N \colon [s,t]\times\mathcal{X}^N \to \mathbb{R}^{dN}$ be a Borel field. The pair $(\nu^N, w^N)$ is called an admissible control if

$$\partial_r \nu_r^N + \nabla\cdot\left(\nu_r^N(v_r^{\oplus N} + w_r^N)\right) = 0 \tag{7}$$

in the sense of distributions, and

$$\int_s^t \left(\frac{1}{2}\int |w_r^N|^2\,\mathrm{d}\nu_r^N + \mathcal{I}(\nu_r^N|m_r^{\otimes N})\right)\mathrm{d}r < \infty. \tag{8}$$

The relative action of an admissible control is

$$\mathcal{A}_N^{s,t}(\nu^N, w^N) = \int_s^t \left(\frac{1}{2}\int |w_r^N|^2\,\mathrm{d}\nu_r^N + \frac{1}{2}\mathcal{I}(\nu_r^N|m_r^{\otimes N}) + N\,\mathbb{E}_{\nu_r^N}[F_r(\mu_{\boldsymbol{X}}|m_r)]\right)\mathrm{d}r. \tag{9}$$

We denote its integrand by $\mathcal{L}_r(\nu^N, w^N)$. For endpoints $\nu_s^N$, $\nu_t^N \in \mathcal{P}_2(\mathcal{X}^N)$ with $\mathcal{H}(\nu_e^N|m_e^{\otimes N}) < \infty$ for $e = s,\, t$, define the minimal action

$$\mathsf{C}_N(s,\nu_s^N;t,\nu_t^N) = \inf_{\substack{(\tilde\nu^N, w^N)\ \text{admissible} \\ \tilde\nu_s^N = \nu_s^N,\ \tilde\nu_t^N = \nu_t^N}} \mathcal{A}_N^{s,t}(\tilde\nu^N, w^N).$$

The following theorem, proved in Section 4, justifies both the change of reference and the passage from path laws to marginal flows.

**Theorem 1** (Equivalence of path and control problems)**.** *Let Assumptions Cost and Ref hold.*

*(a)* Change of reference. *Suppose Assumption End holds. The objectives in (2) and (6), on path laws $P^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$ with $P_0^N = m_0^N$ and $P_T^N = m_T^N$, have the same effective domain and differ there by a finite constant depending only on $N$, $T$, the endpoints and $P_{\mathrm{ref}}$.*

*(b)* Path laws to flows. *Every $P^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$ with $\mathcal{H}(P^N|P_{\mathrm{ref}}^{\otimes N}) < \infty$ has an admissible control $(\nu^N, w^N)$ with the same marginals and*

$$\mathcal{H}(\nu_0^N|m_0^{\otimes N}) + \mathcal{H}(\nu_T^N|m_T^{\otimes N}) + \int_0^T \left(\frac{1}{2}\int |w_r^N|^2 \,\mathrm{d}\nu_r^N + \frac{1}{2}\mathcal{I}(\nu_r^N|m_r^{\otimes N})\right) \mathrm{d}r \leqslant 2\mathcal{H}(P^N|P_{\mathrm{ref}}^{\otimes N}). \quad (10)$$

*If $P^N$ is Markov, equality holds for its relative velocity.*

*(c)* Flows to path laws. *Every admissible control $(\nu^N, w^N)$ on $[0,T]$ has a Markov diffusion law $P^N$ with the same marginals, noise $\sqrt{2}$ and drift*

$$\beta_r^N = (b_{\mathrm{ref},r})^{\oplus N} + w_r^N + \nabla \log \frac{\nu_r^N}{m_r^{\otimes N}}.$$

*Moreover, $\mathcal{H}(P^N|P_{\mathrm{ref}}^{\otimes N}) < \infty$, and (10) holds with equality.*

*(d)* Entropy chain rule. *For every $0 \leqslant s < t \leqslant T$ and every admissible control $(\nu^N, w^N)$ on $[s,t]$, the map $r \mapsto \mathcal{H}(\nu_r^N|m_r^{\otimes N})$ is absolutely continuous on $[s,t]$, with*

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathcal{H}(\nu_r^N|m_r^{\otimes N}) = \int \nabla \log \frac{\nu_r^N}{m_r^{\otimes N}} \cdot w_r^N \,\mathrm{d}\nu_r^N \qquad \text{for a.e. } r.$$

For endpoints satisfying Assumption End, parts (a)–(c) identify the three infima up to fixed endpoint constants, including infinite values. The two path problems have the same finite-cost minimizers. Their marginal flows are precisely those of the finite-cost optimal controls. On each subinterval every such control minimizes the relative action among all admissible controls with the same endpoints.

## 2.3 Turnpike-type chaoticity

We now state two quantitative turnpike estimates for the optimal particle marginals: Section 2.3.1 bounds the relative entropy along a general reference flow, and Section 2.3.2 bounds the modulated free energy for an equilibrium reference, attaining decay exponents arbitrarily close to the rate supplied by the free energy dissipation inequality.

### 2.3.1 Entropic turnpike

Let $\rho > 0$, $\eta_N \in (0,1]$ and $M_N$, $\Delta_F^-$, $\Delta_F^+ \geqslant 0$. For the entropy estimate we assume the following modulated energy bound and log-Sobolev inequality for all $\nu \in \mathcal{P}_2(\mathcal{X})$, $\nu^N \in \mathcal{P}_2(\mathcal{X}^N)$ and $t \in [0,T]$:

$$-\frac{1-\eta_N}{2}\mathcal{I}(\nu^N|m_t^{\otimes N}) - \Delta_F^- \leqslant N\,\mathbb{E}_{\nu^N}[F_t(\mu_{\boldsymbol{X}}|m_t)] \leqslant M_N\mathcal{H}(\nu^N|m_t^{\otimes N}) + \Delta_F^+. \quad \text{(EB)}$$

$$\mathcal{I}(\nu|m_t) \geqslant 2\rho\mathcal{H}(\nu|m_t). \quad \text{(LS)}$$

Here $\Delta_F^-$ and $\Delta_F^+$ are nonnegative additive defects in the lower and upper energy bounds, respectively; they need not be related.

The two sides of (EB) and the log-Sobolev inequality play distinct roles. The lower bound in (EB) ensures that the effective potential

$$\nu^N \mapsto \frac{1}{2}\mathcal{I}(\nu^N|m_t^{\otimes N}) + N\,\mathbb{E}_{\nu^N}[F_t(\mu_{\boldsymbol{X}}|m_t)],$$

in the relative action (9) is coercive in relative Fisher information, up to an additive defect. This coercivity is central to the long-time relaxation mechanism. In particular, if $F_t$ is flat convex, then the modulated interaction $F_t(\cdot|m_t)$ is nonnegative, so the lower bound in (EB) holds with $\eta_N = 1$ and $\Delta_F^- = 0$. The log-Sobolev inequality (LS) tensorizes to $m_t^{\otimes N}$, so the relative Fisher information controls the joint relative entropy uniformly on $[0,T]$. Finally, the upper bound in (EB) controls the modulated interaction by relative entropy, in the spirit of the Jabin–Z. Wang estimate [33]; see also [48, Corollary 3 and Remark 5].

**Theorem 2** (Entropic turnpike estimate)**.** *Under Assumptions Cost, Ref, and End, assume (EB), (LS), and $T \geqslant c_N^{-1}$. Then the $N$-particle Schrödinger problem (2) admits a unique finite-cost optimizer $P_{\mathrm{opt}}^N$, which is Markov and is symmetric if both endpoint laws are symmetric. Moreover, writing $m_t^N = P_{\mathrm{opt},t}^N$ and setting $h_t = \mathcal{H}(m_t^N|m_t^{\otimes N})$, we have for every $t \in [0,T]$,*

$$h_t \leqslant C_N\left(e^{-c_N t}h_0 + e^{-c_N(T-t)}h_T + \frac{\Delta_F^- + \Delta_F^+}{\eta_N \rho}\right), \tag{11}$$

*where*

$$c_N = \frac{\eta_N\rho}{8\sqrt{1+\frac{M_N}{\rho}}}, \quad C_N = 100\sqrt{\frac{1+M_N/\rho}{\eta_N}}.$$

The theorem is proved in Section 5.3.

The quantitative estimate (11) depends only on the constants in (EB) and (LS). When $\rho$ and $\eta_N$ are bounded away from zero and $M_N$ and the defects are bounded above uniformly in $N$ and $T$, it gives entropy relaxation towards $m_t^{\otimes N}$ on a time scale independent of $N$ and $T$, up to a finite-particle error. For symmetric laws, let $m_t^{N,k}$ denote the $k$-particle marginal. Entropy superadditivity gives

$$\mathcal{H}(m_t^{N,k}|m_t^{\otimes k}) \leqslant \frac{\mathcal{H}(m_t^N|m_t^{\otimes N})}{\lfloor N/k \rfloor} = \frac{h_t}{\lfloor N/k \rfloor}.$$

Thus (11) also bounds the relative entropy of each $k$-particle marginal. For product endpoint laws matching the reference flow, only the defect term remains, giving an $O(k/N)$ bound for each fixed $k$, uniformly in time, and hence chaoticity as $N \to \infty$. With endpoint entropies of order $N$, the same $O(k/N)$ bound holds once the distances from both endpoints are at least $c_N^{-1}\log N$, yielding generation of chaos in the interior.

### 2.3.2 Free-energetic turnpike

When the mean-field reference is at equilibrium, we further exploit the free energy structure to obtain a near-optimal relaxation rate. We impose the following additional assumption.

**Assumption Eq** (Equilibrium reference and corrector)**.** The cost is autonomous, $F_t = F$, and the reference marginals in Assumption Ref are stationary, $m_t = m_*$. There is a mean-field functional $\chi\colon \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$, twice continuously flat differentiable with jointly continuous flat derivatives, such that for all $\nu \in \mathcal{P}_2(\mathcal{X})$,

$$\int \left( \nabla_x \cdot D\chi(\nu, x) - \frac{1}{2} |D\chi(\nu, x)|^2 \right) \nu(\mathrm{d}x) + F(\nu) = \text{const.} \tag{12}$$

Moreover, for some $C \geqslant 0$,

$$-C \leqslant \chi(\nu) \leqslant C(1 + M_2(\nu)),$$

and for every $A \geqslant 0$ there is $C_A \geqslant 0$ such that, whenever $M_2(\nu) \leqslant A$,

$$|\delta\chi(\nu, x)| \leqslant C_A \langle x \rangle_{\mathcal{X}}^2, \qquad |\delta^2\chi(\nu, x, y)| \leqslant C_A \langle x \rangle_{\mathcal{X}}^2 \langle y \rangle_{\mathcal{X}}^2.$$

Its intrinsic derivatives are jointly continuous in measure and space, including at empirical measures, and satisfy

$$\begin{aligned} |D\chi(\nu, x)| &\leqslant C(\langle x \rangle_{\mathcal{X}} + M_2(\nu)^{1/2}), \\ |\nabla_x D\chi(\nu, x)| &\leqslant C(1 + M_2(\nu) + \langle x \rangle_{\mathcal{X}}^2), \\ |D^2\chi(\nu, x, y)| &\leqslant C(1 + M_2(\nu) + \langle x \rangle_{\mathcal{X}}^2 + \langle y \rangle_{\mathcal{X}}^2). \end{aligned}$$

The corrector is compatible with the equilibrium reference:

$$D\chi(m_*, x) = -\nabla \log m_*(x), \qquad x \in \mathcal{X}.$$

Finally, there is $C_\chi \geqslant 0$ such that

$$|\operatorname{tr} D^2\chi(\nu, x, x)| \leqslant C_\chi \tag{13}$$

for all $\nu \in \mathcal{P}_2(\mathcal{X})$ and $x \in \mathcal{X}$.

Such a functional $\chi$ is called an *ergodic corrector* for $(F, m_*)$. The associated modulated free energies at the mean-field and particle levels are, respectively,

$$\begin{aligned} \mathcal{F}_*(\nu) &= \mathcal{H}(\nu | m_*) + \chi(\nu | m_*), \\ \mathcal{F}_{*,N}(\nu^N) &= \mathcal{H}(\nu^N | m_*^{\otimes N}) + N\, \mathbb{E}_{\nu^N}\, \chi(\mu_{\boldsymbol{X}} | m_*). \end{aligned} \tag{14}$$

As noted in Section 1, for measures $\nu$ with finite relative Fisher information, (12) gives $\frac{1}{2}\mathcal{I}(\nu | m_*) + F(\nu | m_*) = \frac{1}{2}\|D\mathcal{F}_*(\nu, \cdot)\|^2_{L^2(\nu)}$.

We further assume that there exist $\alpha_N$, $\lambda_N > 0$ and $\Delta_{\mathrm{ent}}$, $\Delta_{\mathrm{rel}} \geqslant 0$ such that, for every $\nu^N \in \mathcal{P}_2(\mathcal{X}^N)$,

$$\frac{1}{2} \int |D\mathcal{F}_{*,N}(\nu^N, \cdot)|^2 \, \mathrm{d}\nu^N \geqslant \lambda_N \mathcal{F}_{*,N}(\nu^N) - \Delta_{\mathrm{rel}}, \tag{PL}$$

$$\mathcal{F}_{*,N}(\nu^N) \geqslant \alpha_N \mathcal{H}(\nu^N | m_*^{\otimes N}) - \Delta_{\mathrm{ent}}. \tag{Ent}$$

Condition (PL) is a Polyak–Łojasiewicz inequality in Wasserstein space with an additive defect. Along the gradient flow of $\mathcal{F}_{*,N}$, it gives free energy relaxation at rate $2\lambda_N$, up to the defect $\Delta_{\mathrm{rel}}/\lambda_N$. The following theorem establishes a two-sided counterpart for the optimal curve, attaining every exponent $2\beta_N < 2\lambda_N$.

**Theorem 3** (Free-energetic turnpike estimate)**.** *Under Assumptions Cost, Ref, End, and Eq, assume* (LS)*, the upper bound in* (EB)*,* (Ent)*,* (PL)*, and* $T \geqslant \lambda_N^{-1}$*. If both endpoint laws are symmetric,* (PL) *is required only for symmetric joint laws. Then* (2) *admits a unique finite-cost optimizer* $P^N_{\mathrm{opt}}$*, which is Markov and is symmetric if both endpoint laws are symmetric. Moreover, writing* $m_t^N = P^N_{\mathrm{opt},t}$ *and setting* $f_t = \mathcal{F}_{*,N}(m_t^N)$*, we have for every* $0 < \beta_N < \lambda_N$*,* $\varepsilon > 0$*, and* $t \in [0,T]$*,*

$$\begin{aligned} f_t + \Delta_{\mathrm{ent}} \leqslant C_{N,\varepsilon,T}\big(e^{-2\beta_N t}(f_0 + \Delta_{\mathrm{ent}}) + e^{-2\beta_N(T-t)}(f_T + \Delta_{\mathrm{ent}})\big) \\ + 20A_N(1+\varepsilon^{-1})^2 D_{N,\beta}, \end{aligned} \tag{15}$$

*where*

$$\begin{aligned} A_N &= 16\max\left\{1, \frac{\sqrt{1+M_N/\rho}+2\lambda_N/\rho}{\alpha_N}\right\}, \\ D_{N,\beta} &= \Delta_{\mathrm{ent}} + \frac{\Delta_{\mathrm{rel}} + \Delta_F^+ + C_\chi}{\lambda_N} + \frac{C_\chi}{\lambda_N - \beta_N}, \\ C_{N,\varepsilon,T} &= \min\big\{A_N,\, 1+\varepsilon + 32A_N^3(1+\varepsilon^{-1})^3 e^{-(\lambda_N-\beta_N)T}\big\}. \end{aligned}$$

The theorem is proved in Section 6.3.

By (Ent), $\mathcal{H}(m_t^N | m_*^{\otimes N}) \leqslant (f_t + \Delta_{\mathrm{ent}})/\alpha_N$, so (15) also gives a turnpike estimate for relative entropy, as in (11).

## 2.4 Transfer of functional inequalities

This subsection provides sufficient conditions for the particle functional inequalities (EB), (Ent), and (PL). More precisely, under suitable regularity assumptions, we derive these inequalities from their mean-field counterparts, with a small loss in the multiplicative constants and an additive defect.

To accommodate general running costs $F$ and correctors $\chi$, we extend the conditional approximation method developed for pair interactions in [48] by imposing the following bound on second flat derivatives.

Let $\mathcal{M}_0(\mathcal{X})$ be the space of finite signed measures of total mass zero, with finite second moment with respect to their total variation on $\mathbb{R}^d$. A symmetric kernel $K\colon \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ is called positive if $\langle K, \sigma^{\otimes 2}\rangle \geqslant 0$ for every $\sigma \in \mathcal{M}_0(\mathcal{X})$. We define its *double oscillation* as

$$\operatorname{Osc}_2 K = \sup_{x,y,z,w\in\mathcal{X}} \langle K, (\delta_x - \delta_y)\otimes(\delta_z - \delta_w)\rangle.$$

Let $A\colon \mathcal{P}_2(\mathcal{X}) \to \mathbb{R}$ be twice continuously flat differentiable. For $M_A$, $L_A \geqslant 0$, we say that $A$ is $(M_A, L_A)$-*regular* if there is a symmetric positive kernel $K_A\colon \mathcal{X}\times\mathcal{X}\to\mathbb{R}$ such that for all $\nu \in \mathcal{P}_2(\mathcal{X})$ and $\sigma \in \mathcal{M}_0(\mathcal{X})$,

$$|\langle \delta^2 A(\nu), \sigma^{\otimes 2}\rangle| \leqslant \langle K_A, \sigma^{\otimes 2}\rangle, \tag{16}$$

and the following bounds hold:

- For $\mathcal{X} = \mathbb{T}^d$, $\operatorname{Osc}_2 K_A \leqslant 4M_A$; set $L_A = 0$.
- For $\mathcal{X} = \mathbb{R}^d$, $K_A$ admits a decomposition $K_A = K_{A,\mathrm{b}} + K_{A,\mathrm{q}}$ into symmetric positive kernels satisfying

$$\operatorname{Osc}_2 K_{A,\mathrm{b}} \leqslant 4M_A, \quad \|\nabla^2_{1,2} K_{A,\mathrm{q}}\|_\infty \leqslant L_A.$$

Here the matrix norm is the operator norm.

Fix $m \in \mathcal{P}_2(\mathcal{X})$ satisfying the log-Sobolev inequality $\mathcal{I}(\nu|m) \geqslant 2\rho\mathcal{H}(\nu|m)$ with $\rho > 0$, and put $\ell_A = L_A/\rho$. For a functional $\chi$ and an arbitrary reference $m$, write

$$\begin{aligned} \mathcal{F}(\nu|m) &= \mathcal{H}(\nu|m) + \chi(\nu|m), \\ \mathcal{F}_N(\nu^N; m) &= \mathcal{H}(\nu^N|m^{\otimes N}) + N\,\mathbb{E}_{\nu^N}\,\chi(\mu_{\boldsymbol{X}}|m). \end{aligned} \tag{17}$$

Here the semicolon in $\mathcal{F}_N(\nu^N; m)$ marks $m$ as a reference parameter: the corrector is modulated before taking its empirical lift. At $m = m_*$ these are $\mathcal{F}_*$ and $\mathcal{F}_{*,N}$ from (14). The mean-field conditions use constants $\eta$, $\alpha \in (0, 1]$, $M \geqslant 0$, and $\lambda > 0$ and are imposed for every $\nu \in \mathcal{P}_2(\mathcal{X})$:

$$-\frac{1-\eta}{2}\mathcal{I}(\nu|m) \leqslant F(\nu|m) \leqslant M\mathcal{H}(\nu|m). \tag{MF-EB}$$

$$\frac{1}{2}\mathcal{I}(\nu|m) + F(\nu|m) \geqslant \lambda\big(\mathcal{H}(\nu|m) + \chi(\nu|m)\big), \tag{MF-PL}$$

$$\mathcal{F}(\nu|m) \geqslant \alpha\mathcal{H}(\nu|m). \tag{MF-Ent}$$

The following theorem transfers the mean-field conditions (MF-EB)–(MF-Ent) to the $N$-particle setting with fixed losses in the multiplicative constants and $O(1)$ additive defects for all sufficiently large $N$.

**Theorem 4** (Transfer of mean-field inequalities)**.** *Let $m$ satisfy the log-Sobolev inequality above. Each conclusion holds for every $\nu^N \in \mathcal{P}_2(\mathcal{X}^N)$, with symmetry required in* (d)*.*

*(a) If $F$ is $(M_F, L_F)$-regular and the lower bound in (MF-EB) holds, then for every $\theta > 0$,*

$$\frac{1}{2}\mathcal{I}(\nu^N|m^{\otimes N}) + N\,\mathbb{E}_{\nu^N}[F(\mu_{\boldsymbol{X}}|m)] \geqslant \frac{\eta_N}{2}\mathcal{I}(\nu^N|m^{\otimes N}) - \Delta_F^-,$$

*where*

$$\begin{aligned} \eta_N &= \eta - \frac{\theta(M_F + \ell_F)}{\rho} - \frac{2(1+\theta^{-1})\ell_F}{\rho N}, \\ \Delta_F^- &= 2(1+\theta^{-1})(M_F + d\ell_F). \end{aligned}$$

*(b) If $F$ is $(M_F, L_F)$-regular and the upper bound in (MF-EB) holds, then for every $\theta > 0$,*

$$N\,\mathbb{E}_{\nu^N}[F(\mu_{\boldsymbol{X}}|m)] \leqslant M_N\mathcal{H}(\nu^N|m^{\otimes N}) + \Delta_F^+,$$

*where*

$$\begin{aligned} M_N &= M + \theta(M_F + \ell_F) + \frac{2(1+\theta^{-1})\ell_F}{N}, \\ \Delta_F^+ &= 2(1+\theta^{-1})(M_F + d\ell_F). \end{aligned}$$

*(c) If $\chi$ is $(M_\chi, L_\chi)$-regular and (MF-Ent) holds, then for every $\vartheta > 0$,*

$$\mathcal{F}_N(\nu^N; m) \geqslant \alpha_N\mathcal{H}(\nu^N|m^{\otimes N}) - \Delta_{\mathrm{ent}},$$

*where*

$$\begin{aligned} \alpha_N &= \alpha - \vartheta(M_\chi + \ell_\chi) - \frac{2(1+\vartheta^{-1})\ell_\chi}{N}, \\ \Delta_{\mathrm{ent}} &= 2(1+\vartheta^{-1})(M_\chi + d\ell_\chi). \end{aligned}$$

(d) *Suppose the lower bound in* (MF-EB) *and* (MF-PL) *hold, with* $F$ $(M_F, L_F)$*-regular and* $\chi$ $(M_\chi, L_\chi)$*-regular. Choose* $\theta > 0$ *with* $\eta_N > 0$ *in* (a)*, and set*

$$M_\lambda = M_F + \lambda M_\chi, \qquad \ell_\lambda = \ell_F + \lambda \ell_\chi.$$

*Then for every* $\varepsilon > 0$*,*

$$\frac{1}{2}\mathcal{I}(\nu^N | m^{\otimes N}) + N\,\mathbb{E}_{\nu^N}[F(\mu_{\boldsymbol{X}}|m)] \geqslant \lambda_N \mathcal{F}_N(\nu^N; m) - \Delta_{\mathrm{rel}}.$$

*Here*

$$\begin{aligned} z_N &= \frac{\varepsilon(M_\lambda + \ell_\lambda) + 5(1+\varepsilon^{-1})\ell_\lambda/N}{\rho\eta_N}, \\ \lambda_N &= \frac{\lambda}{1+z_N}, \\ \Delta_{\mathrm{rel}} &= \frac{5(1+\varepsilon^{-1})(M_\lambda + d\ell_\lambda) + z_N \Delta_F^-}{1+z_N}. \end{aligned}$$

The theorem is proved in Section 7.

Applying parts (a)–(b) with $m = m_t$ and $F = F_t$ gives (EB), uniformly in time if the kernels and constants are common to all $t$. With $m = m_*$ and an ergodic corrector $\chi$ as in Assumption Eq, part (c) gives (Ent). The free energy identity (31) in Section 6.1 turns part (d) into (PL), with an additional defect $C_\chi$.

**Near-optimality.** The argument following [48, Theorem 5] already explains the asymptotic optimality of the passage from mean-field PL to particle PL inequalities. We focus here on the further passage from (PL) to the Schrödinger turnpike estimate. Consider the equilibrium setting $m = m_*$, and let $\lambda$ be the optimal constant in (MF-PL). Under the other assumptions above, for every $0 < \beta < \lambda$ and all sufficiently large $N$, the two theorems give

$$\text{(MF-PL)} \xRightarrow[\lambda > \lambda_N]{\text{Theorem 4}} \text{(PL)} \xRightarrow[\lambda_N > \beta]{\text{Theorem 3}} (15).$$

For each fixed loss in the exponent and each $0 < \varepsilon \leqslant 1$, the defects are $O(1)$ and $C_{N,\varepsilon,T} \to 1 + \varepsilon$ as $T \to \infty$, uniformly for large $N$.

Suppose that in (15) the exponent $2\beta_N$ could be replaced by $2\gamma > 2\lambda$, with $\gamma$ independent of $N$ and $T$. Assume that, for each fixed $\varepsilon$, the additive error remains bounded uniformly in $N$ and $T$, and the endpoint prefactor is $1 + \varepsilon + o_{T\to\infty}(1)$, uniformly for large $N$. Let $\nu$ be any regular law with finite relative Fisher information. Take $m_0^N = \nu^{\otimes N}$ and $m_T^N = m_*^{\otimes N}$. By the same calculation as in [48, Lemma 16], the present regularity assumptions give $f_0/N \to \mathcal{F}_*(\nu)$ and $f_T = O(1)$. The assumed estimate, divided by $N$, would therefore imply

$$\frac{f_t}{N} \leqslant \big(1 + \varepsilon + o_{T\to\infty}(1)\big) e^{-2\gamma t} \frac{f_0}{N} + O_\varepsilon(N^{-1}).$$

Formally, letting $N \to \infty$ and then $T \to \infty$ at fixed $t$ would give an optimizer $\tilde{\nu}$ of the infinite-horizon mean-field ergodic problem starting at $\nu$. Letting $\varepsilon \downarrow 0$ would then yield

$$\mathcal{F}_*(\tilde{\nu}_t) \leqslant e^{-2\gamma t} \mathcal{F}_*(\nu)$$

for $t \geqslant 0$. Thus its free energy would relax at rate $2\gamma > 2\lambda$.

The infinite-horizon ergodic optimizer is expected to follow the gradient flow of $\mathcal{F}_*$, in analogy with the convergence of entropic interpolations to gradient flows established in [22, Theorems 4.2 and 4.7], namely

$$\partial_t \tilde{\nu}_t = \nabla \cdot \big(\tilde{\nu}_t D\mathcal{F}_*(\tilde{\nu}_t, \cdot)\big), \qquad \tilde{\nu}_0 = \nu.$$

For this gradient flow, differentiating the above decay estimate at $t = 0$ would give, for every regular initial law $\nu$,

$$\|D\mathcal{F}_*(\nu, \cdot)\|^2_{L^2(\nu)} = -\left.\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_*(\tilde{\nu}_t)\right|_{t=0} \geqslant 2\gamma\mathcal{F}_*(\nu).$$

By (30), this is (MF-PL) with constant $\gamma$, so $\gamma \leqslant \lambda$, contrary to what is assumed above. This heuristic argument suggests that the exponent in (15) cannot be improved.

As a side note, Porretta [40] and Cirant and De Bernardi [20] obtained long-time convergence rates for mean-field control by analyzing the linearized dynamics at stationarity and its associated spectral gap. The PL constant $\lambda$ here is a global nonlinear counterpart to this spectral gap.

# 3 Examples

This section presents examples to which the results in Section 2 can be applied. Section 3.1 treats mildly flat-semiconvex running costs on the torus $\mathbb{T}^d$, establishing an entropic turnpike estimate along general mean-field references and a near-optimal free-energetic turnpike estimate around the stationary equilibrium. Section 3.2 treats displacement convex costs on $\mathbb{R}^d$, establishing an entropic turnpike estimate around the stationary equilibrium.

## 3.1 Mildly flat-semiconvex costs on the torus

Let $\mathcal{X} = \mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$ and fix $s > d/2$. Let $\mathcal{H}_0^{-s}$ be the space of zero-mean distributions in $\mathcal{H}^{-s}(\mathbb{T}^d)$, with norm $\|\cdot\|_{-s}$ and reproducing kernel $k_s$ of diagonal value $\kappa_s = k_s(0,0) < \infty$; see Appendix A for details. We consider running costs $(F_t)_{t\in[0,T]}$ satisfying Assumption Cost together with uniform first-variation oscillation and $\mathcal{H}^{-s}$ second-variation bounds: for some $C_1$, $B_F$, $\kappa \geqslant 0$ and all $t \in [0,T]$, $\nu \in \mathcal{P}(\mathbb{T}^d)$, and $\sigma \in \mathcal{H}_0^{-s}$,

$$\operatorname{Osc} \delta F_t(\nu, \cdot) \leqslant 2C_1, \quad -\kappa\|\sigma\|^2_{-s} \leqslant \langle \delta^2 F_t(\nu), \sigma^{\otimes 2}\rangle \leqslant B_F\|\sigma\|^2_{-s}. \tag{18}$$

Concrete models satisfying (18) include cylindrical costs

$$F_t(\nu) = \int V_t \,\mathrm{d}\nu + \Phi_t\big(\langle \ell, \nu\rangle\big),$$

where $V_t$ is uniformly bounded in $\mathcal{H}^s(\mathbb{T}^d)$, $\ell \in \mathcal{H}^s(\mathbb{T}^d; \mathbb{R}^r)$, and $\Phi_t$ is uniformly bounded in $\mathcal{C}^2$ on a neighborhood of the convex hull of $\ell(\mathbb{T}^d)$. If $-a\mathrm{Id} \leqslant \nabla^2\Phi_t \leqslant b\mathrm{Id}$ there, with $a$, $b \geqslant 0$, then (18) holds with $\kappa = a\|\ell\|^2_{\mathcal{H}^s}$, $B_F = b\|\ell\|^2_{\mathcal{H}^s}$, and some finite $C_1$.

Fix endpoint marginals $m_0$, $m_T \in \mathcal{P}(\mathbb{T}^d)$ with $\log m_0$, $\log m_T \in L^\infty(\mathbb{T}^d)$; we first verify the hypotheses of Theorem 2 along a general reference bridge. By

Lemma 21(a) in Appendix A, the mean-field problem (1) admits a finite-cost minimizer $P_{\mathrm{ref}}$ satisfying (5) with $\varphi_0$, $\varphi_T \in L^\infty(\mathbb{T}^d)$. This verifies Assumption Ref and reduces Assumption End to $\mathcal{H}(m_e^N|m_e^{\otimes N}) < \infty$ for $e = 0, T$. When $T \geqslant 2/\pi^2$, Lemma 21(b) bounds $\sup_{t\in[0,T]} \operatorname{Osc} \log m_t$ for $m_t = P_{\mathrm{ref},t}$ and applies the Holley–Stroock perturbation principle [32] on $(\mathbb{T}^d, \mathrm{d}x)$ to verify (LS) uniformly on $[0,T]$, with the constant $\rho \in (0, 4\pi^2]$ given in (54) in Appendix A.

We next verify the mean-field energy bound (MF-EB) along $m_t$. By Taylor's formula for $F_t(\nu|m_t)$, (18), and the Pinsker bound (48) in Appendix A, we have

$$-\kappa\kappa_s \mathcal{H}(\nu|m_t) \leqslant F_t(\nu|m_t) \leqslant B_F \kappa_s \mathcal{H}(\nu|m_t), \qquad \nu \in \mathcal{P}(\mathbb{T}^d).$$

Assuming $\kappa\kappa_s < \rho$ from now on, (LS) implies the lower bound in (MF-EB) with $\eta = 1 - \kappa\kappa_s/\rho > 0$, while the upper bound in (MF-EB) holds with $M = B_F\kappa_s$. In addition, Lemma 21(c) shows that $P_{\mathrm{ref}}$ is the unique finite-cost minimizer of (1).

By (47) in Appendix A and (18), every $F_t$ is $(M_F, 0)$-regular with $M_F = \max\{B_F, \kappa\}\kappa_s$. Theorem 4(a)–(b) therefore gives (EB) with constants independent of $N$ and $T$. Applying Theorem 2 gives the following proposition.

**Proposition 5** (Entropic turnpike on the torus)**.** *Under the above assumptions with $\kappa\kappa_s < \rho$, there exist constants $c$, $C > 0$, independent of $N$ and $T$, such that for all $N \geqslant 1$, $T \geqslant c^{-1}$, and $m_0^N$, $m_T^N \in \mathcal{P}((\mathbb{T}^d)^N)$ with $\mathcal{H}(m_e^N|m_e^{\otimes N}) < \infty$ ($e = 0$, $T$), the unique finite-cost optimizer $P_{\mathrm{opt}}^N$ of (2) satisfies, for all $t \in [0,T]$,*

$$\mathcal{H}(m_t^N|m_t^{\otimes N}) \leqslant C\big(e^{-ct}\mathcal{H}(m_0^N|m_0^{\otimes N}) + e^{-c(T-t)}\mathcal{H}(m_T^N|m_T^{\otimes N}) + 1\big),$$

*where $m_t = P_{\mathrm{ref},t}$ and $m_t^N = P_{\mathrm{opt},t}^N$.*

We now turn to the autonomous equilibrium setting $F_t = F$ and verify the hypotheses of Theorem 3. Suppose $F$ has Sobolev regularity with constants $C_F$, $B_F < \infty$ in the sense of Appendix A and satisfies (18) with $\kappa \leqslant \kappa_0$, where $\kappa_0 = \kappa(d, s, C_F, B_F) > 0$ is the threshold in Proposition 20 in Appendix A. That proposition provides the unique minimizer $m_*$ of $\frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x) + F$, which satisfies $\log m_* \in \mathcal{H}^{s+2}(\mathbb{T}^d)$ and (LS) for some $\rho \geqslant \rho_0(d, s, C_F) > 0$ with $\kappa\kappa_s \leqslant \rho/2$, together with an ergodic corrector $\chi$ satisfying Assumption Eq and $(M_\chi, 0)$-regularity. For $m_0 = m_T = m_*$, applying Itô's formula to $\frac{1}{2}\log m_*(X_t)$ along the Euler equation for $\sqrt{m_*}$ and Girsanov's theorem shows that the stationary diffusion $P_{\mathrm{ref}}$ with invariant law $m_*$ and drift $\nabla \log m_*$ satisfies (5) with bounded $\varphi_0$, $\varphi_T$. The bounded drift makes $P_{\mathrm{ref}}$ finite-cost, verifying Assumption Ref. The lower bound in (MF-EB) holds at $m_*$ with $\eta = 1 - \kappa\kappa_s/\rho \geqslant 1/2$, so Lemma 21(c) shows that $P_{\mathrm{ref}}$ is the unique finite-cost minimizer of (1). Assumption End again reduces to $\mathcal{H}(m_e^N|m_*^{\otimes N}) < \infty$ ($e = 0, T$).

Next, we verify the mean-field inequalities (MF-Ent) and (MF-PL) at $m_*$. Combining (50) and (51) with (48) in Appendix A gives, whenever $\mathcal{H}(\nu|m_*) < \infty$,

$$-\frac{1}{2}\mathcal{H}(\nu|m_*) \leqslant \chi(\nu|m_*) \leqslant \frac{B_F\kappa_s}{2\rho}\mathcal{H}(\nu|m_*),$$

where the lower bound holds for all $\nu \in \mathcal{P}(\mathbb{T}^d)$ and improves to 0 when $F$ is flat convex. Thus (MF-Ent) holds with $\alpha = 1/2$ ($\alpha = 1$ if $F$ is flat convex). Moreover, let $\lambda$ be the optimal constant in (MF-PL) at $m = m_*$, namely

$$\lambda = \inf_{\substack{\nu\in\mathcal{P}(\mathbb{T}^d)\setminus\{m_*\} \\ \mathcal{H}(\nu|m_*)<\infty}} \frac{\frac{1}{2}\mathcal{I}(\nu|m_*) + F(\nu|m_*)}{\mathcal{F}_*(\nu)}. \tag{19}$$

Testing at any smooth $\nu \neq m_*$ gives $\lambda < \infty$, while the upper bound on $\chi(\nu|m_*)$ and $\frac{1}{2}\mathcal{I}(\nu|m_*) + F(\nu|m_*) \geqslant \eta\rho\mathcal{H}(\nu|m_*)$ yield

$$\lambda \geqslant \frac{\eta\rho}{1 + B_F\kappa_s/(2\rho)} > 0,$$

verifying (MF-PL) with $\lambda \in (0, \infty)$.

Finally, since $F$ is $(M_F, 0)$-regular and $\chi$ is $(M_\chi, 0)$-regular by Proposition 20(c), for any $0 < \beta < \lambda$, Theorem 4 (together with (31); see Section 2.4) yields the upper bound in (EB) and (Ent) on $\mathcal{P}((\mathbb{T}^d)^N)$, and (PL) on symmetric laws with $\lambda_N \in (\beta, \lambda)$, all with constants independent of $N$. Applying Theorem 3 with $\beta_N = \beta$ gives the following proposition.

**Proposition 6** (Free-energetic turnpike on the torus)**.** *Under the above equilibrium assumptions, let $\lambda \in (0, \infty)$ be the optimal constant in (MF-PL) at $m = m_*$, given by (19). There exists $\Delta_{\mathrm{ent}} \geqslant 0$, and for every $0 < \beta < \lambda$ there exists $C_\beta > 0$, both independent of $N$ and $T$, such that for all $N \geqslant 1$, $T \geqslant \beta^{-1}$, and symmetric $m_0^N$, $m_T^N \in \mathcal{P}((\mathbb{T}^d)^N)$ with $\mathcal{H}(m_e^N|m_*^{\otimes N}) < \infty$, $e = 0, T$, the unique finite-cost optimizer $P_{\mathrm{opt}}^N$ of (2) with $m_0 = m_T = m_*$ satisfies, for all $t \in [0, T]$,*

$$f_t + \Delta_{\mathrm{ent}} \leqslant C_\beta\big(e^{-2\beta t}(f_0 + \Delta_{\mathrm{ent}}) + e^{-2\beta(T-t)}(f_T + \Delta_{\mathrm{ent}}) + 1\big),$$

*where $f_t = \mathcal{F}_{*,N}(m_t^N)$ and $m_t^N = P_{\mathrm{opt},t}^N$; the same bound holds with $\mathcal{H}(m_t^N|m_*^{\otimes N})$ on the left-hand side.*

*Remark* 1. We restrict this example to $\mathbb{T}^d$ for two reasons. Both Theorems 2 and 3 require log-Sobolev bounds (LS) for the reference bridges uniformly in time and horizon, which are more delicate to establish on $\mathbb{R}^d$; the perturbation approach of a previous work of the first-listed author [39] and the weak log-concavity estimates of Chaintron, Conforti and Eichinger [16] may help here. In addition, obtaining particle PL constants arbitrarily close to the optimal mean-field constant in Theorem 3 through our transfer argument requires $(M, L)$-regularity of the running costs and corrector, whose verification on $\mathbb{R}^d$ is also more involved.

## 3.2 Displacement convex costs on Euclidean space

Let $\mathcal{X} = \mathbb{R}^d$ and consider an autonomous running cost $F$ satisfying Assumption Cost. We assume that $F$ is strongly displacement convex and semiconcave along arbitrary coupling interpolations: for some $0 < \lambda \leqslant K$, every $\nu_0$, $\nu_1 \in \mathcal{P}_2(\mathbb{R}^d)$, every coupling $\pi$ of $\nu_0$, $\nu_1$, and every $r \in [0, 1]$, the interpolated measure $\nu_r = ((x, y) \mapsto (1 - r)x + ry)_\#\pi$ satisfies

$$\begin{aligned}\frac{\lambda}{2}r(1-r)\int |y-x|^2\,\pi(\mathrm{d}x,\mathrm{d}y)& \\ \leqslant (1-r)F(\nu_0) + rF(\nu_1) - F(\nu_r)& \\ \leqslant \frac{K}{2}r(1-r)\int |y-x|^2\,\pi(\mathrm{d}x,\mathrm{d}y).& \quad (20)\end{aligned}$$

We also assume that the mixed intrinsic derivative $D^2F$ is jointly continuous and for all $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $x$, $y \in \mathbb{R}^d$,

$$\|D^2F(\nu, x, y)\|_{\mathrm{op}} \leqslant K. \quad (21)$$

Concrete models include potential and interaction costs

$$F(\nu) = \int_{\mathbb{R}^d} U(x)\,\nu(\mathrm{d}x) + \frac{1}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} W(x-y)\,\nu(\mathrm{d}x)\,\nu(\mathrm{d}y) \tag{22}$$

with $U$, $W \in \mathcal{C}^2(\mathbb{R}^d)$, $W$ even, $\lambda \mathrm{Id} \leqslant \nabla^2 U \leqslant K_U \mathrm{Id}$ and $0 \leqslant \nabla^2 W \leqslant K_W \mathrm{Id}$, for which we can take $K = K_U + K_W$.

We verify the hypotheses of Theorem 2 for an equilibrium reference. To keep the presentation short, the technical results are collected in Appendix B. By Lemma 22, $\mathcal{E} = \frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x) + F$ admits a minimizer $m_*$, which satisfies

$$\sqrt{\lambda}\,\mathrm{Id} \leqslant -\nabla^2 \log m_* \leqslant \sqrt{K}\,\mathrm{Id}. \tag{23}$$

In particular, the Bakry–Émery criterion verifies (LS) with $\rho = \sqrt{\lambda}$. Lemma 23(a) in the same appendix shows that $m_*$ is unique and satisfies the lower bound in (MF-EB) with $\eta = \lambda/(\lambda + K)$. For $m_0 = m_T = m_*$, let $P_{\mathrm{ref}}$ be the stationary diffusion with invariant law $m_*$ and drift $\nabla \log m_*$. Writing $c_*$ for the eigenvalue in (55), Itô's formula and Girsanov's theorem give

$$\frac{\mathrm{d}P_{\mathrm{ref}}}{\mathrm{d}R}(x_{[0,T]}) = \sqrt{\frac{m_*(x_T)}{m_*(x_0)}}\exp\left(\frac{c_* T}{2} - \frac{1}{2}\int_0^T \delta F(m_*, x_t)\,\mathrm{d}t\right).$$

Thus (5) holds with $\varphi_0 = -\log m_*$ and $\varphi_T = \log m_* + c_* T$. The quadratic growth of the endpoint potentials follows from (23). The linear growth of the drift and Gaussian tails of $m_*$ make $P_{\mathrm{ref}}$ finite-cost. Thus Assumption Ref holds, and Lemma 21(c) shows that $P_{\mathrm{ref}}$ is the unique finite-cost minimizer of (1). Assumption End reduces to $\mathcal{H}(m_e^N|m_*^{\otimes N}) < \infty$ for $e = 0, T$.

Lemma 23(b) yields the particle energy bound (EB) at $m_t = m_*$ with constants depending only on $d$, $\lambda$, and $K$, independently of $N$ and $T$. Theorem 2 therefore gives the following proposition.

**Proposition 7** (Entropic turnpike for displacement convex costs)**.** *Under Assumption Cost, (20), and (21), there exist $c$, $C > 0$, depending only on $d$, $\lambda$, and $K$, such that for all $N \geqslant 1$, $T \geqslant c^{-1}$, and endpoint laws $m_0^N$, $m_T^N \in \mathcal{P}_2(\mathbb{R}^{dN})$ with $\mathcal{H}(m_e^N|m_*^{\otimes N}) < \infty$, $e = 0, T$, the unique finite-cost optimizer $P_{\mathrm{opt}}^N$ of (2) with $m_0 = m_T = m_*$ satisfies, for all $t \in [0,T]$,*

$$\mathcal{H}(m_t^N|m_*^{\otimes N}) \leqslant C\big(e^{-ct}\mathcal{H}(m_0^N|m_*^{\otimes N}) + e^{-c(T-t)}\mathcal{H}(m_T^N|m_*^{\otimes N}) + 1\big),$$

*where $m_t^N = P_{\mathrm{opt},t}^N$.*

*Remark* 2. Under the present assumptions, a regular ergodic corrector exists and the particle inequality (PL) can be verified directly, but the resulting constant is not guaranteed to approach the optimal mean-field constant. Obtaining near-optimal rates in Theorem 3 through our transfer argument requires $(M, L)$-regularity of $F$ and $\chi$, whose verification on $\mathbb{R}^d$ presents the same difficulty as in Remark 1.

# 4 Relative formulations

We prove Theorem 1 by first changing the path reference to $P_{\mathrm{ref}}^{\otimes N}$ and then passing between path laws and marginal flows. The change of reference produces the

modulated interaction; the time-symmetric entropy identity gives the kinetic and Fisher information terms of the relative action (9). Together they identify the three minimization problems up to fixed endpoint terms and give local optimality of finite-cost minimizing controls.

We apply the diffusion results of Appendix C with $S = P_{\mathrm{ref}}^{\otimes N}$. For one particle, take $W = R$, $V_r(x) = \frac{1}{2}\delta F_r(m_r, x)$ and endpoint factors $a_0$, $a_T$ in Assumption Diff. Assumption Ref gives the finite path second moment and entropy; the bound on $V$ follows from Assumption Cost and the bounded second moments of the reference flow. Closure under restriction and tensor products then covers every $P_{\mathrm{ref},[s,t]}^{\otimes N}$ used below.

*Proof of Theorem 1. (a) Change of reference.* The finite relative entropies in Assumption End make the endpoint factors $a_0$, $a_T$ finite and positive at the given particle endpoints. The same assumption makes their logarithms absolutely integrable; the running potential is integrable by its quadratic bound and the path second moment. Changing reference in (5) in Section 2.2 therefore gives

$$2\mathcal{H}(P^N|R^{\otimes N}) = 2\mathcal{H}(P^N|P_{\mathrm{ref}}^{\otimes N}) + \mathbb{E}_{P^N}\sum_{i=1}^{N}\Big(\varphi_0(X_0^i) + \varphi_T(X_T^i) - \int_0^T \delta F_r(m_r, X_r^i)\,\mathrm{d}r\Big),$$

also for infinite entropies. In particular, the two entropies are finite simultaneously. By definition of the modulated interaction,

$$NF_r(\mu_{\boldsymbol{x}}) = NF_r(\mu_{\boldsymbol{x}}|m_r) + NF_r(m_r) + \sum_{i=1}^{N}\delta F_r(m_r, x^i) - N\langle \delta F_r(m_r,\cdot), m_r\rangle.$$

The finite cost of $P_{\mathrm{ref}}$ gives $r \mapsto F_r(m_r) \in L^1(0,T)$. Thus the two running integrals differ by an integrable term, by the quadratic bound on $\delta F_r(m_r,\cdot)$ and the path second moment. Their positive parts are finite by Assumption Cost and the same bounds. Hence the two objectives have the same effective domain. On that domain the linear running terms cancel, leaving the constant

$$\sum_{e\in\{0,T\}}\int_{\mathcal{X}^N}\sum_{i=1}^{N}\varphi_e(x^i)\,m_e^N(\mathrm{d}\boldsymbol{x}) + N\int_0^T \big(F_r(m_r) - \langle \delta F_r(m_r,\cdot), m_r\rangle\big)\,\mathrm{d}r.$$

Every term is finite by Assumptions End and Cost and the finite cost of $P_{\mathrm{ref}}$, and depends only on the given endpoints and the reference, as claimed.

*(b)–(d) Path laws and marginal flows.* For (b), fix $P^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$ with finite entropy relative to $P_{\mathrm{ref}}^{\otimes N}$. Corollary 26 in Appendix C gives a Markov law $\bar{P}^N$ with the same marginals and no larger relative entropy. These marginals are $W_2$ continuous because $P^N$ has a finite path second moment. Lemma 27(i) gives their relative velocity, the continuity equation and the finite kinetic and Fisher energies. The resulting control is admissible, and the time-symmetric entropy identity gives (10). If $P^N$ is already Markov, take $\bar{P}^N = P^N$ to obtain equality.

Conversely, for (c), let $(\nu^N, w^N)$ be any admissible control on $[0,T]$. Its continuity equation and finite energy are the hypotheses of Lemma 27(ii). The resulting

Markov law has the prescribed marginals and relative velocity $w^N$, finite entropy relative to $P_{\mathrm{ref}}^{\otimes N}$, and drift

$$(b_{\mathrm{ref},r})^{\oplus N} + w_r^N + \nabla \log \frac{\nu_r^N}{m_r^{\otimes N}}.$$

The time-symmetric identity gives equality in (10). Moreover, $W_2$ continuity gives finite endpoint and integrated second moments, so Lemma 24(ii) places this law in $\mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$.

For (d), fix $0 \leqslant s < t \leqslant T$ and an admissible control on $[s,t]$. Use the restricted reference $P_{\mathrm{ref},[s,t]}^{\otimes N}$ in Lemma 27(ii). Its entropy chain rule holds between every pair of times in the closed interval, including the actual endpoint entropies.

*Equivalence and local optimality.* Assume now that the endpoints satisfy Assumption End. The endpoint logarithms are then integrable, so Lemma 24(ii) also gives finite entropy relative to $R^{\otimes N}$ for the law realized in (c). Adding the modulated interaction to the time-symmetric identity gives

$$2\mathcal{H}(P^N|P_{\mathrm{ref}}^{\otimes N}) + N\int_0^T \mathbb{E}_{\nu_r^N} F_r(\mu_{\boldsymbol{X}}|m_r)\,\mathrm{d}r \\ = \mathcal{H}(\nu_0^N|m_0^{\otimes N}) + \mathcal{H}(\nu_T^N|m_T^{\otimes N}) + \mathcal{A}_N^{0,T}(\nu^N, w^N)$$

for a Markov realization. For any path law in the effective domain, the Markov law obtained in (b) has the same running interaction and no larger entropy. The two endpoint entropies are fixed. Thus (a) and this identity identify the three infima, including either infinite value, and the marginal flows of finite-cost minimizers.

By (d), every restriction of an admissible control remains admissible. Actions add under concatenation at matching endpoints, so replacing a segment of a finite-cost optimal control by a cheaper admissible segment would lower the full action. This proves local optimality. □

# 5 Entropic turnpike

We prove the entropic turnpike estimate of Theorem 2 along a possibly time-dependent reference flow, using the functional inequalities of Section 2.3.1. Accelerated flows give an upper bound on the minimal action in terms of the endpoint entropies and the upper-bound defect $\Delta_F^+$ in (EB), as well as a finite-cost competitor for the existence proof. Along the optimizer, this upper bound and the coercive lower bound give a discrete comparison for action averages. A weighted action bound then recovers the entropy at each time.

## 5.1 Connection via accelerated flows

Accelerated diffusion flows connect two endpoint laws through the product law $m_u^{\otimes N}$ at a common interior time $u$. The cost of this connection bounds the minimal action from above. Under the lower bound in (EB) the connecting control also serves as a finite-cost competitor for the particle Schrödinger problem.

**Lemma 8** (Connection to the reference flow)**.** *Let Assumptions Cost and Ref hold, together with* (LS) *and the upper bound in* (EB)*. Fix* $0 < s < u < T$ *and* $\nu_s^N \in$

*$\mathcal{P}_2(\mathcal{X}^N)$ with $\mathcal{H}(\nu_s^N|m_s^{\otimes N}) < \infty$, and set*

$$\omega_N = \sqrt{\rho^2 + \rho M_N}, \qquad q_r = \frac{\omega_N}{\rho}\coth\big(\omega_N(u-r)\big), \qquad s \leqslant r < u.$$

*There is an admissible control on $[s,u]$ with*

$$\tilde\nu_s^N = \nu_s^N, \qquad \tilde\nu_u^N = m_u^{\otimes N}, \qquad \tilde w_r^N = -q_r \nabla \log \frac{\tilde\nu_r^N}{m_r^{\otimes N}}, \qquad s \leqslant r < u.$$

*Writing $\tilde h_r = \mathcal{H}(\tilde\nu_r^N|m_r^{\otimes N})$, $\tilde\imath_r = \mathcal{I}(\tilde\nu_r^N|m_r^{\otimes N})$, we have*

$$\dot{\tilde h}_r = -q_r \tilde\imath_r,$$

*and*

$$\mathcal{A}_N^{s,u}(\tilde\nu^N, \tilde w^N) \leqslant q_s \tilde h_s + \Delta_F^+(u-s).$$

*The same assertions hold in reversed time, with $\check m_r = m_{T-r}$, $\check v_r = -v_{T-r}$ and $\check F_r = F_{T-r}$. Under $\nu_r^N = \check\nu_{T-r}^N$, $w_r^N = -\check w_{T-r}^N$, a control on $[T-t, T-u]$ returns to $[u,t]$ with the same action.*

*Proof of Lemma 8.* The relative velocity in the statement gives the diffusion equation

$$\partial_r \tilde\nu_r^N = q_r \Delta \tilde\nu_r^N - \nabla \cdot \big(\tilde\nu_r^N (v_r^{\oplus N} + q_r \nabla \log m_r^{\otimes N})\big).$$

Using the singularity of $q$ at $u$, we first obtain estimates that force its terminal value $\tilde\nu_u^N$ to be $m_u^{\otimes N}$. We then construct the flow and verify admissibility, first for bounded initial density ratios and then by approximation.

*Step 1: Action and energy estimates.* The coefficient $q$ satisfies

$$\dot q_r = \rho(q_r^2 - 1) - M_N, \quad q_r \geqslant 1, \quad q_r \to \infty \text{ as } r \uparrow u.$$

For this step, suppose that the flow exists on $[s,u)$ and satisfies the dissipation identity $\dot{\tilde h}_r = -q_r \tilde\imath_r$; Steps 2 and 3 construct such a flow. The relative action (9) in Section 2.2 for $(\tilde\nu_r^N, \tilde w_r^N)$ then has density

$$\mathcal{L}_r(\tilde\nu^N, \tilde w^N) = \frac{q_r^2 + 1}{2}\tilde\imath_r + N\,\mathbb{E}_{\tilde\nu_r^N}[F_r(\mu_{\boldsymbol{X}}|m_r)].$$

By (LS), the upper bound in (EB) and the dissipation identity, it satisfies

$$\begin{aligned}\mathcal{L}_r(\tilde\nu^N, \tilde w^N) + \frac{\mathrm{d}}{\mathrm{d}r}(q_r \tilde h_r) &\leqslant \frac{1 - q_r^2}{2}\tilde\imath_r + (M_N + \dot q_r)\tilde h_r + \Delta_F^+ \\ &\leqslant \big(\rho(1 - q_r^2) + M_N + \dot q_r\big)\tilde h_r + \Delta_F^+ = \Delta_F^+.\end{aligned} \tag{24}$$

Inequality (LS) also gives $\dot{\tilde h}_r \leqslant -2\rho q_r \tilde h_r$, and integrating gives

$$\tilde h_r \leqslant \tilde h_s \left(\frac{\sinh(\omega_N(u-r))}{\sinh(\omega_N(u-s))}\right)^2. \tag{25}$$

In particular, $\tilde h_r = O((u-r)^2)$ and $q_r \tilde h_r \to 0$. Integrating (24) gives the asserted action bound. For admissibility at $u$, we also use

$$\int_s^u q_r^2 \tilde\imath_r \,\mathrm{d}r = q_s \tilde h_s + \int_s^u \dot q_r \tilde h_r \,\mathrm{d}r < \infty. \tag{26}$$

The last integral is finite because $\dot{q}_r = O((u-r)^{-2})$ and $\tilde{h}_r = O((u-r)^2)$.

*Step 2: Construction with bounded density ratios.* First assume that the initial density ratio $\nu_s^N/m_s^{\otimes N}$ is bounded. On each compact subinterval of $[s,u)$, the reference marginals $m_r^{\otimes N}$ solve the diffusion equation above. The corresponding drift is $(b_{\mathrm{ref},r})^{\oplus N} + (q_r - 1)\nabla \log m_r^{\otimes N}$. It has finite energy under the reference marginals: the first term is controlled by the entropy $\mathcal{H}(P_{\mathrm{ref}}|R)$; for the score, apply (63) in Appendix C with $P = P_{\mathrm{ref}}$ and $S = R$, and split $\nabla \log m_r = \nabla \log(m_r/R_r) + \nabla \log R_r$. The heat-marginal score $\nabla \log R_r$ has at most linear growth on $[s,u]$, since $0 < s < u < T$, and $q$ is locally bounded on $[s,u)$.

The superposition principle [47, Theorem 2.5] gives a diffusion with these reference marginals and noise $\sqrt{2q_r}$. In the time variable $\tau(r) = \int_s^r q_\ell \,\mathrm{d}\ell$, the noise is $\sqrt{2}$. Brownian Girsanov and strict convexity of entropy give uniqueness for this fixed state drift. Its laws are therefore Markov and consistent on overlapping intervals. Changing the initial law gives the desired marginal flow. Its density ratio is a conditional expectation of the initial ratio, hence remains bounded.

Finite-entropy time reversal in the same time variable gives the stated dissipation identity [14, Theorem 1.16]. Step 1 now applies on compact intervals. Increasing them to $[s,u)$ gives (25) and (26). The original relative velocity is $\tilde{w}_r^N = -q_r \nabla \log(\tilde{\nu}_r^N/m_r^{\otimes N})$. Since $q \geqslant 1$, these estimates bound both terms in (8). For admissibility we also need continuity in $W_2$, which is a hypothesis of the path-law realization in Theorem 1(c). Tensorized Talagrand and the $W_2$ continuity of the reference give

$$W_2(\tilde{\nu}_r^N, m_u^{\otimes N}) \leqslant \sqrt{2\tilde{h}_r/\rho} + W_2(m_r^{\otimes N}, m_u^{\otimes N}) \to 0 \quad \text{as } r \uparrow u.$$

Bounded density ratios and the reference moment bound give $W_2$ continuity at all earlier times. Thus $\tilde{\nu}^N \in \mathcal{C}([s,u];\mathcal{P}_2(\mathcal{X}^N))$, with terminal value $m_u^{\otimes N}$, and the control satisfies (7) and (8).

*Step 3: General finite-entropy initial laws.* For an unbounded initial ratio $p = \mathrm{d}\nu_s^N / \mathrm{d}m_s^{\otimes N}$, set

$$Z_n = \int (p \wedge n)\,\mathrm{d}m_s^{\otimes N}, \qquad p_n = \frac{p \wedge n}{Z_n}, \qquad \nu_s^{N,n} = p_n m_s^{\otimes N}.$$

Then $Z_n \uparrow 1$, and $p_n \leqslant 2p$ for large $n$. The initial laws converge in $W_2$ and their relative entropies converge to $\tilde{h}_s$ by dominated convergence. The energy bounds in Step 1 are uniform over the truncations. To control moment tails, choose a convex $\Psi$ growing faster than $z \log z$ with $\int \Psi(2p)\,\mathrm{d}m_s^{\otimes N} < \infty$. Conditional Jensen bounds the $\Psi$-integrals of every evolving ratio, uniformly over these approximations. Inequality (LS) gives a uniform Gaussian moment for $m_r^{\otimes N}$ on $[s,u]$. For a sufficiently small exponent $\alpha > 0$, Young's inequality gives

$$\int_{|\boldsymbol{x}|>L} |\boldsymbol{x}|^2 p_r \,\mathrm{d}m_r^{\otimes N} \leqslant \frac{1}{\alpha} \int_{|\boldsymbol{x}|>L} (p_r \log p_r)^+ \,\mathrm{d}m_r^{\otimes N} + \frac{1}{\alpha} \int_{|\boldsymbol{x}|>L} e^{\alpha|\boldsymbol{x}|^2} \,\mathrm{d}m_r^{\otimes N}.$$

The $\Psi$ bound makes the first tail uniformly small; a Gaussian moment at a slightly larger exponent does the same for the second. The continuity equation gives local time compactness, so these tails upgrade marginal convergence to $W_2$ convergence, including both endpoint traces.

Take weak limits of the relative velocity fluxes and scores. On compact interior cylinders the weak gradient identity passes to the limit and identifies $\tilde{w}_r^N =$

$-q_r \nabla \log(\tilde{\nu}_r^N / m_r^{\otimes N})$. Lower semicontinuity, followed by increasing the cylinders, gives (8) on $[s,u]$. Theorem 1(d) now gives the entropy dissipation identity for the limiting admissible control. Thus the estimates of Step 1 hold with its actual initial entropy.

On interior intervals the reversed reference has current velocity $\check{v}_r = -v_{T-r}$ and forward drift $\check{v}_r + \nabla \log \check{m}_r$. Both terms have finite energy there by the drift and score bounds of Step 2. The same construction thus applies in reversed time. Reversing both the constructed flow and its reference changes the sign of the relative velocity and preserves the action. □

**Proposition 9** (Entropy bound on the minimal action)**.** *Let Assumptions Cost and Ref hold, together with* (LS) *and the upper bound in* (EB)*. For every* $0 \leqslant s < t \leqslant T$ *and every* $\nu_s^N$, $\nu_t^N \in \mathcal{P}_2(\mathcal{X}^N)$ *with* $\mathcal{H}(\nu_e^N | m_e^{\otimes N}) < \infty$ *for* $e = s,\, t$*,*

$$\mathsf{C}_N(s, \nu_s^N; t, \nu_t^N) \leqslant \left( \sqrt{1 + \frac{M_N}{\rho}} + \frac{2}{\rho(t-s)} \right) \big( \mathcal{H}(\nu_s^N | m_s^{\otimes N}) + \mathcal{H}(\nu_t^N | m_t^{\otimes N}) \big) + \Delta_F^+ (t-s). \quad (27)$$

*Proof of Proposition 9.* Fix $s < t$ and endpoints $\nu_s^N$, $\nu_t^N$, and put $u = (s+t)/2$. We connect each endpoint to $m_u^{\otimes N}$ and concatenate the two controls at $u$:

$$\nu_s^N \xrightarrow{\text{accelerated flow}} m_u^{\otimes N} \xleftarrow{\text{accelerated flow}} \nu_t^N.$$

For the terminal connection, we use the accelerated flow for the reversed reference and then return to the original time orientation. The coefficient in Lemma 8 is singular only at $u$, whereas its construction starts at an interior time. We therefore insert a short ordinary evolution before accelerating. We describe the left half. Write $h_e = \mathcal{H}(\nu_e^N | m_e^{\otimes N})$ for $e = s,\, t$.

Take $r_0 = s + \varepsilon$ with $0 < \varepsilon < u - s$ and start the restricted reference $(P_{\mathrm{ref}}^{\otimes N})_{[s,r_0]}$ from $\nu_s^N$. Corollary 28 in Appendix C applies also when $s = 0$. On $\mathbb{R}^{dN}$ its Gaussian moment condition follows from tensorized (LS) and the bounded reference second moments. The resulting flow $\nu^N$ is $W_2$ continuous, with relative velocity $-\nabla \log(\nu_r^N / m_r^{\otimes N})$. By (64), its relative entropy $h_r$ and Fisher information $i_r$ satisfy

$$h_{r_0} \leqslant h_s, \qquad \int_s^{r_0} i_r \, \mathrm{d}r = h_s - h_{r_0}.$$

The upper bound in (EB) bounds this short segment's action by $h_s - h_{r_0} + \varepsilon M_N h_s + \varepsilon \Delta_F^+$. Apply Lemma 8 from $\nu_{r_0}^N$ to $m_u^{\otimes N}$, with its coefficient $q$ terminating at $u$. By the action bound in Lemma 8, the total left action is at most

$$h_s + (q_{r_0} - 1) h_{r_0} + \varepsilon M_N h_s + \Delta_F^+ (u-s) \leqslant (q_{r_0} + \varepsilon M_N) h_s + \Delta_F^+ (u-s).$$

On the right half, use the terminal endpoint in Corollary 28 on $[t-\varepsilon, t]$ and the reversed connection of Lemma 8 on $[u, t-\varepsilon]$. Both flows have trace $m_u^{\otimes N}$ and are admissible in the original time orientation, so they can be concatenated. Letting $\varepsilon \downarrow 0$ gives

$$\mathsf{C}_N(s, \nu_s^N; t, \nu_t^N) \leqslant \frac{\omega_N}{\rho} \coth\!\left( \frac{\omega_N (t-s)}{2} \right) (h_s + h_t) + \Delta_F^+ (t-s).$$

Finally, $\coth x \leqslant 1 + x^{-1}$ gives (27). □

**Corollary 10** (Existence and uniqueness of the optimizer)**.** *Let Assumptions Cost, Ref, and End hold, together with* (LS) *and* (EB)*. For every $T > 0$, the particle Schrödinger problem* (2) *in Section 2.1 has a unique finite-cost optimizer. It is Markov and is symmetric if both endpoint laws are symmetric.*

*Proof of Corollary 10. Step 1: A finite-cost competitor.* The admissible control $(\nu^N, w^N)$ from Proposition 9 connects $m_0^N$ to $m_T^N$. Set $V_r(\boldsymbol{x}) = NF_r(\mu_{\boldsymbol{x}}|m_r)$ and $i_r = \mathcal{I}(\nu_r^N|m_r^{\otimes N})$. The lower bound in (EB), Assumption Cost and the reference moment bound give, for some $C < \infty$,

$$\mathbb{E}_{\nu_r^N} V_r \geqslant -\frac{1-\eta_N}{2} i_r - \Delta_F^-,$$
$$V_r^+(\boldsymbol{x}) \leqslant C\left(1 + \sum_i \langle x^i\rangle_{\mathcal{X}}^2\right) + N|F_r(m_r)|.$$

The finite cost of $P_{\mathrm{ref}}$ makes $r \mapsto F_r(m_r)$ integrable, and $W_2$ continuity bounds the second moments of the control. Thus $\int_0^T \mathbb{E}_{\nu_r^N} V_r^+ \,\mathrm{d}r < \infty$. Consequently,

$$\int_0^T \mathbb{E}_{\nu_r^N} V_r^- \,\mathrm{d}r \leqslant \int_0^T \mathbb{E}_{\nu_r^N} V_r^+ \,\mathrm{d}r + \frac{1-\eta_N}{2}\int_0^T i_r \,\mathrm{d}r + \Delta_F^- T < \infty.$$

The finite kinetic and Fisher energies therefore make its action finite. Theorem 1(c) realizes it by a Markov law $\bar{P}^N$ with finite entropy relative to $P_{\mathrm{ref}}^{\otimes N}$. The flow has a finite time-integrated second moment by $W_2$ continuity. Applying Lemma 24(ii) in Appendix C to the product Brownian tilt in Assumption Ref, with Assumption End verifying the endpoint logarithm hypothesis under $\bar{P}^N$, yields

$$\mathcal{H}(\bar{P}^N|R^{\otimes N}) < \infty, \qquad \bar{P}^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N)).$$

Theorem 1(a) then makes $\bar{P}^N$ a finite-cost competitor for the original problem.

*Step 2: Entropy minimization with a tilted reference.* Put $S = P_{\mathrm{ref}}^{\otimes N}$ and

$$G = \int_0^T V_r(\boldsymbol{X}_r)\,\mathrm{d}r.$$

Choose $p > 1$ with $p(1-\eta_N) < 1$. Combining (10) in Section 2.2 with the lower bound in (EB) gives

$$2\mathcal{H}(P^N|S) + p\,\mathbb{E}_{P^N} G \geqslant -p\Delta_F^- T$$

for every $P^N \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$ with finite entropy relative to $S$ and integrable $G$: the kinetic term, endpoint entropies and the remaining Fisher coefficient $1 - p(1-\eta_N)$ are nonnegative. By Assumption Cost and the bounded reference second moments, $V_r$ is locally bounded uniformly in time. Restrict $e^{-pG/2}S$ to paths contained in a sufficiently large compact spatial set and normalize it. Its density is bounded, so the preceding inequality applies. Its left-hand side is minus twice the logarithm of the normalizing constant, which is therefore at most $e^{p\Delta_F^- T/2}$. Exhausting $\mathcal{X}^N$ by compact sets and using monotone convergence gives

$$0 < Z_p := \mathbb{E}_S\, e^{-pG/2} \leqslant e^{p\Delta_F^- T/2} < \infty.$$

In particular, $Z = \mathbb{E}_S\, e^{-G/2}$ is finite and positive, and $\hat{S} = Z^{-1}e^{-G/2}S$ is a probability measure.

For path laws in the effective domain of the modulated problem, the exponential bound controls $\mathbb{E}\, G^-$ by relative entropy; the positive part is integrable by the second moments. Hence

$$2\mathcal{H}(P^N|S) + \mathbb{E}_{P^N}\, G = 2\mathcal{H}(P^N|\hat{S}) - 2\log Z.$$

The finite-cost competitor $\bar{P}^N$ has finite entropy relative to $\hat{S}$. Relative entropy is weakly lower semicontinuous with weakly compact sublevel sets, and the endpoint constraints are weakly closed. Hence there is a minimizer $Q$ of $\mathcal{H}(\cdot|\hat{S})$ with the given endpoints [36, Lemma 2.4]. Strict convexity of relative entropy on the convex constraint set makes this minimizer unique.

*Step 3: The original path class and uniqueness.* We must show that $Q$ has a finite path second moment, finite entropy relative to both $S$ and $R^{\otimes N}$, and an integrable running cost $G$. These properties put $Q$ in the effective domain of the original particle Schrödinger problem and justify changing the reference back. Tensorized (LS) and the bounded reference second moments give a uniform Gaussian moment of $m_r^{\otimes N}$ on $[0,T]$. Hölder's inequality and the bound on $Z_p$ give the same property for the marginals of $\hat{S}$, with a smaller exponent. The entropy inequality then bounds the second moments of $Q_r$ uniformly in $r$, so $\mathbb{E}_Q\, G^+ < \infty$. Moreover, $\mathbb{E}_{\hat{S}} \exp((p-1)G^-/2) \leqslant (1+Z_p)/Z$, so $\mathbb{E}_Q\, G^- < \infty$ as well. Changing reference gives $\mathcal{H}(Q|S) < \infty$. Lemma 24(ii) in Appendix C, with the integrated second-moment bound and Assumption End, now gives $Q \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}^N))$ and $\mathcal{H}(Q|R^{\otimes N}) < \infty$. The displayed entropy identity and Theorem 1(a) show that $Q$ is the unique finite-cost optimizer of the original problem.

Markovization with reference $R^{\otimes N}$ preserves all marginals and does not increase the objective, by Corollary 26 in Appendix C; uniqueness therefore makes the optimizer Markov. If both endpoint laws are symmetric, permuting particle indices preserves the reference, objective and constraints, so uniqueness also makes the optimizer symmetric. □

## 5.2 Action lower bounds

A chain rule and a dissipation inequality give a weighted lower bound on the action. The lemma is stated for a general nonnegative quantity so that it applies both to entropy here and to free energy in Section 6. Weights that vanish at the outer endpoints give the pointwise estimate used below; retaining the endpoint terms will allow cancellation against the connection bound in the equilibrium proof.

**Lemma 11** (Weighted action bound)**.** *Along an admissible control $(\nu^N, w^N)$ on $[s,u]$, let $\phi \geqslant 0$ be absolutely continuous and let a vector field $\xi$ satisfy*

$$\dot{\phi}_r = \int \xi_r \cdot w_r^N \,\mathrm{d}\nu_r^N, \qquad \mathcal{D}_r = \frac{1}{2}\int |\xi_r|^2 \,\mathrm{d}\nu_r^N.$$

*Suppose that a measurable density $\ell$ satisfies, for some $\lambda > 0$ and $\Delta \geqslant 0$,*

$$\ell_r \geqslant \frac{1}{2}\int |w_r^N|^2 \,\mathrm{d}\nu_r^N + \mathcal{D}_r, \qquad \mathcal{D}_r \geqslant \lambda\phi_r - \Delta.$$

*Then for every differentiable weight $\theta$ with $|\theta| \leqslant 1$ and $\dot{\theta} = \lambda(1-\theta^2)$,*

$$\int_s^u \ell_r \,\mathrm{d}r \geqslant \theta(u)\phi_u - \theta(s)\phi_s - \frac{\Delta}{\lambda}\big(\theta(u) - \theta(s)\big).$$

*In particular, for $s < u$ and every $t \in [s,u]$,*

$$\phi_t \leqslant \frac{\Delta}{\lambda} + \frac{1}{\tanh(\lambda(u-s))} \int_s^u \ell_r \,\mathrm{d}r. \tag{28}$$

*Proof of Lemma 11.* It suffices to consider $\int_s^u \ell_r \,\mathrm{d}r < \infty$. The first lower bound on $\ell$ then makes $\int \xi_r \cdot w_r^N \,\mathrm{d}\nu_r^N$ integrable in time by Cauchy–Schwarz. Completing the square gives

$$\begin{aligned}\ell_r - \theta(r)\dot{\phi}_r &\geqslant \frac{1}{2}\|w_r^N - \theta(r)\xi_r\|^2_{L^2(\nu_r^N)} + (1-\theta(r)^2)\mathcal{D}_r \\ &\geqslant \dot{\theta}(r)\phi_r - \Delta(1-\theta(r)^2).\end{aligned}$$

Since $\dot{\theta} = \lambda(1-\theta^2)$, integration gives the weighted bound with defect $(\Delta/\lambda)(\theta(u) - \theta(s))$.

For the pointwise bound, take $\theta_-(r) = \tanh(\lambda(r-s))$ on $[s,t]$ and $\theta_+(r) = \tanh(\lambda(r-u))$ on $[t,u]$. The outer endpoint terms vanish, so adding the two bounds gives

$$\int_s^u \ell_r \,\mathrm{d}r \geqslant \big(\tanh(\lambda(t-s)) + \tanh(\lambda(u-t))\big)\Big(\phi_t - \frac{\Delta}{\lambda}\Big).$$

The addition formula for tanh bounds the coefficient from below by $\tanh(\lambda(u-s)) > 0$. Dividing first by the coefficient and using $\int_s^u \ell_r \,\mathrm{d}r \geqslant 0$ gives (28). □

## 5.3 Discrete comparison

We prove a discrete comparison principle and apply it to action averages on neighboring intervals of $[0,T]$. The endpoint entropies serve as boundary values for the recurrence. The pointwise bound of Lemma 11 then converts the estimate for these averages into the entropic turnpike estimate of Theorem 2.

**Lemma 12** (Discrete window comparison)**.** *Let $n \geqslant 1$, $0 < \theta < 1/2$, and $\Delta \geqslant 0$. Suppose the nonnegative sequence $g_{-1}, g_0, \dots, g_n$ satisfies*

$$g_j \leqslant \theta(g_{j-1} + g_{j+1}) + \Delta, \qquad 0 \leqslant j < n.$$

*For every $\sigma > 0$ with $2\theta\cosh\sigma \leqslant 1$,*

$$g_j \leqslant e^{-\sigma(j+1)} g_{-1} + e^{-\sigma(n-j)} g_n + \frac{\Delta}{1-2\theta}, \qquad 0 \leqslant j < n.$$

*Proof of Lemma 12.* Use comparison for $g_j \leqslant \theta(g_{j-1} + g_{j+1}) + \Delta$ with fixed boundary values $g_{-1}$, $g_n$. The right-hand side defines a monotone contraction on the interior values, with constant at most $2\theta < 1$. Let $U_j$ be the asserted upper bound, also at $j = -1, n$. Its boundary values dominate $g_{-1}$, $g_n$, and $2\theta\cosh\sigma \leqslant 1$ gives

$$U_j \geqslant \theta(U_{j-1} + U_{j+1}) + \Delta.$$

Thus $g$ is a subsolution and $U$ a supersolution with ordered boundary values. Iteration of the monotone contraction gives $g \leqslant U$. □

*Proof of Theorem 2.* Existence, uniqueness and the stated Markov and symmetry properties follow from Corollary 10.

*Step 1: Entropy and action bounds.* We use the time scale

$$\tau = (4c_N)^{-1} = \frac{2\sqrt{1 + M_N/\rho}}{\eta_N \rho}.$$

Denote the relative velocity of the optimizer by $w^N_{\mathrm{opt},r}$. Write $h_r = \mathcal{H}(m_r^N | m_r^{\otimes N})$ and $i_r = \mathcal{I}(m_r^N | m_r^{\otimes N})$. The action density is

$$\mathcal{L}_r = \frac{1}{2}\int |w^N_{\mathrm{opt},r}|^2 \,\mathrm{d}m_r^N + \frac{1}{2} i_r + N\,\mathbb{E}_{m_r^N}[F_r(\mu_{\boldsymbol{X}} | m_r)].$$

The log-Sobolev inequality (LS) and the lower bound in (EB) give $\mathcal{L}_r + \Delta_F^- \geqslant \frac{1}{2}\int |w^N_{\mathrm{opt},r}|^2 \,\mathrm{d}m_r^N + \eta_N \rho h_r/2 \geqslant \eta_N \rho h_r$. By Theorem 1, the optimal control is admissible and minimizes the action between its endpoints on every subinterval. The entropy chain rule gives finite entropy at both endpoints of each segment. The coefficient in (27) is at most $2\sqrt{1 + M_N/\rho}$ for $u - s \geqslant 2\tau$. Proposition 9 therefore gives

$$\int_s^u \mathcal{L}_r \,\mathrm{d}r \leqslant 2\sqrt{1 + \frac{M_N}{\rho}}\,(h_s + h_u) + \Delta_F^+ (u - s).$$

*Step 2: Discrete comparison.* Let $n = \lfloor T/(4\tau) \rfloor \geqslant 1$ and $w = T/n$, so $4\tau \leqslant w \leqslant 8\tau$. On the time intervals $J_j = [jw, (j+1)w]$, $0 \leqslant j < n$, define

$$G_j = \frac{1}{\eta_N \rho w} \int_{J_j} (\mathcal{L}_r + \Delta_F^-) \,\mathrm{d}r.$$

For an interior cell $1 \leqslant j \leqslant n-2$, take $s \in J_{j-1}$ and $u \in J_{j+1}$. Then $J_j \subset [s, u]$ and $w \leqslant u - s \leqslant 3w$. Nonnegativity of $\mathcal{L} + \Delta_F^-$ and the upper action bound in Step 1, with $2\sqrt{1 + M_N/\rho} = \eta_N \rho \tau$, give

$$\eta_N \rho w G_j \leqslant \int_s^u (\mathcal{L}_r + \Delta_F^-) \,\mathrm{d}r \leqslant \eta_N \rho \tau (h_s + h_u) + 3(\Delta_F^- + \Delta_F^+) w.$$

Moreover, $\mathcal{L} + \Delta_F^- \geqslant \eta_N \rho h$ gives

$$\frac{1}{w}\int_{J_k} h_r \,\mathrm{d}r \leqslant G_k, \qquad k = j-1,\ j+1.$$

Averaging the action bound over $J_{j-1} \times J_{j+1}$ and dividing by $\eta_N \rho w$ therefore gives

$$G_j \leqslant \frac{\tau}{w}(G_{j-1} + G_{j+1}) + \frac{3(\Delta_F^- + \Delta_F^+)}{\eta_N \rho}, \qquad \frac{\tau}{w} \leqslant \frac{1}{4}.$$

For $n \geqslant 2$, at the first cell fix $s = 0$ and average over $u \in J_1$; the resulting bound is $G_0 \leqslant (\tau/w)(h_0 + G_1) + 2(\Delta_F^- + \Delta_F^+)/(\eta_N \rho)$. The last cell is treated in the same way with $u = T$. When $n = 1$, the bound on $[0, T]$ gives $G_0 \leqslant (\tau/w)(h_0 + h_T) + (\Delta_F^- + \Delta_F^+)/(\eta_N \rho)$. Thus the same inequality holds for all cells with $G_{-1} = h_0$ and $G_n = h_T$.

Take $\sigma = c_N w = w/(4\tau) \in [1, 2]$. The elementary inequality $\cosh \sigma \leqslant 2\sigma$ gives $2(\tau/w)\cosh \sigma \leqslant 1$. Applying Lemma 12 to the action averages with $\theta = \tau/w$ yields

$$G_j \leqslant e^{-c_N (j+1) w} h_0 + e^{-c_N (n-j) w} h_T + \frac{6(\Delta_F^- + \Delta_F^+)}{\eta_N \rho}.$$

*Step 3: Pointwise estimate.* Apply Lemma 11 on $J_j$ to $\sqrt{\eta_N}h$, with density $\mathcal{L}_r+\Delta_F^-$ and field $\sqrt{\eta_N}\nabla\log(m_r^N/m_r^{\otimes N})$. The entropy chain rule in Theorem 1(d) verifies the required chain rule on the closed interval $J_j$. Inequality (LS) gives $\eta_N i_r/2 \geqslant \eta_N\rho h_r$, and Step 1 supplies the action lower bound. Thus (28), with $\lambda=\sqrt{\eta_N}\rho$ and $\Delta=0$, gives for $t\in J_j$

$$h_t \leqslant \frac{\sqrt{\eta_N}\rho w}{\tanh(\sqrt{\eta_N}\rho w)}G_j \leqslant 16\coth(8)\sqrt{\frac{1+M_N/\rho}{\eta_N}}G_j,$$

because $w\leqslant 8\tau$ and $\sqrt{\eta_N}\rho w\geqslant 8$. Since $(j+1)w\geqslant t$ and $(n-j)w\geqslant T-t$, we obtain

$$h_t \leqslant 16\coth(8)\sqrt{\frac{1+M_N/\rho}{\eta_N}}\left(e^{-c_N t}h_0 + e^{-c_N(T-t)}h_T + \frac{6(\Delta_F^-+\Delta_F^+)}{\eta_N\rho}\right).$$

Since $96\coth(8)<100$, this proves (11) in Section 2.3.1. □

# 6 Free-energetic turnpike

Throughout this section Assumption Eq holds, and the functional inequalities are those of Section 2.3. To prove the free-energetic turnpike estimate of Theorem 3, we first express the relative action in terms of the free energy gradient. Gradient flow connections then give an upper bound on the minimal action. Combining it with the weighted lower bound of Lemma 11 yields a window estimate, from which a continuous comparison gives the claimed decay.

## 6.1 Free energy structure

For brevity in this section, write

$$F_* = F(\cdot|m_*), \qquad \chi_* = \chi(\cdot|m_*), \qquad \mathcal{E}_*(\nu) = \frac{1}{2}\mathcal{I}(\nu|m_*) + F_*(\nu).$$

Compatibility in Assumption Eq gives

$$D\chi_*(\nu,x) = D\chi(\nu,x) + \nabla\log m_*(x), \qquad D\chi_*(m_*,x)=0.$$

In particular, $\nabla\log m_*$ has linear growth and

$$|D\chi_*(\nu,x)| \leqslant C\big(\langle x\rangle_{\mathcal{X}} + M_2(\nu)^{1/2}\big).$$

Define the empirical potential and diagonal correction by

$$U_N(\boldsymbol{x}) = N\chi_*(\mu_{\boldsymbol{x}}), \qquad \mathcal{R}_N(\nu^N) = \mathbb{E}_{\nu^N}\int \operatorname{tr} D^2\chi(\mu_{\boldsymbol{X}},x,x)\,\mu_{\boldsymbol{X}}(\mathrm{d}x).$$

By Assumption Eq, $|\mathcal{R}_N|\leqslant C_\chi$.

**Proposition 13** (Free energy identities)**.** *Let Assumptions Cost, Ref, and Eq hold.*

*(a) For every $\nu \in \mathcal{P}_2(\mathcal{X})$,*

$$F_*(\nu) = \int \Big( \frac{1}{2}|D\chi_*(\nu,x)|^2 - \nabla_x \cdot D\chi_*(\nu,x) \\ - \nabla \log m_*(x) \cdot D\chi_*(\nu,x) \Big)\, \nu(\mathrm{d}x). \tag{29}$$

*If $\mathcal{I}(\nu|m_*) < \infty$, then*

$$\frac{1}{2}\|DF_*(\nu,\cdot)\|^2_{L^2(\nu)} = \mathcal{E}_*(\nu). \tag{30}$$

*(b) For every $\nu^N \in \mathcal{P}_2(\mathcal{X}^N)$ with finite relative Fisher information,*

$$\frac{1}{2}\int |DF_{*,N}(\nu^N,\cdot)|^2 \,\mathrm{d}\nu^N = \frac{1}{2}\mathcal{I}(\nu^N|m_*^{\otimes N}) + N\,\mathbb{E}_{\nu^N} F_*(\mu_{\boldsymbol{X}}) - \mathcal{R}_N(\nu^N). \tag{31}$$

*Along an admissible control $(\nu^N, w^N)$, the action density from (9) in Section 2.2 is therefore*

$$\mathcal{L}_r(\nu^N, w^N) = \frac{1}{2}\int \big(|w_r^N|^2 + |DF_{*,N}(\nu_r^N,\cdot)|^2\big)\,\mathrm{d}\nu_r^N + \mathcal{R}_N(\nu_r^N). \tag{32}$$

*This representation is invariant under time reversal.*

*(c) Along every admissible control $(\nu^N, w^N)$ on $[s,t]$, $r \mapsto \mathcal{F}_{*,N}(\nu_r^N)$ is absolutely continuous, and for all $s \leqslant a \leqslant b \leqslant t$,*

$$\mathcal{F}_{*,N}(\nu_b^N) - \mathcal{F}_{*,N}(\nu_a^N) = \int_a^b \int DF_{*,N}(\nu_r^N,\cdot)\cdot w_r^N \,\mathrm{d}\nu_r^N \,\mathrm{d}r. \tag{33}$$

*Proof of Proposition 13. (a) Mean-field identities.* The stationary-reference assertion of Lemma 24(i) in Appendix C gives $v_r = 0$ and $b_{\mathrm{ref},r} = \nabla \log m_*$. Thus its backward Feynman–Kac factor is $h_r^+ = c_r\sqrt{m_*}$. Its equation with potential $\delta F(m_*,\cdot)/2$ gives, modulo constants,

$$\delta F(m_*,\cdot) = \Delta \log m_* + \frac{1}{2}|\nabla \log m_*|^2.$$

Substituting $D\chi(\nu,\cdot) = D\chi_*(\nu,\cdot) - \nabla \log m_*$ into (12) from Section 2.3.2 expresses $F(\nu)$ as the right-hand side of (29) plus an affine function of $\nu$. At $\nu = m_*$, $D\chi_*(m_*,\cdot) = 0$, so this affine function is $F(m_*) + \langle \delta F(m_*,\cdot), \nu - m_*\rangle$. Subtracting it proves (29). The continuity and growth bounds in Assumption Eq justify the calculation for every $\nu \in \mathcal{P}_2(\mathcal{X})$.

For finite relative Fisher information,

$$DF_*(\nu,\cdot) = \nabla \log(\nu/m_*) + D\chi_*(\nu,\cdot).$$

Weighted integration by parts gives

$$\int \nabla \log(\nu/m_*)\cdot D\chi_*(\nu,\cdot)\,\mathrm{d}\nu = -\int \big(\nabla_x \cdot D\chi_*(\nu,\cdot) + \nabla \log m_* \cdot D\chi_*(\nu,\cdot)\big)\,\mathrm{d}\nu.$$

The score term is integrable by Cauchy–Schwarz, and the other terms by the second moment and the corrector bounds. The cutoff calculation in (b) below also justifies this identity for weak scores. Expanding the square now gives (30).

*(b) Particle identities.* The empirical derivative formulas give

$$\nabla_i U_N(\boldsymbol{x}) = D\chi_*(\mu_{\boldsymbol{x}}, x^i),$$
$$\nabla_i \cdot D\chi_*(\mu_{\boldsymbol{x}}, x^i) = \nabla_x \cdot D\chi_*(\mu_{\boldsymbol{x}}, x^i) + \frac{1}{N} \operatorname{tr} D^2\chi(\mu_{\boldsymbol{x}}, x^i, x^i).$$

Write $\operatorname{div}_{m_*^{\otimes N}} \nabla U_N = \Delta U_N + \nabla \log m_*^{\otimes N} \cdot \nabla U_N$. Summing over $i$ and applying (29) at $\mu_{\boldsymbol{x}}$ yields

$$\operatorname{div}_{m_*^{\otimes N}} \nabla U_N = \frac{1}{2}|\nabla U_N|^2 - NF_*(\mu_{\boldsymbol{x}}) + \int \operatorname{tr} D^2\chi(\mu_{\boldsymbol{x}}, x, x)\, \mu_{\boldsymbol{x}}(\mathrm{d}x),$$

This weighted divergence has at most quadratic growth.

On $\mathbb{R}^{dN}$, take smooth cutoffs $\zeta_R$, equal to one on $B_R$, zero outside $B_{2R}$, with $|\nabla \zeta_R| \leqslant C/R$. Weak integration by parts gives

$$\int \zeta_R \nabla \log(\nu^N / m_*^{\otimes N}) \cdot \nabla U_N \,\mathrm{d}\nu^N = -\int \zeta_R \operatorname{div}_{m_*^{\otimes N}} \nabla U_N \,\mathrm{d}\nu^N - \int \nabla \zeta_R \cdot \nabla U_N \,\mathrm{d}\nu^N.$$

The mixed score term is integrable by Cauchy–Schwarz, and the divergence term by its quadratic growth. Since $\nabla U_N$ has linear growth, the remaining term satisfies

$$\left| \int \nabla \zeta_R \cdot \nabla U_N \,\mathrm{d}\nu^N \right| \leqslant C_N \int_{\{|\boldsymbol{x}| \geqslant R\}} (1 + |\boldsymbol{x}|^2) \,\mathrm{d}\nu^N \longrightarrow 0.$$

Thus we may let $R \to \infty$; on the torus no cutoff is needed. Expanding

$$D\mathcal{F}_{*,N}(\nu^N, \cdot) = \nabla \log(\nu^N / m_*^{\otimes N}) + \nabla U_N$$

proves (31) and (32).

*(c) Free energy chain rule.* Testing the continuity equation against cutoffs of $U_N$ gives

$$\int U_N \,\mathrm{d}\nu_b^N - \int U_N \,\mathrm{d}\nu_a^N = \int_a^b \int \nabla U_N \cdot w_r^N \,\mathrm{d}\nu_r^N \,\mathrm{d}r.$$

Cauchy–Schwarz and the bounded second moments make the right-hand side integrable. The $W_2$ continuity of $\nu^N$ and quadratic growth of $U_N$ identify the endpoint integrals for every $a$, $b \in [s,t]$. Adding the entropy chain rule from Theorem 1(d) proves (33). □

For an admissible control on $[s,t]$, completing the square in (32) and using (33) gives

$$\mathcal{A}_N^{s,t}(\nu^N, w^N) = \mathcal{F}_{*,N}(\nu_s^N) - \mathcal{F}_{*,N}(\nu_t^N) + \int_s^t \left( \frac{1}{2} \int |w_r^N + D\mathcal{F}_{*,N}(\nu_r^N, \cdot)|^2 \,\mathrm{d}\nu_r^N + \mathcal{R}_N(\nu_r^N) \right) \mathrm{d}r. \tag{34}$$

Following the negative free energy gradient cancels the square.

**Lemma 14** (Free energy gradient flow)**.** *Let Assumptions Cost, Ref, and Eq hold. For every $\nu^N \in \mathcal{P}_2(\mathcal{X}^N)$ with $\mathcal{H}(\nu^N|m_*^{\otimes N}) < \infty$, there is a $W_2$-continuous flow $\tilde{\nu}^N$ starting from $\nu^N$ such that*

$$\tilde{w}_r^N = -D\mathcal{F}_{*,N}(\tilde{\nu}_r^N, \cdot)$$

*forms an admissible control with $\tilde{\nu}^N$ on every finite interval. For every $t \geqslant 0$,*

$$\mathcal{F}_{*,N}(\tilde{\nu}_t^N) + \int_0^t \int |D\mathcal{F}_{*,N}(\tilde{\nu}_r^N, \cdot)|^2 \, \mathrm{d}\tilde{\nu}_r^N \, \mathrm{d}r = \mathcal{F}_{*,N}(\nu^N). \tag{35}$$

*If $\nu^N$ is symmetric, the flow may be chosen symmetric.*

*Proof of Lemma 14.* Solve the Fokker–Planck equation

$$\partial_r \tilde{\nu}_r^N = \Delta \tilde{\nu}_r^N - \nabla \cdot \big(\tilde{\nu}_r^N (\nabla \log m_*^{\otimes N} - \nabla U_N)\big), \qquad \tilde{\nu}_0^N = \nu^N.$$

Its drift is locally Lipschitz and has linear growth by Assumption Eq. Truncating the drift and using the usual second-moment estimate gives a $W_2$-continuous solution with

$$\sup_{0 \leqslant r \leqslant t} M_2(\tilde{\nu}_r^N) \leqslant C_{N,t}\big(1 + M_2(\nu^N)\big), \qquad t < \infty.$$

In particular, $\int_0^t \int |\nabla U_N|^2 \, \mathrm{d}\tilde{\nu}_r^N \, \mathrm{d}r < \infty$.

Apply Lemma 25(ii) and Lemma 27(i) from Appendix C, with reference $P_{\mathrm{ref}}^{\otimes N}$ and drift perturbation $-\nabla U_N$, on an interval of length at most $T$. They give an admissible control with relative velocity

$$-\nabla \log(\tilde{\nu}_r^N / m_*^{\otimes N}) - \nabla U_N = -D\mathcal{F}_{*,N}(\tilde{\nu}_r^N, \cdot).$$

Their entropy chain rule also gives finite relative entropy at the terminal time. Repeating the argument from that law gives admissibility on every finite interval. The chain rule (33) now gives (35), including $t = 0$. Uniqueness for the locally Lipschitz drift and permutation invariance give symmetry when the initial law is symmetric. □

## 6.2 Connection via gradient flows

For the upper bound, we run a gradient flow from each endpoint and join the resulting laws using Proposition 9 in Section 5.1. Along either gradient flow, (34) identifies its action with the drop in free energy plus the diagonal correction. At the inner endpoints, (Ent) converts the connection's entropy bound into a free energy bound.

**Lemma 15** (Free energy bound on the minimal action)**.** *Let Assumptions Cost, Ref, and Eq hold, together with (LS), the upper bound in (EB), (Ent), and (PL) from Section 2.3. The latter may be restricted to symmetric joint laws. Set*

$$L = \lambda_N^{-1}, \qquad \Delta = \Delta_{\mathrm{rel}} + \lambda_N \Delta_{\mathrm{ent}}, \qquad q = \left( \frac{\sqrt{1 + M_N/\rho} + 2\lambda_N/\rho}{\alpha_N} - 1 \right)_+.$$

*For $0 \leqslant s < u \leqslant T$ with $u - s \geqslant L$, let $\nu_s^N$, $\nu_u^N \in \mathcal{P}_2(\mathcal{X}^N)$ have finite entropy relative to $m_*^{\otimes N}$, and suppose that both are symmetric if (PL) is assumed only for symmetric laws. Put $f_e = \mathcal{F}_{*,N}(\nu_e^N)$ for $e = s,\, u$. Then, for every $t \in [s,u]$,*

$$\begin{aligned}
&\mathsf{C}_N(s,\nu_s^N;u,\nu_u^N) + C_\chi(u-s)\\
&\quad\leqslant f_s + f_u + 2\Delta_{\mathrm{ent}} + qe^2\big(e^{-2\lambda_N(t-s)}(f_s+\Delta_{\mathrm{ent}}) + e^{-2\lambda_N(u-t)}(f_u+\Delta_{\mathrm{ent}})\big)\\
&\qquad + \frac{\Delta_F^+ + C_\chi + 2q\Delta}{\lambda_N} + 2C_\chi(u-s-L).
\end{aligned}$$

*Proof of Lemma 15.* Let $(\tilde\nu^{N,-},\tilde w^{N,-})$ and $(\tilde\nu^{N,+},\tilde w^{N,+})$ be the gradient flows of Lemma 14 starting from $\nu_s^N$ and $\nu_u^N$. Write $\tilde f_\pm(r) = \mathcal{F}_{*,N}(\tilde\nu_r^{N,\pm})$, where $r$ is elapsed time from the corresponding endpoint. The flows are symmetric when their initial laws are, so (PL) applies throughout. By (35),

$$\dot{\tilde f}_\pm(r) = -\int |D\mathcal{F}_{*,N}(\tilde\nu_r^{N,\pm},\cdot)|^2\,\mathrm{d}\tilde\nu_r^{N,\pm} \leqslant -2\lambda_N\tilde f_\pm(r) + 2\Delta_{\mathrm{rel}}.$$

Integrating gives

$$\begin{aligned}
\tilde f_-(r) + \Delta_{\mathrm{ent}} &\leqslant e^{-2\lambda_N r}(f_s+\Delta_{\mathrm{ent}}) + \Delta/\lambda_N,\\
\tilde f_+(r) + \Delta_{\mathrm{ent}} &\leqslant e^{-2\lambda_N r}(f_u+\Delta_{\mathrm{ent}}) + \Delta/\lambda_N.
\end{aligned} \tag{36}$$

Since their velocities cancel the square in (34) and $\mathcal{R}_N + C_\chi \leqslant 2C_\chi$,

$$\mathcal{A}_N^{0,r}(\tilde\nu^{N,\pm},\tilde w^{N,\pm}) + C_\chi r \leqslant \tilde f_\pm(0) - \tilde f_\pm(r) + 2C_\chi r.$$

This also holds trivially for $r = 0$.

Reserve a duration $L$ for the connection and take $\tau_- + \tau_+ = u - s - L$ with $\tau_\pm \geqslant 0$. We join the two gradient segments by a connection of duration $L$:

$$\nu_s^N \xrightarrow[\tau_-]{\text{gradient flow}} \tilde\nu_{\tau_-}^{N,-} \xleftarrow[L]{\text{connection}} \tilde\nu_{\tau_+}^{N,+} \xleftarrow[\tau_+]{\text{gradient flow}} \nu_u^N.$$

The last segment is traversed in reverse, which preserves the action by (32). Proposition 9 in Section 5.1 and (Ent) give

$$\begin{aligned}
&\mathsf{C}_N(s+\tau_-,\tilde\nu_{\tau_-}^{N,-};u-\tau_+,\tilde\nu_{\tau_+}^{N,+}) + C_\chi L\\
&\quad\leqslant \big(\sqrt{1+M_N/\rho} + 2\lambda_N/\rho\big)\sum_\pm \mathcal{H}(\tilde\nu_{\tau_\pm}^{N,\pm}|m_*^{\otimes N}) + (\Delta_F^+ + C_\chi)L\\
&\quad\leqslant (1+q)\big(\tilde f_-(\tau_-) + \tilde f_+(\tau_+) + 2\Delta_{\mathrm{ent}}\big) + (\Delta_F^+ + C_\chi)L.
\end{aligned}$$

The shifted free energies are nonnegative by (Ent). Adding the three segment bounds gives

$$\begin{aligned}
\mathsf{C}_N(s,\nu_s^N;u,\nu_u^N) + C_\chi(u-s) &\leqslant f_s + f_u + 2\Delta_{\mathrm{ent}}\\
&\quad + q\big(\tilde f_-(\tau_-) + \tilde f_+(\tau_+) + 2\Delta_{\mathrm{ent}}\big)\\
&\quad + (\Delta_F^+ + C_\chi)L + 2C_\chi(u-s-L)\\
&\leqslant f_s + f_u + 2\Delta_{\mathrm{ent}}\\
&\quad + q\big(e^{-2\lambda_N\tau_-}(f_s+\Delta_{\mathrm{ent}}) + e^{-2\lambda_N\tau_+}(f_u+\Delta_{\mathrm{ent}})\big)\\
&\quad + \frac{\Delta_F^+ + C_\chi + 2q\Delta}{\lambda_N} + 2C_\chi(u-s-L),
\end{aligned}$$

where the last inequality uses (36). For the given $t$, choose

$$\tau_- = \Big(1 - \frac{L}{u-s}\Big)(t-s), \quad \tau_+ = \Big(1 - \frac{L}{u-s}\Big)(u-t).$$

Then $\tau_- \geqslant t - s - L$ and $\tau_+ \geqslant u - t - L$. Since $\lambda_N L = 1$, their exponentials are at most $e^2 e^{-2\lambda_N(t-s)}$ and $e^2 e^{-2\lambda_N(u-t)}$, respectively. Substitution proves the bound. □

## 6.3 Continuous comparison

We first prove a continuous comparison principle for window estimates. In the proof of Theorem 3, the gradient flow connection and the weighted action lower bound give such an estimate for the shifted free energy. The comparison principle first gives a coarse estimate on the full interval. An asymmetric window comparison then improves its endpoint prefactor on long intervals.

**Lemma 16** (Continuous window comparison)**.** *Let $g \geqslant 0$ be continuous on $[0,T]$, and let $\lambda$, $L > 0$, $p \geqslant 1$, and $\Delta$, $\Omega \geqslant 0$, with $T \geqslant L$. Suppose*

$$g_t \leqslant p\big(e^{-2\lambda(t-s)} g_s + e^{-2\lambda(u-t)} g_u\big) + \Delta + \Omega(u-s)$$

*whenever $0 \leqslant s \leqslant t \leqslant u \leqslant T$ and $u - s \geqslant L$. For $0 < \beta < \lambda$, set*

$$\ell = \max\left\{L, \frac{\log(4p)}{2(\lambda-\beta)}\right\}.$$

*Then*

$$g_t \leqslant 2p\big(e^{-2\beta t} g_0 + e^{-2\beta(T-t)} g_T\big) + 2\Delta + 4\ell\Omega.$$

*Proof of Lemma 16.* We use comparison for the window inequality. Set

$$U(t) = 2p\big(e^{-2\beta t} g_0 + e^{-2\beta(T-t)} g_T\big) + 2\Delta + 4\ell\Omega.$$

If $T \leqslant 2\ell$, applying the hypothesis on $[0,T]$ directly gives $g \leqslant U$. Suppose $T > 2\ell$ and choose

$$s(t) = \max(0, t-\ell), \qquad u(t) = \min(T, t+\ell).$$

These windows have length at least $\ell \geqslant L$. For a bounded function $h$ on $(0,T)$, extend it by the fixed boundary values $h(0) = g_0$, $h(T) = g_T$, and define

$$(\mathcal{T}h)(t) = pe^{-2\lambda(t-s(t))} h(s(t)) + pe^{-2\lambda(u(t)-t)} h(u(t)) + \Delta + \Omega(u(t) - s(t)).$$

The hypothesis says $g \leqslant \mathcal{T}g$ on $(0,T)$. Each unknown interior value is at distance $\ell$, so $\mathcal{T}$ is monotone and is a contraction in the supremum norm, with constant

$$2pe^{-2\lambda\ell} \leqslant \frac{1}{2}, \qquad pe^{-2(\lambda-\beta)\ell} \leqslant \frac{1}{4}.$$

It remains to check that $U$ is a supersolution. Its boundary values dominate $g_0$, $g_T$. At an interior window,

$$pe^{-2\lambda\ell} U(t \pm \ell) \leqslant U(t)/4,$$

while $\Delta + \Omega(u(t) - s(t)) \leqslant \Delta + 2\ell\Omega \leqslant U(t)/2$. Thus $\mathcal{T}U \leqslant U$ there. Near the left boundary, write $U = U_0 + U_T + C$, where $U_0$, $U_T$ are the two endpoint exponentials

and $C = 2\Delta + 4\ell\Omega$. The fixed endpoint contributes $pe^{-2\lambda t} g_0 \leqslant U_0(t)/2$, the interior neighbor contributes at most $U(t)/4$, and the defect is at most $C/2$. Their sum is

$$\frac{3}{4}U_0(t) + \frac{1}{4}U_T(t) + \frac{3}{4}C \leqslant U(t).$$

The right boundary is identical by time reversal. The comparison principle for the monotone contraction $\mathcal{T}$ therefore gives $g \leqslant U$: iterate $g \leqslant \mathcal{T}g$ and $\mathcal{T}U \leqslant U$ to their common fixed point. □

*Proof of Theorem 3. Step 1: Existence and symmetry.* The assumptions of the theorem imply the lower bound in (EB) from Section 2.3.1 with qualitative constants. Indeed, the linear growth of $\nabla U_N$ and tensorized Talagrand give

$$\mathcal{I}(\nu^N | m_*^{\otimes N}) \leqslant C\left(1 + \mathcal{H}(\nu^N | m_*^{\otimes N}) + \int |D\mathcal{F}_{*,N}(\nu^N, \cdot)|^2 \, \mathrm{d}\nu^N\right).$$

Conditions (Ent) and (PL) give

$$\alpha_N \mathcal{H}(\nu^N | m_*^{\otimes N}) \leqslant \mathcal{F}_{*,N}(\nu^N) + \Delta_{\mathrm{ent}} \leqslant \frac{1}{2\lambda_N} \int |D\mathcal{F}_{*,N}(\nu^N, \cdot)|^2 \, \mathrm{d}\nu^N + \frac{\Delta_{\mathrm{rel}}}{\lambda_N} + \Delta_{\mathrm{ent}}.$$

Consequently, for some $C \geqslant 1$,

$$\mathcal{I}(\nu^N | m_*^{\otimes N}) \leqslant C\left(1 + \int |D\mathcal{F}_{*,N}(\nu^N, \cdot)|^2 \, \mathrm{d}\nu^N\right).$$

Using (31) and $|\mathcal{R}_N| \leqslant C_\chi$, we obtain

$$\frac{1}{2}\mathcal{I}(\nu^N | m_*^{\otimes N}) + N \, \mathbb{E}_{\nu^N} F_*(\mu_{\boldsymbol{X}}) \geqslant \frac{1}{2C}\mathcal{I}(\nu^N | m_*^{\otimes N}) - \frac{1}{2} - C_\chi.$$

This is the lower bound in (EB) with $\eta_N = C^{-1}$ and finite defect. If (PL) is assumed only for symmetric laws, first obtain the lower bound in (EB) on that class. Symmetrization preserves the interaction and decreases Fisher information. Since the coefficient of Fisher information in the lower bound in (EB) is $-(1 - \eta_N)/2 \leqslant 0$, this bound therefore extends to all laws. Corollary 10 in Section 5.1 now gives existence, uniqueness, and the Markov property, as well as symmetric optimal marginals when the endpoints are symmetric. These qualitative constants are used only for existence; all subsequent uses of (PL) are on symmetric laws when that restriction is imposed.

*Step 2: Comparing the action bounds.* Let $L$, $\Delta$, and $q$ be as in Lemma 15, and write $f_r = \mathcal{F}_{*,N}(m_r^N)$ as in the theorem. Denote the optimal relative velocity by $w^N_{\mathrm{opt},r}$. Theorem 1 gives admissibility, finite endpoint entropies on every segment, and local optimality. Hence

$$\int_s^u \left(\mathcal{L}_r(m^N, w^N_{\mathrm{opt}}) + C_\chi\right) \mathrm{d}r = \mathsf{C}_N(s, m_s^N; u, m_u^N) + C_\chi(u - s).$$

For $u - s \geqslant L$, Lemma 15 bounds this quantity from above. To bound it from below, use (32), $\mathcal{R}_N + C_\chi \geqslant 0$, and

$$f_r + \Delta_{\mathrm{ent}} \geqslant 0, \qquad \frac{1}{2}\int |D\mathcal{F}_{*,N}(m_r^N, \cdot)|^2 \, \mathrm{d}m_r^N \geqslant \lambda_N(f_r + \Delta_{\mathrm{ent}}) - \Delta.$$

Together with the chain rule (33), these verify the hypotheses of Lemma 11 in Section 5.2 for $\phi = f + \Delta_{\mathrm{ent}}$, $\lambda = \lambda_N$ and density $\mathcal{L}_r(m^N, w^N_{\mathrm{opt}}) + C_\chi$. For $t \in (s,u)$, choose

$$\begin{aligned} \theta_-(r) &= \tanh(\lambda_N(r-t) + \operatorname{arctanh}(1/2)) \qquad \text{on } [s,t], \\ \theta_+(r) &= \tanh(\lambda_N(r-t) - \operatorname{arctanh}(1/2)) \qquad \text{on } [t,u]. \end{aligned}$$

Both weights satisfy $\dot\theta = \lambda_N(1-\theta^2)$ and $|\theta| \leqslant 1$, with $\theta_-(t) - \theta_+(t) = 1$. Adding the two weighted bounds and using

$$1 + \theta_-(s) = \frac{6}{3 + e^{2\lambda_N(t-s)}} \leqslant 6e^{-2\lambda_N(t-s)}, \qquad 1 - \theta_+(u) \leqslant 6e^{-2\lambda_N(u-t)}$$

gives

$$\begin{aligned} \int_s^u \big(\mathcal{L}_r(m^N, w^N_{\mathrm{opt}}) + C_\chi\big)\,\mathrm{d}r \geqslant\ & f_t + \Delta_{\mathrm{ent}} \\ & + \big(1 - 6e^{-2\lambda_N(t-s)}\big)(f_s + \Delta_{\mathrm{ent}}) \\ & + \big(1 - 6e^{-2\lambda_N(u-t)}\big)(f_u + \Delta_{\mathrm{ent}}) - \frac{4\Delta}{\lambda_N}. \end{aligned} \tag{37}$$

Each weight increases by at most two, giving the defect $4\Delta/\lambda_N$. Continuity includes $t = s$, $u$. Cancelling $f_s + f_u + 2\Delta_{\mathrm{ent}}$ against the upper bound yields

$$\begin{aligned} f_t + \Delta_{\mathrm{ent}} \leqslant\ & p\big(e^{-2\lambda_N(t-s)}(f_s + \Delta_{\mathrm{ent}}) + e^{-2\lambda_N(u-t)}(f_u + \Delta_{\mathrm{ent}})\big) \\ & + \frac{(2q+4)\Delta + \Delta_F^+ + C_\chi}{\lambda_N} + 2C_\chi(u-s), \end{aligned} \tag{38}$$

where $p = 6 + qe^2$ and we used $u - s - L \leqslant u - s$.

*Step 3: Continuous comparison.* Fix $0 < \beta_N < \lambda_N$. Since $p \geqslant 6$ and $\log(4p) > 2$, the window length in Lemma 16 is $\ell = \log(4p)/(2(\lambda_N - \beta_N)) \geqslant L$. Apply it to $f + \Delta_{\mathrm{ent}}$ using (38), with $\lambda = \lambda_N$, $\Omega = 2C_\chi$ and constant defect $((2q+4)\Delta + \Delta_F^+ + C_\chi)/\lambda_N$. We obtain

$$\begin{aligned} f_t + \Delta_{\mathrm{ent}} \leqslant\ & 2p\big(e^{-2\beta_N t}(f_0 + \Delta_{\mathrm{ent}}) + e^{-2\beta_N(T-t)}(f_T + \Delta_{\mathrm{ent}})\big) \\ & + \frac{(4q+8)\Delta + 2(\Delta_F^+ + C_\chi)}{\lambda_N} + \frac{4\log(4p)C_\chi}{\lambda_N - \beta_N}. \end{aligned}$$

The coefficients $2p$, $4q+8$, and $2$ are at most $16(1+q)$, and

$$\log(4p) \leqslant \log 24 + \frac{e^2}{6}q \leqslant 4(1+q).$$

Using $\Delta = \Delta_{\mathrm{rel}} + \lambda_N\Delta_{\mathrm{ent}}$ and

$$1 + q = \max\left\{1, \frac{\sqrt{1 + M_N/\rho} + 2\lambda_N/\rho}{\alpha_N}\right\}$$

gives the following preliminary estimate for the nonnegative shifted free energy $g_r = f_r + \Delta_{\mathrm{ent}}$:

$$g_r \leqslant A_N\big(e^{-2\beta_N r}g_0 + e^{-2\beta_N(T-r)}g_T + D_{N,\beta}\big), \qquad r \in [0,T]. \tag{39}$$

*Step 4: Improving the endpoint prefactor.* Fix $\varepsilon > 0$ and choose

$$K = 32A_N^2(1+\varepsilon^{-1})^3, \qquad \ell = \frac{\log(256A_N^3(1+\varepsilon^{-1})^4)}{2(\lambda_N - \beta_N)}.$$

The coarse estimate proves the conclusion whenever $C_{N,\varepsilon,T} = A_N$. Hence suppose $C_{N,\varepsilon,T} < A_N$. Then $\varepsilon < A_N - 1$ and $Ke^{-(\lambda_N-\beta_N)T} < 1$. For the asymmetric windows, take

$$R = \frac{\log(4A_N(1+\varepsilon^{-1})^2)}{2\lambda_N}, \qquad h = \frac{\varepsilon}{4(1+\varepsilon)}, \qquad m = h(1-h/4).$$

These choices satisfy $\ell \geqslant R + L$ and $K \geqslant e^{2\lambda_N(R+L)}$, so $T > 2(R+L)$.

We first obtain an asymmetric window estimate. For $u - t \geqslant R + L$, choose the Riccati weights in Lemma 11 with

$$\theta_-(t) = -1 + h, \qquad \theta_+(t) = -1 + h^2/4.$$

Their jump at $t$ is $m$. In the gradient-flow connection from the proof of Lemma 15, take durations $\tau_- = t - s + R$ and $\tau_+ = u - t - R - L$. Combining the weighted lower bound with this connection gives

$$\begin{aligned} g_t \leqslant (1+\varepsilon/2)e^{-2\lambda_N(t-s)}g_s + Ke^{-2\lambda_N(u-t)}g_u \\ + \frac{5}{2}A_N(1+\varepsilon^{-1})D_{N,\beta} + \frac{2C_\chi}{m}(u-s). \end{aligned} \tag{40}$$

Indeed, the two endpoint coefficients, after their exponentials are factored out, are bounded respectively by

$$\begin{aligned} \frac{1}{(1-h/2)(1-h/4)} + \frac{q}{m}e^{-2\lambda_N R} &\leqslant 1 + \frac{\varepsilon}{2}, \\ \frac{2(2-h^2/4)}{(h^2/4)m} + \frac{q}{m}e^{2\lambda_N(R+L)} &\leqslant K. \end{aligned}$$

Here $1/m \leqslant 8(1+\varepsilon^{-1})$, $q/m \leqslant A_N(1+\varepsilon^{-1})/2$, and the first term in the second coefficient is at most $2^{11}(1+\varepsilon^{-1})^3$. The constant defect is bounded by

$$\frac{(2q+4)\Delta + \Delta_F^+ + C_\chi}{\lambda_N m} \leqslant \frac{2q+5}{m}D_{N,\beta} \leqslant \frac{5}{2}A_N(1+\varepsilon^{-1})D_{N,\beta},$$

which proves (40).

By time reversal it suffices to consider $t \leqslant T/2$. Set

$$w = \min(\ell, T/2), \qquad s = (t-w)_+, \qquad u = t + w.$$

Then $w \geqslant R + L$ and $u - s \leqslant 2\ell$, so (40) applies. Use (39) at $u$, and also at $s$ when $s > 0$. If $w = T/2$, then $s = 0$, and the endpoint contributions are at most

$$\begin{aligned} \big(1 + \varepsilon/2 + A_N K e^{-(\lambda_N-\beta_N)T}\big)e^{-2\beta_N t}g_0 \\ + A_N K e^{-(\lambda_N-\beta_N)T}e^{-2\beta_N(T-t)}g_T. \end{aligned}$$

Both coefficients are at most $C_{N,\varepsilon,T}$. The propagated defect is at most $A_N D_{N,\beta}$, using $Ke^{-\lambda_N T} \leqslant 1$. If $w = \ell$, use

$$A_N K e^{-2(\lambda_N-\beta_N)\ell} = \frac{1}{8(1+\varepsilon^{-1})} \leqslant \frac{\varepsilon}{8},$$

and $1+\varepsilon/2 \leqslant A_N \leqslant K$. When $s=0$, the two endpoint coefficients are bounded by $1+5\varepsilon/8$ and $\varepsilon/8$; when $s>0$, both are at most

$$A_N(1+\varepsilon/2+K)e^{-2(\lambda_N-\beta_N)\ell} \leqslant \frac{1}{4(1+\varepsilon^{-1})} \leqslant \frac{\varepsilon}{4}.$$

Thus they are again at most $C_{N,\varepsilon,T}$, and the propagated defect is at most $A_N D_{N,\beta}$.

The remaining defect is therefore at most

$$\begin{aligned}
&A_N D_{N,\beta} + \frac{5}{2}A_N(1+\varepsilon^{-1})D_{N,\beta} + \frac{4C_\chi \ell}{m} \\
&\quad\leqslant 4A_N(1+\varepsilon^{-1})D_{N,\beta} + 32(1+\varepsilon^{-1})C_\chi \ell \\
&\quad\leqslant A_N(1+\varepsilon^{-1})D_{N,\beta}\left(4 + \frac{16}{A_N}\log(256A_N^3(1+\varepsilon^{-1})^4)\right) \\
&\quad\leqslant 20A_N(1+\varepsilon^{-1})^2 D_{N,\beta}.
\end{aligned}$$

Here $D_{N,\beta} \geqslant C_\chi/(\lambda_N-\beta_N)$ and $A_N \geqslant 16$; the latter gives

$$4 + \frac{16}{A_N}\log(256A_N^3(1+\varepsilon^{-1})^4) \leqslant 18 + 4\log(1+\varepsilon^{-1}) \leqslant 20(1+\varepsilon^{-1}).$$

This proves (15). Finally, $\alpha_N \mathcal{H}(m_t^N | m_*^{\otimes N}) \leqslant f_t + \Delta_{\mathrm{ent}}$ gives the entropy estimate. □

# 7 Transfer of functional inequalities

To transfer the mean-field inequalities of Section 2.4 to particle laws (Theorem 4), we use the conditional approximation of S. Wang [48, Section 4]. We first approximate nonlinear functionals of the empirical measure, with errors expressed as kernel quadratic forms. We then apply the mean-field inequalities to the conditional average or to the individual conditional laws and bound the errors by entropy and Fisher information.

For the approximation lemmas below, let $A$ be $(M_A, L_A)$-regular and fix a corresponding kernel $K_A$. Let $m \in \mathcal{P}_2(\mathcal{X})$ be arbitrary. Let $\boldsymbol{X} = (X^1, \dots, X^N)$ have joint law $\nu^N \in \mathcal{P}_2(\mathcal{X}^N)$, and put

$$\mathcal{G}_j = \sigma(X^1, \dots, X^j), \qquad \nu_j = \mathcal{L}(X^j | \mathcal{G}_{j-1}), \qquad \bar{\nu} = \frac{1}{N}\sum_{j\in[N]} \nu_j,$$

with $\mathcal{G}_0$ the trivial $\sigma$-algebra. For a symmetric kernel $U$, write

$$Q_U(\sigma) = \langle U, \sigma^{\otimes 2}\rangle, \qquad \sigma \in \mathcal{M}_0(\mathcal{X}).$$

The martingale argument of S. Wang [48, Lemma 13] gives, for every symmetric kernel $U$ of at most quadratic growth,

$$\mathbb{E}\, Q_U(\mu_{\boldsymbol{X}} - \bar{\nu}) = \frac{1}{N^2}\sum_j \mathbb{E}\, Q_U(\delta_{X^j} - \nu_j). \tag{41}$$

Centering the components of $K_A$ at a fixed point gives kernels of at most quadratic growth, without changing their quadratic forms on mass-zero measures.

**Lemma 17** (Taylor and Jensen bounds)**.** *For every $\nu$, $\nu' \in \mathcal{P}_2(\mathcal{X})$ and $\sigma \in \mathcal{M}_0(\mathcal{X})$,*

$$|A(\nu|\nu')| \leqslant \frac{1}{2} Q_{K_A}(\nu - \nu'),$$
$$|\langle \delta A(\nu) - \delta A(m), \sigma \rangle| \leqslant \sqrt{Q_{K_A}(\nu - m) Q_{K_A}(\sigma)}.$$

*Consequently, for probabilities $\nu_1, \dots, \nu_N$ and $\bar{\nu} = N^{-1} \sum_j \nu_j$,*

$$\Big| \sum_j A(\nu_j|m) - N A(\bar{\nu}|m) \Big| \leqslant \frac{1}{2} \Big( \sum_j Q_{K_A}(\nu_j - m) - N Q_{K_A}(\bar{\nu} - m) \Big).$$

*Proof of Lemma 17.* Polarization of (16) in Section 2.4, followed by rescaling the two arguments, gives, for every $\sigma$, $\tau \in \mathcal{M}_0(\mathcal{X})$,

$$|\langle \delta^2 A(\nu), \sigma \otimes \tau \rangle| \leqslant \sqrt{Q_{K_A}(\sigma) Q_{K_A}(\tau)}.$$

The integral Taylor formula reads

$$A(\nu|\nu') = \int_0^1 (1-t) \langle \delta^2 A(\nu' + t(\nu - \nu')), (\nu - \nu')^{\otimes 2} \rangle \, \mathrm{d}t.$$

Using (16) proves the first bound. Integrating the polarized bound along the segment from $m$ to $\nu$ proves the second.

For the last assertion, expansion at $\bar{\nu}$ gives

$$\sum_j A(\nu_j|m) - N A(\bar{\nu}|m) = \sum_j A(\nu_j|\bar{\nu}),$$

since $\sum_j (\nu_j - \bar{\nu}) = 0$ cancels the linear terms. The first bound therefore controls the absolute value by

$$\frac{1}{2} \sum_j Q_{K_A}(\nu_j - \bar{\nu}) = \frac{1}{2} \Big( \sum_j Q_{K_A}(\nu_j - m) - N Q_{K_A}(\bar{\nu} - m) \Big),$$

as claimed. □

We first approximate the empirical energy by evaluating the modulated functional at the conditional average $\bar{\nu}$.

**Lemma 18** (Approximation by the conditional average)**.** *For every $\theta > 0$,*

$$N \big| \mathbb{E}\, A(\mu_{\boldsymbol{X}}|m) - \mathbb{E}\, A(\bar{\nu}|m) \big| \leqslant \frac{\theta N}{2} \, \mathbb{E}\, Q_{K_A}(\bar{\nu} - m) + \frac{1 + \theta^{-1}}{2N} \sum_j \mathbb{E}\, Q_{K_A}(\delta_{X^j} - \nu_j).$$

*Proof of Lemma 18.* Expansion at $\bar{\nu}$ gives

$$A(\mu_{\boldsymbol{X}}|m) - A(\bar{\nu}|m) = A(\mu_{\boldsymbol{X}}|\bar{\nu}) + \langle \delta A(\bar{\nu}) - \delta A(m), \mu_{\boldsymbol{X}} - \bar{\nu} \rangle.$$

Lemma 17 bounds the absolute value of the remainder by $Q_{K_A}(\mu_{\boldsymbol{X}} - \bar{\nu})/2$ and that of the linear term by $\sqrt{Q_{K_A}(\bar{\nu} - m) Q_{K_A}(\mu_{\boldsymbol{X}} - \bar{\nu})}$. Young's inequality therefore yields, pointwise,

$$|A(\mu_{\boldsymbol{X}}|m) - A(\bar{\nu}|m)| \leqslant \frac{\theta}{2} Q_{K_A}(\bar{\nu} - m) + \frac{1 + \theta^{-1}}{2} Q_{K_A}(\mu_{\boldsymbol{X}} - \bar{\nu}).$$

Taking expectations, multiplying by $N$, and using (41) proves the claim. □

For a symmetric joint law we can instead evaluate the functional at each conditional law $\nu_j$ and sum the resulting energies.

**Lemma 19** (Approximation by individual conditional laws)**.** *Suppose $\nu^N$ is symmetric. For every $\varepsilon > 0$,*

$$\left| N\,\mathbb{E}\,A(\mu_{\boldsymbol{X}}|m) - \sum_j \mathbb{E}\,A(\nu_j|m) \right| \leqslant \frac{\varepsilon N}{2}\,\mathbb{E}\,Q_{K_A}(\nu_N - m) + \frac{5(1+\varepsilon^{-1})}{4N} \sum_j \mathbb{E}\,Q_{K_A}(\delta_{X^j} - \nu_j).$$

*Proof of Lemma 19.* For every $j \in [N]$, the tower rule and symmetry give

$$\mathbb{E}[\nu_N|\mathcal{G}_{j-1}] = \mathbb{E}[\delta_{X^N}|\mathcal{G}_{j-1}] = \mathbb{E}[\delta_{X^j}|\mathcal{G}_{j-1}] = \nu_j.$$

As in the proof of S. Wang [48, Lemma 14], we compare the quadratic energies of the conditional laws with that of their average. The tower identity yields

$$\mathbb{E}\langle K_A, (\nu_N - m) \otimes (\bar{\nu} - m)\rangle = \frac{1}{N} \sum_j \mathbb{E}\,Q_{K_A}(\nu_j - m).$$

Conditioning on $\mathcal{G}_{N-1}$ and then using symmetry gives

$$\begin{aligned} &\mathbb{E}\langle K_A, (\nu_N - m) \otimes (\mu_{\boldsymbol{X}} - m)\rangle \\ &\quad = \frac{1}{N} \sum_{j<N} \mathbb{E}\langle K_A, (\delta_{X^N} - m) \otimes (\delta_{X^j} - m)\rangle + \frac{1}{N}\,\mathbb{E}\,Q_{K_A}(\nu_N - m) \\ &\quad = \mathbb{E}\,Q_{K_A}(\mu_{\boldsymbol{X}} - m) - \frac{1}{N}\,\mathbb{E}\,Q_{K_A}(\delta_{X^N} - \nu_N). \end{aligned}$$

Denote the Jensen difference by

$$G = \sum_j \mathbb{E}\,Q_{K_A}(\nu_j - m) - N\,\mathbb{E}\,Q_{K_A}(\bar{\nu} - m) \geqslant 0.$$

Subtracting the two pairing identities and expanding $Q_{K_A}(\mu_{\boldsymbol{X}} - m)$ about $\bar{\nu} - m$, we obtain

$$\begin{aligned} G = {} & N\,\mathbb{E}\langle K_A, (2\bar{\nu} - \nu_N - m) \otimes (\mu_{\boldsymbol{X}} - \bar{\nu})\rangle \\ & + N\,\mathbb{E}\,Q_{K_A}(\mu_{\boldsymbol{X}} - \bar{\nu}) - \mathbb{E}\,Q_{K_A}(\delta_{X^N} - \nu_N). \end{aligned}$$

The last term is nonpositive. To bound the other two, put

$$u = \mathbb{E}\,Q_{K_A}(\nu_N - m), \qquad v = \mathbb{E}\,Q_{K_A}(\mu_{\boldsymbol{X}} - \bar{\nu}).$$

The tower identity above yields

$$\mathbb{E}\,Q_{K_A}(2\bar{\nu} - \nu_N - m) = u - \frac{4G}{N} \leqslant u,$$

while the tower rule and Jensen's inequality give $\mathbb{E}\,Q_{K_A}(\bar{\nu} - m) \leqslant u$. Cauchy–Schwarz therefore implies $G \leqslant N\sqrt{uv} + Nv$.

We now return to $A$. Expansion at $\bar{\nu}$ and Lemma 17 give

$$N\big|\mathbb{E}\,A(\mu_{\boldsymbol{X}}|m) - \mathbb{E}\,A(\bar{\nu}|m)\big| \leqslant N\sqrt{uv} + \frac{N}{2}v,$$
$$\left|\sum_j \mathbb{E}\,A(\nu_j|m) - N\,\mathbb{E}\,A(\bar{\nu}|m)\right| \leqslant \frac{G}{2}.$$

Combining these bounds and applying Young's inequality, we obtain

$$\begin{aligned}\left|N\,\mathbb{E}\,A(\mu_{\boldsymbol{X}}|m) - \sum_j \mathbb{E}\,A(\nu_j|m)\right| &\leqslant \frac{3N}{2}\sqrt{uv} + Nv \\ &\leqslant \frac{\varepsilon N}{2}u + \frac{5}{4}(1+\varepsilon^{-1})Nv.\end{aligned}$$

Finally, (41) gives

$$v = \frac{1}{N^2}\sum_j \mathbb{E}\,Q_{K_A}(\delta_{X^j} - \nu_j),$$

which proves the stated bound. □

*Proof of Theorem 4.* We relate the entropy and Fisher information of the joint law to those of the conditional laws defined above. Parts (a)–(c) use the conditional average $\bar{\nu}$, while part (d) uses each law separately. The entropy chain rule and conditional Jensen inequalities give

$$\mathcal{H}(\nu^N|m^{\otimes N}) = \sum_j \mathbb{E}\,\mathcal{H}(\nu_j|m) \geqslant N\,\mathbb{E}\,\mathcal{H}(\bar{\nu}|m), \tag{42}$$
$$\mathcal{I}(\nu^N|m^{\otimes N}) \geqslant \sum_j \mathbb{E}\,\mathcal{I}(\nu_j|m) \geqslant N\,\mathbb{E}\,\mathcal{I}(\bar{\nu}|m). \tag{43}$$

We make the computations for finite relative entropy and, in the estimates involving Fisher information, finite relative Fisher information.

Let $U$ be a symmetric positive kernel satisfying the bounds in the definition of $(M, L)$-regularity, with constants $M_U$, $L_U$, and put $\ell_U = L_U/\rho$. Centering each component against $m$ makes it positive on all signed measures without changing its quadratic form on mass-zero measures or these bounds. Since (LS) supplies the Talagrand and Poincaré inequalities with the same constant $\rho$, S. Wang [48, Lemma 15] gives

$$Q_U(\nu - m) \leqslant 2(M_U + \ell_U)\mathcal{H}(\nu|m), \tag{44}$$
$$\sum_j \mathbb{E}\,Q_U(\delta_{X^j} - \nu_j) \leqslant 4(M_U + d\ell_U)N + 4\ell_U\mathcal{H}(\nu^N|m^{\otimes N}). \tag{45}$$

Combining these bounds and (42) with Lemma 18, we obtain, for any $(M_A, L_A)$-regular functional $A$ and every $\theta > 0$,

$$\begin{aligned}N\big|\mathbb{E}\,A(\mu_{\boldsymbol{X}}|m) - \mathbb{E}\,A(\bar{\nu}|m)\big| \leqslant \left(\theta(M_A + \ell_A) + \frac{2(1+\theta^{-1})\ell_A}{N}\right)\mathcal{H}(\nu^N|m^{\otimes N}) \\ + 2(1+\theta^{-1})(M_A + d\ell_A).\end{aligned} \tag{46}$$

For part (a), the lower bound in (MF-EB) gives

$$F(\bar{\nu}|m) \geqslant -\frac{1-\eta}{2}\mathcal{I}(\bar{\nu}|m).$$

Since $0 < \eta \leqslant 1$, inequality (43) implies

$$\frac{1}{2}\mathcal{I}(\nu^N|m^{\otimes N}) + N\,\mathbb{E}\,F(\bar{\nu}|m) \geqslant \frac{\eta}{2}\mathcal{I}(\nu^N|m^{\otimes N}).$$

Applying (46) with $A = F$, we obtain

$$\begin{aligned}\frac{1}{2}\mathcal{I}(\nu^N|m^{\otimes N}) + N\,\mathbb{E}\,F(\mu_{\boldsymbol{X}}|m) &\geqslant \frac{\eta}{2}\mathcal{I}(\nu^N|m^{\otimes N}) \\ &\quad - \left(\theta(M_F + \ell_F) + \frac{2(1+\theta^{-1})\ell_F}{N}\right)\mathcal{H}(\nu^N|m^{\otimes N}) \\ &\quad - 2(1+\theta^{-1})(M_F + d\ell_F) \\ &\geqslant \frac{\eta_N}{2}\mathcal{I}(\nu^N|m^{\otimes N}) - \Delta_F^-,\end{aligned}$$

where the last inequality uses $\mathcal{H}(\nu^N|m^{\otimes N}) \leqslant \mathcal{I}(\nu^N|m^{\otimes N})/(2\rho)$.

For part (b), the upper bound in (MF-EB) and (42) give

$$N\,\mathbb{E}\,F(\bar{\nu}|m) \leqslant MN\,\mathbb{E}\,\mathcal{H}(\bar{\nu}|m) \leqslant M\mathcal{H}(\nu^N|m^{\otimes N}).$$

The same approximation bound yields

$$\begin{aligned}N\,\mathbb{E}\,F(\mu_{\boldsymbol{X}}|m) &\leqslant \left(M + \theta(M_F + \ell_F) + \frac{2(1+\theta^{-1})\ell_F}{N}\right)\mathcal{H}(\nu^N|m^{\otimes N}) \\ &\quad + 2(1+\theta^{-1})(M_F + d\ell_F) \\ &= M_N\mathcal{H}(\nu^N|m^{\otimes N}) + \Delta_F^+.\end{aligned}$$

For part (c), condition (MF-Ent) gives $\chi(\bar{\nu}|m) \geqslant -(1-\alpha)\mathcal{H}(\bar{\nu}|m)$. As $0 < \alpha \leqslant 1$, inequality (42) implies

$$\mathcal{H}(\nu^N|m^{\otimes N}) + N\,\mathbb{E}\,\chi(\bar{\nu}|m) \geqslant \alpha\mathcal{H}(\nu^N|m^{\otimes N}).$$

Applying (46) with $A = \chi$ and $\theta = \vartheta$ now gives

$$\begin{aligned}\mathcal{F}_N(\nu^N; m) &\geqslant \left(\alpha - \vartheta(M_\chi + \ell_\chi) - \frac{2(1+\vartheta^{-1})\ell_\chi}{N}\right)\mathcal{H}(\nu^N|m^{\otimes N}) \\ &\quad - 2(1+\vartheta^{-1})(M_\chi + d\ell_\chi) \\ &= \alpha_N\mathcal{H}(\nu^N|m^{\otimes N}) - \Delta_{\mathrm{ent}}.\end{aligned}$$

For part (d), let $\nu^N$ be symmetric. We apply (MF-PL) to each conditional law separately, so that (42) recovers the full joint entropy. The last-coordinate score and symmetry give

$$\mathcal{I}(\nu^N|m^{\otimes N}) = N\,\mathbb{E}\,\mathcal{I}(\nu_N|m).$$

Consequently, for any $(M_A, L_A)$-regular functional $A$ and a corresponding kernel $K_A$, the kernel bounds above imply

$$\begin{aligned}N\,\mathbb{E}\,Q_{K_A}(\nu_N - m) &\leqslant \frac{M_A + \ell_A}{\rho}\mathcal{I}(\nu^N|m^{\otimes N}), \\ \frac{1}{N}\sum_j \mathbb{E}\,Q_{K_A}(\delta_{X^j} - \nu_j) &\leqslant 4(M_A + d\ell_A) + \frac{2\ell_A}{\rho N}\mathcal{I}(\nu^N|m^{\otimes N}).\end{aligned}$$

The second bound also uses the tensorized log-Sobolev inequality.

Write

$$\mathcal{E}_N(\nu^N; m) = \frac{1}{2}\mathcal{I}(\nu^N|m^{\otimes N}) + N\,\mathbb{E}\,F(\mu_{\boldsymbol{X}}|m),$$

and put

$$M_\lambda = M_F + \lambda M_\chi, \quad \ell_\lambda = \ell_F + \lambda \ell_\chi.$$

Apply (MF-PL) to each $\nu_j$, take expectations, and sum. The entropy equality in (42) and the Fisher inequality in (43) give

$$\frac{1}{2}\mathcal{I}(\nu^N|m^{\otimes N}) + \sum_j \mathbb{E}\,F(\nu_j|m) \geqslant \lambda\mathcal{H}(\nu^N|m^{\otimes N}) + \lambda\sum_j \mathbb{E}\,\chi(\nu_j|m).$$

We replace the sums of conditional energies by their empirical counterparts using Lemma 19, separately for $F$ and $\chi$. Inserting the kernel bounds just obtained gives, for every $\varepsilon > 0$,

$$\begin{aligned}\mathcal{E}_N(\nu^N; m) \geqslant\ & \lambda\mathcal{F}_N(\nu^N; m) \\ & - \frac{\varepsilon(M_\lambda + \ell_\lambda) + 5(1+\varepsilon^{-1})\ell_\lambda/N}{2\rho}\mathcal{I}(\nu^N|m^{\otimes N}) \\ & - 5(1+\varepsilon^{-1})(M_\lambda + d\ell_\lambda).\end{aligned}$$

It remains to absorb the Fisher-information error into $\mathcal{E}_N(\nu^N; m)$. Choose $\theta > 0$ in part (a) so that $\eta_N > 0$. That part gives

$$\frac{\eta_N}{2}\mathcal{I}(\nu^N|m^{\otimes N}) \leqslant \mathcal{E}_N(\nu^N; m) + \Delta_F^-.$$

Thus, setting

$$z_N = \frac{\varepsilon(M_\lambda + \ell_\lambda) + 5(1+\varepsilon^{-1})\ell_\lambda/N}{\rho\eta_N},$$

we obtain

$$(1+z_N)\mathcal{E}_N(\nu^N; m) \geqslant \lambda\mathcal{F}_N(\nu^N; m) - 5(1+\varepsilon^{-1})(M_\lambda + d\ell_\lambda) - z_N\Delta_F^-.$$

Dividing by $1 + z_N$ gives the claimed inequality with

$$\lambda_N = \frac{\lambda}{1+z_N}, \qquad \Delta_{\mathrm{rel}} = \frac{5(1+\varepsilon^{-1})(M_\lambda + d\ell_\lambda) + z_N\Delta_F^-}{1+z_N}.$$

Fix $0 < \lambda' < \lambda$. Choose $\theta$ small enough that $\theta(M_F + \ell_F) < \eta\rho/2$, and then $N$ large enough that $\eta_N \geqslant \eta/2$. Choosing $\varepsilon$ small and increasing $N$ further makes $z_N \leqslant \lambda/\lambda' - 1$, hence $\lambda_N \geqslant \lambda'$. With $\theta$ and $\varepsilon$ fixed, $\Delta_{\mathrm{rel}}$ is bounded independently of $N$. □

# A Mean-field estimates on the torus

This appendix has two goals. First, under Sobolev bounds on an autonomous running cost and a smallness condition on the negative part of its second flat derivative, we obtain the unique ergodic state on $\mathbb{T}^d$ and construct a stationary corrector with the regularity required in Assumption Eq and Sections 2.4 and 3.1 (Proposition 20). Second, for possibly time-dependent running costs, we establish existence, the tilt

representation (5) in Assumption Ref, and uniform log-Sobolev bounds for mean-field Schrödinger bridges on $\mathbb{T}^d$, together with a uniqueness criterion that also applies on $\mathbb{R}^d$ (Lemma 21).

We normalize the volume of $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ to one. Fix $s > d/2$ and, for a distribution $\sigma$ on $\mathbb{T}^d$ with Fourier coefficients $\hat{\sigma}(k) = \langle e^{-2\pi i k\cdot x}, \sigma\rangle$, write

$$\|\sigma\|_{-s}^2 = \sum_{k\in\mathbb{Z}^d} (1 + 4\pi^2|k|^2)^{-s}|\hat{\sigma}(k)|^2, \qquad \mathcal{H}_0^{-s} = \{\sigma \in \mathcal{H}^{-s} : \langle 1, \sigma\rangle = 0\}.$$

Every finite signed measure belongs to $\mathcal{H}^{-s}$. Moreover, $\mathcal{P}(\mathbb{T}^d)$ is compact in this space, and its $\mathcal{H}^{-s}$ topology is the weak topology. Set

$$k_s(x, y) = \sum_{k\in\mathbb{Z}^d} (1 + 4\pi^2|k|^2)^{-s} e^{2\pi i k\cdot(x-y)}, \qquad \kappa_s = k_s(x, x) < \infty. \tag{47}$$

This is the reproducing kernel of $\mathcal{H}^s(\mathbb{T}^d)$. It is real, continuous, and positive, with $Q_{k_s}(\sigma) = \|\sigma\|_{-s}^2$ and $\operatorname{Osc}_2 k_s \leqslant 4\kappa_s$. Furthermore, for any $\nu$, $m \in \mathcal{P}(\mathbb{T}^d)$, Pinsker's inequality gives $|\hat{\nu}(k) - \hat{m}(k)| \leqslant \|\nu - m\|_{\mathrm{TV}} \leqslant \sqrt{2\mathcal{H}(\nu|m)}$ for $k \neq 0$, hence

$$\|\nu - m\|_{-s}^2 \leqslant 2\kappa_s \mathcal{H}(\nu|m). \tag{48}$$

The first flat derivative is defined modulo constants, and the second flat derivative modulo functions of either spatial argument. We equip $\mathcal{H}^s/\mathbb{R}$ with the norm of the representative with zero Lebesgue mean. For a twice flat differentiable functional $F$, write

$$A_\nu\sigma(x) = \langle \delta^2 F(\nu, x, \cdot), \sigma\rangle.$$

The operator $A_\nu$ represents the second flat derivative on mass-zero directions.

We say that $F$ has *Sobolev regularity* if it is twice continuously flat differentiable along probability segments, including segments with empirical endpoints, the maps

$$\nu \mapsto \delta F(\nu) \in \mathcal{H}^s/\mathbb{R}, \qquad \nu \mapsto A_\nu \in \mathcal{L}(\mathcal{H}_0^{-s}, \mathcal{H}^s/\mathbb{R})$$

are continuous for the $\mathcal{H}^{-s}$ topology and the indicated norm topologies, and there are bounds $C_F$, $B_F < \infty$ such that, for every $\nu \in \mathcal{P}(\mathbb{T}^d)$ and $\sigma \in \mathcal{H}_0^{-s}$,

$$\|\delta F(\nu)\|_{\mathcal{H}^s/\mathbb{R}} \leqslant C_F, \qquad \big|\langle A_\nu\sigma, \sigma\rangle\big| \leqslant B_F\|\sigma\|_{-s}^2.$$

**Proposition 20** (Regularity of the stationary corrector)**.** *For every $d$, $s$, $C_F$, $B_F$ as above there is $\kappa = \kappa(d, s, C_F, B_F) > 0$ with the following property. Let $F$ have the Sobolev regularity just described, and suppose that*

$$\langle A_\nu\sigma, \sigma\rangle \geqslant -\kappa\|\sigma\|_{-s}^2, \qquad \nu \in \mathcal{P}(\mathbb{T}^d), \quad \sigma \in \mathcal{H}_0^{-s}.$$

*Then the following conclusions hold.*

*(a) The stationary cost $\frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x) + F$ has a unique minimizer, the ergodic state $m_*$, satisfying $\log m_* \in \mathcal{H}^{s+2}(\mathbb{T}^d)$ and $\mathcal{I}(\cdot|m_*) \geqslant 2\rho_0\mathcal{H}(\cdot|m_*)$ for some $\rho_0 = \rho_0(d, s, C_F) > 0$ with $\kappa\kappa_s \leqslant \rho_0/2$. The normalized stationary corrector $\chi$ is finite and twice continuously flat differentiable, including at empirical measures, with $\chi(m_*) = 0$ and $\delta\chi(m_*) = -\log m_*$ modulo constants. There is $C = C(d, s, C_F, B_F)$ with the bounds below.*

*(b) The second flat derivative of $\chi$ satisfies*

$$\left|\langle \delta^2\chi(\nu), \sigma^{\otimes 2}\rangle\right| \leqslant C\|\sigma\|_{-s}^2, \qquad \nu \in \mathcal{P}(\mathbb{T}^d), \quad \sigma \in \mathcal{H}_0^{-s}. \tag{49}$$

*Moreover, for every $\nu \in \mathcal{P}(\mathbb{T}^d)$,*

$$\chi(\nu|m_*) \geqslant -\frac{1}{4\kappa_s}\|\nu - m_*\|_{-s}^2, \tag{50}$$

*and whenever $\mathcal{I}(\cdot|m_*) \geqslant 2\rho\mathcal{H}(\cdot|m_*)$ with $\rho > 0$ and $\mathcal{H}(\nu|m_*) < \infty$,*

$$\chi(\nu|m_*) \leqslant \frac{B_F\kappa_s}{2\rho}\mathcal{H}(\nu|m_*). \tag{51}$$

*(c) The first flat derivative is uniformly $\mathcal{C}^2$ in space, and the second flat derivative has a continuous mixed spatial derivative satisfying*

$$\sup_{\nu,x,y}\left|\nabla_x\nabla_y\delta^2\chi(\nu,x,y)\right| \leqslant C. \tag{52}$$

*Consequently $F$ is $(M_F, 0)$-regular and $\chi$ is $(M_\chi, 0)$-regular, with kernels $K_F = B_F k_s$ and $K_\chi = Ck_s$. In the notation of Section 2.4, one may take*

$$M_F = B_F\kappa_s, \qquad M_\chi = C\kappa_s, \qquad L_F = L_\chi = 0.$$

*The corrector $\chi$ satisfies (12) in Section 2.3.2, its empirical potential has the regularity in Assumption Eq, and (13) holds with $C_\chi \leqslant dC$.*

*If $F$ is flat convex, $\chi$ is flat convex and $\chi(\cdot|m_*)$ is nonnegative.*

*Proof of Proposition 20.* Throughout the proof, $C$ denotes a constant depending on $d$, $s$, $C_F$, $B_F$, which may change from line to line. We use the abbreviations $F_* = F(\cdot|m_*)$, $\chi_* = \chi(\cdot|m_*)$, and $\mathcal{E}_*(\nu) = \frac{1}{2}\mathcal{I}(\nu|m_*) + F_*(\nu)$.

*Step 1: The ergodic state.* Polarization and Sobolev duality give

$$\|A_\nu\sigma\|_{\mathcal{H}^s/\mathbb{R}} \leqslant B_F\|\sigma\|_{-s}.$$

Define the ergodic state by

$$m_* \in \operatorname*{argmin}_{\nu\in\mathcal{P}(\mathbb{T}^d)}\left\{\frac{1}{2}\mathcal{I}(\nu|\,\mathrm{d}x) + F(\nu)\right\}.$$

Existence follows from lower semicontinuity and compactness of bounded Fisher-information sublevels. Writing $w_* = \sqrt{m_*}$, its Euler equation is

$$-2\Delta w_* + \delta F(m_*)w_* = cw_*.$$

Since $\mathcal{H}^s \hookrightarrow \mathcal{C}^0$, the positive ground state and Harnack's inequality give uniform positive upper and lower bounds on $w_*$; see Gilbarg and Trudinger [29, Chapters 8–9]. Standard elliptic bootstrapping in the algebra $\mathcal{H}^s$ gives uniform bounds on $w_* \in \mathcal{H}^{s+2}$ and $\nabla\log m_* \in \mathcal{H}^{s+1}$. The bounded perturbation principle [32] therefore gives $\rho_0 = \rho_0(d, s, C_F) > 0$ such that every minimizing density satisfies $\mathcal{I}(\cdot|m_*) \geqslant 2\rho_0\mathcal{H}(\cdot|m_*)$. We use the equilibrium generator $L_* = \Delta + \nabla\log m_* \cdot \nabla$. Center flat derivatives against $m_*$. The Euler equation gives the exact modulation identity

$$\mathcal{E}_*(\nu) = \frac{1}{2}\mathcal{I}(\nu|m_*) + F_*(\nu) = \frac{1}{2}\mathcal{I}(\nu|\,\mathrm{d}x) + F(\nu) - \left(\frac{1}{2}\mathcal{I}(m_*|\,\mathrm{d}x) + F(m_*)\right).$$

Taylor's formula, the lower bound on the second flat derivative of $F$, and Pinsker's inequality give

$$F_*(\nu) \geqslant -\frac{\kappa}{2}\|\nu - m_*\|_{-s}^2 \geqslant -\kappa\kappa_s \mathcal{H}(\nu|m_*).$$

Consequently,

$$\mathcal{E}_*(\nu) \geqslant \frac{1}{2}\Big(1 - \frac{\kappa\kappa_s}{\rho_0}\Big)\mathcal{I}(\nu|m_*).$$

We choose $\kappa\kappa_s \leqslant \rho_0/2$. The modulation identity then shows that any other minimizer equals $m_*$.

*Step 2: The finite-horizon problem.* For $T > 0$, introduce the finite-horizon counterpart of $\chi_*$, the modulated value function

$$\chi_{*,T}(\nu) = \inf_{\tilde\nu_0 = \nu} \int_0^T \left(\frac{1}{2}\int |b_t|^2 \,\mathrm{d}\tilde\nu_t + F_*(\tilde\nu_t)\right) \mathrm{d}t, \qquad \partial_t \tilde\nu = L_*^* \tilde\nu - \nabla\cdot(\tilde\nu b).$$

Its infinite-horizon limit will be denoted by $\chi_*$; the original corrector $\chi$ is recovered by the affine correction in Step 11. Writing $m$ for an optimal flow, its optimality system, with $b = -\nabla u$, is

$$\begin{cases} -\partial_t u = L_* u - \frac{1}{2}|\nabla u|^2 + \delta F(m_t) - \delta F(m_*), & u_T = 0, \\ \partial_t m = L_*^* m + \nabla\cdot(m\nabla u), & m_0 = \nu. \end{cases}$$

The direct method applies in the flux variable $j = \tilde\nu b$, since $|j|^2/(2\tilde\nu)$ is convex and lower semicontinuous and $F_*(\cdot)$ is bounded and continuous. The displayed system is its standard first-order optimality condition; Steps 3 and 5–6 supply regularity and uniqueness. Singular initial probabilities are allowed by Lemma 25(ii) in Appendix C, using Brownian motion started at $\nu$, so that the initial relative entropy is zero.

*Step 3: Spatial regularity of the optimal drift.* Fix $\varepsilon$, depending only on $d$, $s$, such that

$$0 < \varepsilon < \min\{1/2, (s - d/2)/4\}.$$

The Hopf–Cole transform $w = e^{-u/2}$ satisfies

$$-\partial_t w = L_* w - \frac{1}{2}\big(\delta F(m_t) - \delta F(m_*)\big)w.$$

Standard heat-kernel bounds [2, Sections 7–8] bound $\operatorname{Osc} u_t$ uniformly on unit intervals, independently of the positive terminal datum; near $T$ use $u_T = 0$. With $\nabla \log m_* \in \mathcal{H}^{s+1}$ and $\delta F(m_t) \in \mathcal{H}^s/\mathbb{R}$ uniformly, Duhamel's formula and Sobolev multiplication then give

$$\sup_{T,\nu,\, 0\leqslant t\leqslant T} \|u_t\|_{\mathcal{H}^{s+2-\varepsilon}/\mathbb{R}} \leqslant C.$$

Here the smoothing factor $t^{-1+\varepsilon/2}$ is integrable; the heat-kernel bounds permit normalization of $w$ before composition with log. Sobolev embedding, since $s+2-\varepsilon > d/2+2$, gives

$$\sup_{T,\nu,\, 0\leqslant t\leqslant T} \|u_t\|_{\mathcal{C}^2/\mathbb{R}} \leqslant C.$$

In particular the controlled generator $L_t = \Delta + (\nabla \log m_* - \nabla u_t)\cdot\nabla$ has drift uniformly bounded in $\mathcal{H}^{s+1-\varepsilon}$, hence in $\mathcal{C}^{1,\beta}$ for some $\beta > 0$.

*Step 4: Semigroup estimates.* Write $P^*_{h,t}$ for its forward evolution. The standard exponential mixing estimates for bounded torus drifts [11, Section 1B], combined with unit-time smoothing and duality, give, for $\sigma \in \mathcal{H}_0^{-s}$,

$$\|P^*_{h,t}\sigma\|_{-s} \leqslant Ce^{-\omega(t-h)}\|\sigma\|_{-s}.$$

For short times, Duhamel's formula and Sobolev multiplication give boundedness on $\mathcal{H}^s$ and gains of $0 \leqslant \theta \leqslant 2-\varepsilon$ derivatives with norm $Ct^{-\theta/2}$, using Step 3 with loss $\varepsilon/2$. Set

$$\alpha = \max\{0, (1+d/2+\varepsilon-s)/2\} < \frac{1}{2}, \qquad g(t) = C(1+t^{-\alpha})e^{-\omega t}.$$

For $\varphi \in \mathcal{H}^s$, duality gives

$$\big\langle P^*_{h,t}\nabla\cdot(\nu_h q_h), \varphi\big\rangle = -\int q_h \cdot \nabla P_{h,t}\varphi \,\mathrm{d}\nu_h.$$

The gradient bound $\|\nabla P_{h,t}\varphi\|_\infty \leqslant g(t-h)\|\varphi\|_{\mathcal{H}^s}$ and Cauchy–Schwarz therefore give

$$\|P^*_{h,t}\nabla\cdot(\nu_h q_h)\|_{-s} \leqslant g(t-h)\|q_h\|_{L^2(\nu_h)}.$$

Here $g \in L^1 \cap L^2$, and $\nu_h$ can be any probability. The constants are common to the controlled drifts, their averages, and extensions by $L_*$ after the horizon. To quantify the response to a flux and to an initial perturbation, respectively, use

$$B = \|g\|_{L^1(0,\infty)}, \qquad C_0 = \sup_{\substack{L,\,\sigma\in\mathcal{H}_0^{-s}\\ \|\sigma\|_{-s}=1}} \int_0^\infty \|P^*_{0,t}\sigma\|^2_{-s}\,\mathrm{d}t < \infty,$$

where $L$ ranges over this common class of generators. The bounds on $\delta F$ and on the drift, hence $B$ and $C_0$, can be fixed before choosing $\kappa$. We may choose

$$0 < \kappa \leqslant \min\left\{\frac{\rho_0}{2\kappa_s}, \frac{1}{4B^2}, \frac{3}{8C_0\kappa_s}\right\}. \tag{53}$$

The last restriction yields (50) upon passing to the infinite-horizon limit in Steps 9 and 11.

*Step 5: The second-variation energy.* Fix an initial direction $\sigma \in \mathcal{H}_0^{-s}$. For a control $q \in L^2(m_t\,\mathrm{d}t)$, let $n$ denote the measure perturbation solving

$$\partial_t n = L_t^* n - \nabla\cdot(m_t q_t), \qquad n_0 = \sigma,$$

and set

$$Q(q) = \int_0^T \left(\int |q_t|^2 \,\mathrm{d}m_t + \langle A_{m_t} n_t, n_t\rangle\right)\mathrm{d}t.$$

Write $Q_0$ for the same quadratic action with zero initial datum. Separate the free evolution $\zeta_t = P^*_{0,t}\sigma$ from the response to the flux:

$$n = \zeta + \mathcal{K}q, \qquad (\mathcal{K}q)_t = -\int_0^t P^*_{h,t}\nabla\cdot(m_h q_h)\,\mathrm{d}h.$$

By the flux estimate of Step 4 and Young's inequality,

$$\|\mathcal{K}q\|_{L^2_t H^{-s}} \leqslant B\|q\|_{L^2(m_t\,\mathrm{d}t)}.$$

The lower bound on $A$ and the choice of $\kappa$ in (53) therefore give

$$Q_0(q) \geqslant (1-\kappa B^2)\|q\|^2_{L^2(m_t\,\mathrm{d}t)} \geqslant \frac{3}{4}\|q\|^2_{L^2(m_t\,\mathrm{d}t)}.$$

Thus $Q$ has a unique minimizer. Its first-order condition gives $q = -\nabla z$, where $(n, z)$ solves

$$\begin{cases} -\partial_t z = L_t z + A_{m_t} n_t, & z_T = 0, \\ \partial_t n = L_t^* n + \nabla\cdot(m_t \nabla z), & n_0 = \sigma. \end{cases}$$

This linearizes the optimality system of Step 2. Its backward representation is $z_t = \int_t^T P_{t,h}(A_{m_h} n_h)\,\mathrm{d}h$ modulo constants. Integration by parts gives

$$\langle z_0, \sigma\rangle = Q(-\nabla z) = \inf_q Q(q).$$

Completing the square, in the corresponding space and time $L^2$ norms, gives

$$Q(q) \geqslant \|q\|^2 - \kappa(\|\zeta\| + B\|q\|)^2 \geqslant -\frac{\kappa}{1-\kappa B^2}\|\zeta\|^2.$$

The competitor $q = 0$ gives $\inf Q \leqslant B_F\|\zeta\|^2$. Since $\|\zeta\|^2 \leqslant C_0\|\sigma\|^2_{-s}$, we obtain

$$-\frac{\kappa C_0}{1-\kappa B^2}\|\sigma\|^2_{-s} \leqslant \langle z_0, \sigma\rangle \leqslant B_F C_0 \|\sigma\|^2_{-s}.$$

The first-order condition and the coercivity of $Q_0$ also give $\|q\| \leqslant B_F B\|\zeta\|/(1-\kappa B^2)$, hence

$$\int_0^T \left(\|q_t\|^2_{L^2(m_t)} + \|n_t\|^2_{-s}\right)\mathrm{d}t \leqslant C\int_0^T \|\zeta_t\|^2_{-s}\,\mathrm{d}t.$$

These estimates hold for any $\zeta \in L^2_t H^{-s}_0$ and for averaged operators $A$ with the same bounds, including the singular directions used in Step 8.

*Step 6: Differentiability in the initial measure.* The envelope formula and its derivative are

$$\delta\chi_{*,T}(\nu) = u_0 \pmod{\mathbb{R}}, \qquad \langle \delta^2\chi_{*,T}(\nu), \sigma\otimes\tau\rangle = \langle z_0^\sigma, \tau\rangle.$$

To justify these identities, take two optimality systems $(m^j, u^j)$ starting at $\nu^j$, and set

$$\bar m = \frac{m^0+m^1}{2}, \quad \bar u = \frac{u^0+u^1}{2}, \quad n = m^1 - m^0, \quad z = u^1 - u^0,$$

and define

$$\bar A_t = \int_0^1 A_{(1-\theta)m_t^0 + \theta m_t^1}\,\mathrm{d}\theta, \qquad \bar L_t = \Delta + (\nabla \log m_* - \nabla \bar u_t)\cdot\nabla.$$

Their difference satisfies exactly

$$\begin{cases} -\partial_t z = \bar L_t z + \bar A_t n_t, & z_T = 0, \\ \partial_t n = \bar L_t^* n + \nabla\cdot(\bar m_t \nabla z), & n_0 = \nu^1 - \nu^0. \end{cases}$$

Step 5 applies with weight $\bar m_t$ and the same constants. For equal initial data, $Q_0(-\nabla z) = 0$ proves uniqueness by the coercivity bound of Step 5. In general, the energy bound and forward and backward representations give

$$\sup_{0\leqslant t\leqslant T}\|n_t\|_{-s} + \sup_{0\leqslant t\leqslant T}\|z_t\|_{\mathcal{H}^{s+2-\varepsilon}/\mathbb{R}} \leqslant C\|\nu^1 - \nu^0\|_{-s}.$$

The standard envelope argument, testing the drift optimal for $\nu$ at initial datum $\nu'$ and then reversing their roles, gives

$$\left|\chi_{*,T}(\nu') - \chi_{*,T}(\nu) - \langle u_0, \nu' - \nu\rangle\right| \leqslant C\|\nu' - \nu\|_{-s}^2.$$

Here Taylor's formula and the semigroup estimate of Step 4 control the remainder. As $\nu' \to \nu$, the averaged coefficients converge to the base coefficients. Cross-testing the averaged and base linear systems, as in Step 10, and using coercivity and Sobolev multiplication gives, with the linearized solution at $\nu$ in direction $\nu' - \nu$,

$$\|u_0^{\nu',T} - u_0^{\nu,T} - z_0^{\nu'-\nu}\|_{\mathcal{H}^s/\mathbb{R}} = o(\|\nu' - \nu\|_{-s}).$$

The same comparison for arbitrary initial directions gives operator-norm continuity of $\delta^2\chi_{*,T}$. This proves the envelope identities and continuity of the derivatives; compare the linearized representation in Cardaliaguet and Porretta [11, Proposition 5.3, equation (90)]. For singular initial laws, regularize in space: on $[h, H]$ the densities are uniformly positive, while the contribution of $[0, h]$ vanishes with $\|g\|_{L^2(0,h)}$. Thus only continuity of $A_\nu$ is needed.

*Step 7: Decay of the optimality systems.* Extend the coefficients after $T$ by $L_t = L_*$, $A_{m_t} = 0$, and let $m$ evolve without control there; extend $q$ by zero. Duhamel's formula, the two integrability properties of $g$, and the energy estimate of Step 5 give

$$\int_h^\infty \|n_t\|_{-s}^2 \,\mathrm{d}t + \sup_{t \geqslant h} \|n_t\|_{-s}^2 \leqslant C\|n_h\|_{-s}^2.$$

Indeed, the restricted control is still minimizing, so the energy bound applies from time $h$. Thus $Y(h) = \int_h^\infty \|n_t\|_{-s}^2 \,\mathrm{d}t$ satisfies $Y \leqslant -C\dot{Y}$. Exponential decay of $Y$, followed by the supremum bound averaged over $[h-1, h]$, gives pointwise decay of $n_h$. The backward representation and its integrable smoothing factor $(t-h)^{-1+\varepsilon/2}$ then give, for some $\gamma > 0$,

$$\|n_h\|_{-s} + \|z_h\|_{\mathcal{H}^{s+2-\varepsilon}/\mathbb{R}} \leqslant Ce^{-\gamma h}\|\sigma\|_{-s}.$$

Apply this also to the difference system of Step 6 and compare with the equilibrium solution $(m_*, 0)$ to obtain

$$\|m_h^{\nu,T} - m_*\|_{-s} + \|u_h^{\nu,T}\|_{\mathcal{H}^{s+2-\varepsilon}/\mathbb{R}} \leqslant Ce^{-\gamma h}\|\nu - m_*\|_{-s}, \qquad 0 \leqslant h \leqslant T.$$

*Step 8: Mixed spatial derivatives.* To obtain the mixed spatial derivative, introduce the mass-zero distribution

$$\eta_{y,i}(\varphi) = \partial_i\varphi(y), \qquad i \in [d].$$

The gradient estimate of Step 4 gives

$$\|P_{0,t}^*\eta_{y,i}\|_{-s} \leqslant g(t), \qquad \sup_{y,i} \int_0^\infty \|P_{0,t}^*\eta_{y,i}\|_{-s}^2 \,\mathrm{d}t < \infty.$$

The distribution $\eta_{y,i}$ need not belong to $\mathcal{H}^{-s}$, but its free evolution has the required time-integrated bound. Regularizing in space and using the energy estimate of Step 5 therefore gives

$$\int_0^\infty \|n_t^{y,i}\|_{-s}^2 \,\mathrm{d}t \leqslant C.$$

The backward representation and Cauchy–Schwarz then give

$$\|\nabla z_0^{y,i}\|_\infty \leqslant B_F \int_0^T g(t)\|n_t^{y,i}\|_{-s}\,\mathrm{d}t \leqslant B_F\|g\|_{L^2}\|n^{y,i}\|_{L_t^2 H^{-s}} \leqslant C.$$

Dominated convergence gives convergence of the regularized free evolutions in $L_t^2 H^{-s}$. Coercivity transfers this convergence, and continuity in $y$ and $\nu$, to the quadratic minimizers. Symmetry and the envelope formula identify

$$\partial_{y_i}\delta^2\chi_{*,T}(\nu,x,y) = z_0^{y,i}(x) \quad (\text{mod a function of } y).$$

This proves (52) for $\chi_{*,T}$. Duhamel's formula bounds $\|n_1^{y,i}\|_{-s}$ uniformly, so Step 7 also gives uniform exponential tails from time one.

*Step 9: The infinite-horizon value.* The equilibrium pair $(m_*, 0)$ solves the finite-horizon system. By Step 6 it is its unique solution for initial datum $m_*$, so $\chi_{*,T}(m_*) = 0$ and $\delta\chi_{*,T}(m_*) = 0$ modulo constants. The second-variation bound of Step 5 and Taylor's formula therefore give, uniformly in $T$,

$$-\frac{\kappa C_0}{2(1-\kappa B^2)}\|\nu - m_*\|_{-s}^2 \leqslant \chi_{*,T}(\nu) \leqslant \frac{B_F C_0}{2}\|\nu - m_*\|_{-s}^2.$$

To compare different horizons, use dynamic programming. For $S > 0$, append an $S$-horizon optimizer to the $T$-horizon optimizer; conversely, restrict a $(T+S)$-horizon optimizer to $[0,T]$. With $m^{\nu,H}$ denoting the $H$-horizon optimal flow, this gives

$$\chi_{*,S}(m_T^{\nu,T+S}) \leqslant \chi_{*,T+S}(\nu) - \chi_{*,T}(\nu) \leqslant \chi_{*,S}(m_T^{\nu,T}).$$

The quadratic value bounds and the decay estimate of Step 7 at time $T$, applied to both optimal flows, now give

$$|\chi_{*,T+S}(\nu) - \chi_{*,T}(\nu)| \leqslant Ce^{-2\gamma T}\|\nu - m_*\|_{-s}^2.$$

Thus $\chi_{*,T}$ converges uniformly on $\mathcal{P}(\mathbb{T}^d)$ to a finite continuous functional $\chi_*$. We identify the modulated limit with

$$\chi_*(\nu) = \inf_{\tilde\nu_0=\nu}\int_0^\infty \left(\frac{1}{2}\int |b_t|^2\,\mathrm{d}\tilde\nu_t + F_*(\tilde\nu_t)\right)\mathrm{d}t, \qquad \partial_t\tilde\nu_t = L_*^*\tilde\nu_t - \nabla\cdot(\tilde\nu_t b_t).$$

The action means the limit of the finite-interval actions when it exists. Since $|F_*(\nu)| \leqslant B_F\|\nu - m_*\|_{-s}^2/2$, the decay estimate of Step 7 implies, for $0 \leqslant H \leqslant T$,

$$\int_H^T \left(\frac{1}{2}\|\nabla u_t^{\nu,T}\|_{L^2(m_t^{\nu,T})}^2 + |F_*(m_t^{\nu,T})|\right)\mathrm{d}t \leqslant Ce^{-2\gamma H}\|\nu - m_*\|_{-s}^2.$$

Standard parabolic compactness on bounded intervals and this tail bound produce an infinite-horizon optimizer of cost $\chi_*(\nu)$. Every competitor costs at least $\chi_{*,T}(\nu)$ on $[0,T]$, proving the infinite-horizon control representation by passage to the limit. The difference identity of Step 6, now with a terminal pairing at $H$, proves uniqueness as $H \to \infty$: the pairing vanishes by exponential decay, and the coercivity bound of Step 5 forces the difference to vanish. The same argument applies to the linearized systems with $z_t \to 0$ modulo constants. Thus both characteristics and derivative systems converge on compact time intervals.

*Step 10: Convergence of the derivatives.* We now pass the derivative estimates to the infinite horizon. Fix $\nu$ and use horizon superscripts for the characteristics and generators. Write $A_t^T = A_{m_t^{\nu,T}}$ for the operator representing the second flat derivative of $F$ along the corresponding characteristic. Let $(n^T, z^T)$ and $(n^S, z^S)$ be the linearized solutions with initial directions $\sigma$, $\tau \in \mathcal{H}_0^{-s}$, respectively. Cross-testing their equations and using symmetry gives, for $H < \min(T, S)$,

$$\begin{aligned}\langle(\delta^2\chi_{*,T}(\nu) - \delta^2\chi_{*,S}(\nu))\sigma, \tau\rangle = \langle z_H^T, n_H^S\rangle - \langle z_H^S, n_H^T\rangle \\ + \int_0^H \Big(\langle (L^T - L^S)z^T, n^S\rangle \\ + \langle (L^T - L^S)z^S, n^T\rangle + \langle (A^T - A^S)n^T, n^S\rangle \\ + \langle m^S - m^T, \nabla z^T \cdot \nabla z^S\rangle\Big)\,\mathrm{d}t,\end{aligned}$$

where time subscripts in the integrand are suppressed. Step 7 bounds the terminal pairings by $Ce^{-2\gamma H}\|\sigma\|_{-s}\|\tau\|_{-s}$. Since $\mathcal{H}^s$ is an algebra and $s + 2 - \varepsilon > s + 1$, the same estimates bound the integrand by $C\|\sigma\|_{-s}\|\tau\|_{-s}$. On every fixed $[h, H] \subset (0, \infty)$, compactness and Sobolev interpolation give convergence of $m^T$ in $\mathcal{H}^{-s}$ and $\nabla u^T$ in $\mathcal{H}^s$, uniformly in time; operator-norm continuity of $A_\nu$ then makes the integrand $o_{T,S}(1)\|\sigma\|_{-s}\|\tau\|_{-s}$. Consequently,

$$\|\delta^2\chi_{*,T}(\nu) - \delta^2\chi_{*,S}(\nu)\|_{\mathcal{L}(\mathcal{H}_0^{-s}, \mathcal{H}^s/\mathbb{R})} \leqslant C(h + e^{-2\gamma H}) + o_{T,S;h,H}(1).$$

Choosing $h$ small, $H$ large and then $T$, $S$ large proves operator-norm convergence. The same comparison for different initial measures, using Step 6, gives continuity in $\nu$. For the mixed derivatives use Step 8: the short-time piece is bounded by $C\|g\|_{L^2(0,h)}$, the tail decays exponentially, and on $[h, H]$ smoothing and coercivity give convergence. Passing to the limit in the segment formulas now proves twice continuous flat differentiability, including at empirical measures, and the infinite-horizon envelope identities. The estimates (49) and (52) follow from Steps 5 and 8.

*Step 11: The corrector equation.* Define the original stationary corrector by

$$\chi(\nu) = \chi_*(\nu) - \int \log m_*(x)(\nu - m_*)(\mathrm{d}x).$$

The finite-horizon normalization passes to the limit: $\chi_*(m_*) = 0$ and $\delta\chi_*(m_*) = 0$ modulo constants. Consequently $\chi(m_*) = 0$, $\delta\chi(m_*) = -\log m_*$ modulo constants, and $\chi(\nu|m_*) = \chi_*(\nu)$. Passing the lower bound of Step 9 to the limit and using (53) proves (50), while testing $\chi_{*,T}(\nu)$ in Step 2 with $b = 0$, using $F_*(\tilde{\nu}_t) \leqslant B_F\kappa_s e^{-2\rho t}\mathcal{H}(\nu|m_*)$ when $\mathcal{H}(\nu|m_*) < \infty$, and letting $T \to \infty$ gives (51). If $F$ is flat convex, $Q(q) \geqslant 0$, so the envelope identity shows that each $\chi_{*,T}$ is flat convex. Its normalization makes it nonnegative, and both properties pass to $\chi_*$; the affine correction preserves flat convexity of $\chi$. Dynamic programming in the infinite-horizon control representation, followed by minimization over the initial drift, gives

$$0 = F_*(\nu) + \int L_*\delta\chi_*(\nu, x)\nu(\mathrm{d}x) - \frac{1}{2}\int |\nabla\delta\chi_*(\nu, x)|^2\nu(\mathrm{d}x).$$

The spatial estimates justify this identity also for singular $\nu$. It is exactly (29). Undoing the modulation gives (12) in Section 2.3.2, and integration by parts for

smooth densities gives (30). Finally, $\operatorname{Osc}_2 k_s \leqslant 4\kappa_s$, so both kernels satisfy the bounds for $(M,0)$-regularity on the torus. The mixed derivative estimate gives $|\operatorname{tr} D^2\chi(\nu,x,x)| \leqslant dC$. The empirical potential $N\chi_*(\mu_{\boldsymbol{x}})$ is $\mathcal{C}^2$ by these same derivative formulas. For each fixed $N$, the gradient flow of $\mathcal{F}_{*,N}$ satisfies the Fokker–Planck equation on $(\mathbb{T}^d)^N$ with drift

$$\big(-D\chi(\mu_{\boldsymbol{x}}, x^i)\big)_{i\in[N]}.$$

This drift is bounded and $\mathcal{C}^1$, so it defines a nonexplosive diffusion. Lemma 14 in Section 6.1 provides its dissipation identity and admissibility for every initial law with finite relative entropy. □

**Lemma 21** (Mean-field Schrödinger bridges)**.** *(a) Let $\mathcal{X} = \mathbb{T}^d$ and $s > d/2$. Suppose $(F_t)_{t\in[0,T]}$ satisfies Assumption Cost and, for all $t \in [0,T]$, $\nu \in \mathcal{P}(\mathbb{T}^d)$, and $\sigma \in \mathcal{H}_0^{-s}$,*

$$\operatorname{Osc} \delta F_t(\nu,\cdot) \leqslant 2C_1, \qquad \big|\langle \delta^2 F_t(\nu), \sigma^{\otimes 2}\rangle\big| \leqslant B_F \|\sigma\|_{-s}^2.$$

*If $m_0$, $m_T \in \mathcal{P}(\mathbb{T}^d)$ satisfy $\log m_0$, $\log m_T \in L^\infty(\mathbb{T}^d)$, then (1) admits at least one finite-cost minimizer $P_{\mathrm{ref}}$, and every such minimizer satisfies (5) with $\varphi_0$, $\varphi_T \in L^\infty(\mathbb{T}^d)$.*

*(b) Under the hypotheses of* (a)*, let $p_\tau$ denote the heat kernel of $\Delta$ on $\mathbb{T}^d$ at time $\tau > 0$, and set*

$$\Omega_0 = \max\{\operatorname{Osc}\log m_0, \operatorname{Osc}\log m_T\} + \frac{3C_1}{\pi^2} + 2\log\frac{\max p_{1/\pi^2}}{\min p_{1/\pi^2}}.$$

*If $T \geqslant 2/\pi^2$, the marginals $m_t = P_{\mathrm{ref},t}$ of any minimizer satisfy $\operatorname{Osc}\log m_t \leqslant \Omega_0$ and (LS) for all $t \in [0,T]$, with*

$$\rho = 4\pi^2 e^{-\Omega_0}. \tag{54}$$

*(c) On either $\mathbb{T}^d$ or $\mathbb{R}^d$, suppose $(F_t)_{t\in[0,T]}$ satisfies Assumption Cost, and let $P_{\mathrm{ref}} \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}))$ have finite cost in (1) and satisfy (5) with $\varphi_e \in L^1(m_e)$ $(e = 0,\ T)$, where $m_t = P_{\mathrm{ref},t}$. If $m_t$ satisfies (LS) and the lower bound in (MF-EB) for all $t \in [0,T]$ with $\rho$, $\eta > 0$, then $P_{\mathrm{ref}}$ is the unique finite-cost minimizer of (1).*

*Proof of Lemma 21. (a) Existence and tilt representation.* Integrating the Brownian bridge between independent endpoints distributed according to $m_0 \otimes m_T$ gives a competitor $P^0$ with $\mathcal{H}(P^0|R) = \mathcal{H}(m_0 \otimes m_T|R_{0,T}) < \infty$, because $p_T > 0$ is continuous on $\mathbb{T}^d \times \mathbb{T}^d$ and $\log m_T \in L^\infty(\mathbb{T}^d)$. Let $\bar\nu$ be the fixed measure in Assumption Cost. The flat derivative bound and integration along the segment from $\bar\nu$ to $\nu$ give

$$|F_t(\nu)| \leqslant |F_t(\bar\nu)| + C_T, \qquad \nu \in \mathcal{P}(\mathbb{T}^d).$$

The right-hand side is integrable in time. Consequently $P^0$ has finite running cost, and dominated convergence, using continuity of $F_t$ and of the evaluation maps $P \mapsto P_t$, shows that $P \mapsto \int_0^T F_t(P_t)\,\mathrm{d}t$ is bounded and weakly continuous on $\mathcal{P}(\mathcal{C}([0,T];\mathbb{T}^d))$. Since $\mathcal{H}(\cdot|R)$ is lower semicontinuous with compact sublevels, the direct method gives a finite-cost minimizer $P_{\mathrm{ref}}$ of (1) with marginals $m_t = P_{\mathrm{ref},t}$.

For any such minimizer $P_{\mathrm{ref}}$, set $\bar{V}_t(x) = \frac{1}{2}\big(\delta F_t(m_t, x) - \int_{\mathbb{T}^d} \delta F_t(m_t, y)\,\mathrm{d}y\big)$, so $|\bar{V}_t| \leqslant C_1$, and define the tilted reference path law $Q$ by

$$\frac{\mathrm{d}Q}{\mathrm{d}R}(x_{[0,T]}) = \frac{1}{Z}\exp\Big(-\int_0^T \bar{V}_t(x_t)\,\mathrm{d}t\Big).$$

For any competitor $P$ with $P_0 = m_0$, $P_T = m_T$, and $\mathcal{H}(P|R) < \infty$, put $P^\varepsilon = (1-\varepsilon)P_{\mathrm{ref}} + \varepsilon P$ for $\varepsilon \in (0,1)$. Changing reference from $R$ to $Q$ rewrites the mean-field objective at $P^\varepsilon$, up to an $\varepsilon$-independent constant, as $2\mathcal{H}(P^\varepsilon|Q) + \int_0^T F_t(P_t^\varepsilon|m_t)\,\mathrm{d}t$. Since $|\langle \delta^2 F_t(\nu), \sigma^{\otimes 2}\rangle| \leqslant B_F\|\sigma\|_{-s}^2$, Lemma 17 gives $|F_t(P_t^\varepsilon|m_t)| = O(\varepsilon^2)$ uniformly in $t$. Convexity of $\mathcal{H}(\cdot|Q)$ and optimality of $P_{\mathrm{ref}}$ as $\varepsilon \downarrow 0$ show that $P_{\mathrm{ref}}$ minimizes $\mathcal{H}(\cdot|Q)$ subject to $P_0 = m_0$ and $P_T = m_T$. Because $p_T$ and $\mathrm{d}Q/\,\mathrm{d}R$ are bounded above and away from zero and $\log m_0$, $\log m_T \in L^\infty(\mathbb{T}^d)$, we have $\log(\mathrm{d}Q_{0,T}/\,\mathrm{d}(m_0 \otimes m_T)) \in L^\infty$. By [36, Theorem 2.8], $\mathrm{d}P_{\mathrm{ref}}/\,\mathrm{d}Q = \tilde{a}_0(x_0)\tilde{a}_T(x_T)$ with positive factors satisfying $\log \tilde{a}_0$, $\log \tilde{a}_T \in L^\infty(\mathbb{T}^d)$. Absorbing the spatial constants from $\bar{V}_t$ and $Z$ yields (5) with $\varphi_0$, $\varphi_T \in L^\infty(\mathbb{T}^d)$.

*(b) Uniform log-Sobolev bound.* Set $\tau_0 = 1/\pi^2$ and $V_r = \frac{1}{2}\delta F_r(m_r, \cdot)$, so $\operatorname{Osc} V_r \leqslant C_1$ and $T \geqslant 2\tau_0$. By Lemma 24(i), $m_t = h_t^- h_t^+$, where under Brownian motion $B$ on $\mathbb{T}^d$ with generator $\Delta$,

$$h_t^+(x) = \mathbb{E}_x\Big[a_T(B_{T-t})\exp\Big(-\int_0^{T-t} V_{t+r}(B_r)\,\mathrm{d}r\Big)\Big],$$
$$h_t^-(x) = \mathbb{E}_x\Big[(a_0 m_0)(B_t)\exp\Big(-\int_0^{t} V_{t-r}(B_r)\,\mathrm{d}r\Big)\Big].$$

Put $L_0 = C_1\tau_0 + \log(\max p_{\tau_0}/\min p_{\tau_0})$. For $t \in [\tau_0, T]$, applying the Markov property at time $\tau_0$ in $h_t^-$ and conditioning on $B_{\tau_0} = y$ bounds $h_t^-(x)/h_t^-(x')$ by $e^{L_0}$, so $\operatorname{Osc}\log h_t^- \leqslant L_0$; symmetrically, $\operatorname{Osc}\log h_t^+ \leqslant L_0$ for $t \in [0, T-\tau_0]$. Since $T \geqslant 2\tau_0$, these intervals overlap on $[\tau_0, T-\tau_0]$, where $\operatorname{Osc}\log m_t \leqslant 2L_0$. On $[0,\tau_0]$, the identity $h_0^- = m_0/h_0^+$ gives $\operatorname{Osc}\log h_0^- \leqslant \operatorname{Osc}\log m_0 + L_0$, hence $\operatorname{Osc}\log h_t^- \leqslant \operatorname{Osc}\log m_0 + L_0 + C_1\tau_0$ and $\operatorname{Osc}\log m_t \leqslant \operatorname{Osc}\log m_0 + 2L_0 + C_1\tau_0$. Time reversal gives the same bound on $[T-\tau_0, T]$ with $m_T$ in place of $m_0$. Since $(\mathbb{T}^d, \mathrm{d}x)$ satisfies the log-Sobolev inequality with constant $4\pi^2$ [4, Propositions 5.2.7 and 5.7.5], the Holley–Stroock perturbation principle [32] yields (LS) with (54).

*(c) Optimality and uniqueness.* Applying the $N = 1$ change-of-reference and path-to-flow estimates in the proof of Theorem 1(a)–(b) to the finite-cost reference $P_{\mathrm{ref}}$ shows that every finite-cost competitor $P \in \mathcal{P}_2(\mathcal{C}([0,T];\mathcal{X}))$ with $P_0 = m_0$ and $P_T = m_T$ satisfies

$$\begin{aligned}
&2\mathcal{H}(P|R) + \int_0^T F_t(P_t)\,\mathrm{d}t - \Big(2\mathcal{H}(P_{\mathrm{ref}}|R) + \int_0^T F_t(m_t)\,\mathrm{d}t\Big)\\
&\quad = 2\mathcal{H}(P|P_{\mathrm{ref}}) + \int_0^T F_t(P_t|m_t)\,\mathrm{d}t\\
&\quad \geqslant \eta\rho\int_0^T \mathcal{H}(P_t|m_t)\,\mathrm{d}t \geqslant 0.
\end{aligned}$$

Hence $P_{\mathrm{ref}}$ minimizes (1). If equality holds, then $P_t = m_t$ and $F_t(P_t|m_t) = 0$ for a.e. $t \in [0,T]$, which forces $\mathcal{H}(P|P_{\mathrm{ref}}) = 0$ and thus $P = P_{\mathrm{ref}}$. □

# B Estimates for displacement convex costs

We establish the equilibrium, coercivity, and particle estimates used in Section 3.2. Throughout this appendix, $F$ satisfies all the assumptions of Section 3.2.

**Lemma 22** (Curvature of the ergodic state)**.** *The functional $\frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x)+F$ admits a minimizer $m_* \in \mathcal{P}_2(\mathbb{R}^d)$. Every minimizer has a positive $\mathcal{C}^3$ density, and $w_* = \sqrt{m_*}$ is the ground state of $\mathsf{H}_* = -2\Delta + V_*$, where $V_* = \delta F(m_*, \cdot)$. In particular, for some $c_* \in \mathbb{R}$,*

$$-2\Delta w_* + V_* w_* = c_* w_*, \qquad \sqrt{\lambda}\,\mathrm{Id} \leqslant -\nabla^2 \log m_* \leqslant \sqrt{K}\,\mathrm{Id}. \tag{55}$$

*Proof of Lemma 22.* Set $\mathcal{E} = \frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x) + F$. Applying (20) to the coupling between $\delta_0$ and $\nu$ gives $F(\nu) \geqslant \lambda M_2(\nu)/4 - C$. Hence minimizing sequences for $\mathcal{E}$ have bounded second moments and Fisher information. By (20), the lift $X \mapsto F(\mathcal{L}(X))$ on an atomless $L^2$ space is convex and continuous, hence weakly lower semicontinuous. If $\nu_n$ converges weakly with bounded second moments, a Skorohod representation gives random variables $X_n \to X$ almost surely with bounded $L^2$ norms. They therefore converge weakly in $L^2$, so $F(\mathcal{L}(X)) \leqslant \liminf_n F(\mathcal{L}(X_n))$. Together with lower semicontinuity of Fisher information, this proves existence of $m_*$.

Fix $x$, $y \in \mathbb{R}^d$ and $r \in [0,1]$, and put $z_r = (1-r)x + ry$. Apply (20) to

$$\pi_\varepsilon = (1-\varepsilon)(\mathrm{Id}, \mathrm{Id})_{\#} m_* + \varepsilon \delta_{(x,y)}.$$

Its interpolated measures are $\nu_{\varepsilon,r} = (1-\varepsilon)m_* + \varepsilon\delta_{z_r}$. Flat differentiability gives $F(\nu_{\varepsilon,r}) = F(m_*) + \varepsilon\langle V_*, \delta_{z_r} - m_*\rangle + o(\varepsilon)$. Dividing (20) by $\varepsilon$ and letting $\varepsilon \downarrow 0$ therefore yields

$$\frac{\lambda}{2} r(1-r)|y-x|^2 \leqslant (1-r)V_*(x) + rV_*(y) - V_*(z_r) \leqslant \frac{K}{2} r(1-r)|y-x|^2.$$

Thus $V_*$ is $\lambda$-convex and $K$-semiconcave, or equivalently, almost everywhere,

$$\lambda \mathrm{Id} \leqslant \nabla^2 V_* \leqslant K \mathrm{Id}. \tag{56}$$

In particular, $V_* \in \mathcal{C}^{1,1}(\mathbb{R}^d)$ and grows quadratically. For every finite-energy $\nu$, testing $\mathcal{E}$ along $(1-\varepsilon)m_* + \varepsilon\nu$ and using the flat convexity of Fisher information gives

$$0 \leqslant \frac{1}{2}\mathcal{I}(\nu|\,\mathrm{d}x) - \frac{1}{2}\mathcal{I}(m_*|\,\mathrm{d}x) + \int V_*\,\mathrm{d}(\nu - m_*).$$

For $w \in \mathcal{H}^1(\mathbb{R}^d)$ with $\|w\|_{L^2} = 1$, the identity $\mathcal{I}(w^2|\,\mathrm{d}x) = 4\int |\nabla w|^2\,\mathrm{d}x$ identifies $\frac{1}{2}\mathcal{I}(w^2|\,\mathrm{d}x) + \int V_* w^2\,\mathrm{d}x$ with the quadratic form

$$2\int |\nabla w|^2\,\mathrm{d}x + \int V_* w^2\,\mathrm{d}x$$

of $\mathsf{H}_*$. The preceding inequality therefore shows that $w_*$ minimizes this form under the constraint $\|w\|_{L^2} = 1$. After adding a constant to $V_*$, assume $V_* \geqslant 1$. Its quadratic confinement gives compact resolvent; elliptic regularity and the positivity of the ground state yield $w_* > 0$, $w_* \in \mathcal{C}^3$, and the first identity in (55); see [41, Theorem X.28] and [29, Chapters 8–9].

For $a \in \{\sqrt{\lambda}, \sqrt{K}\}$, put $h_a(x) = e^{-a|x|^2/4}$ and $G_a(x) = V_*(x) - a^2|x|^2/2$. The Gaussian ground-state transform and Feynman–Kac formula give

$$e^{dat}\frac{(e^{-t\mathsf{H}_*}h_a)(x)}{h_a(x)} = \mathbb{E}\exp\Big(-\int_0^t G_a(e^{-2as}x + Z_s^a)\,\mathrm{d}s\Big),$$

where $Z_s^a = 2\int_0^s e^{-2a(s-r)}\,\mathrm{d}B_r$. For $a = \sqrt{\lambda}$, $G_a$ is convex, so the exponential integrand is log-concave jointly in $x$ and the Gaussian noise path. The Gaussian Prékopa theorem [5] therefore makes its expectation log-concave in $x$. For $a = \sqrt{K}$, $G_a$ is concave, so the integrand is log-convex in $x$ for each noise path. Hölder's inequality then makes its expectation log-convex in $x$. The expectation is finite by the original Feynman–Kac formula with $V_* \geqslant 1$ and bounded $h_a$. Spectral theory gives

$$e^{c_* t}e^{-t\mathsf{H}_*}h_a \longrightarrow \langle h_a, w_*\rangle w_*$$

in $L^2$. Choose a subsequence converging almost everywhere. The preceding log-concavity and log-convexity pass to the limit: $w_*/h_{\sqrt{\lambda}}$ is log-concave, and $w_*/h_{\sqrt{K}}$ is log-convex. These are precisely the curvature bounds in (55). □

**Lemma 23** (Coercivity and particle bounds)**.** *Let $m_*$ be as in Lemma 22.*

*(a) For every $\nu \in \mathcal{P}_2(\mathbb{R}^d)$,*

$$\frac{1}{2}\mathcal{I}(\nu|m_*) + F(\nu|m_*) \geqslant \frac{\lambda}{2(\lambda+K)}\mathcal{I}(\nu|m_*). \tag{57}$$

*In particular, $m_*$ is the unique minimizer of $\frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x) + F$.*

*(b) For every $N \geqslant 1$, (EB) holds for $F_t = F$ and $m_t = m_*$ with*

$$\eta_N = \frac{\lambda}{\lambda+2K}, \quad M_N = \frac{K}{\sqrt{\lambda}}, \quad \Delta_F^- = \Delta_F^+ = \frac{34dK^2}{\lambda^{3/2}}.$$

*Proof of Lemma 23. (a) Mean-field coercivity.* For brevity, write $F_* = F(\cdot|m_*)$ and $\mathcal{E}_*(\nu) = \frac{1}{2}\mathcal{I}(\nu|m_*) + F_*(\nu)$. Set $\mathcal{E} = \frac{1}{2}\mathcal{I}(\cdot|\,\mathrm{d}x) + F$ and $b_* = \nabla \log m_*$. First let $\nu$ be a smooth positive density on $\mathbb{R}^d$ with globally Lipschitz score $b_\nu = \nabla \log \nu$. Couple the stationary Langevin diffusions with laws $\nu$ and $m_*$ using the same Brownian motion, and choose a joint invariant law $\mathcal{L}(X, Y)$. Such a law exists by averaging the joint transition laws in time: their marginals remain $\nu$ and $m_*$, so the averages are tight, and the joint diffusion is Feller. Write

$$\Delta = X - Y, \qquad U = b_\nu(X) - b_*(Y), \qquad Z_s = (1-s)Y + sX, \quad \mu_s = \mathcal{L}(Z_s).$$

The process $Z_s$ is stationary with diffusion coefficient $\sqrt{2}$ and conditional drift $\beta_s(Z_s) = \mathbb{E}[(1-s)b_*(Y) + sb_\nu(X) \mid Z_s]$. Testing its stationary Fokker–Planck equation against the log density, by convolution and cutoff if necessary, gives

$$\mathcal{I}(\mu_s|\,\mathrm{d}x) \leqslant \int |\beta_s|^2\,\mathrm{d}\mu_s \leqslant \mathbb{E}\,|(1-s)b_*(Y) + sb_\nu(X)|^2. \tag{58}$$

Expanding the square on the right gives

$$\mathcal{I}(\mu_s|\,\mathrm{d}x) \leqslant (1-s)\mathcal{I}(m_*|\,\mathrm{d}x) + s\mathcal{I}(\nu|\,\mathrm{d}x) - s(1-s)\,\mathbb{E}\,|U|^2.$$

Adding the displacement convexity bound (20), applied to $\mathcal{L}(Y, X)$, yields

$$\mathcal{E}(\mu_s) \leqslant (1-s)\mathcal{E}(m_*) + s\mathcal{E}(\nu) - \frac{s(1-s)}{2}\big(\mathbb{E}\,|U|^2 + \lambda\,\mathbb{E}\,|\Delta|^2\big).$$

Since $\mathcal{E}(\mu_s) \geqslant \mathcal{E}(m_*)$, rearranging and letting $s \downarrow 0$ gives

$$\mathcal{E}(\nu) - \mathcal{E}(m_*) \geqslant \frac{1}{2}\big(\mathbb{E}\,|U|^2 + \lambda\,\mathbb{E}\,|\Delta|^2\big).$$

The reference score $b_*$ is $\sqrt{K}$-Lipschitz by (55). Thus weighted Cauchy–Schwarz gives

$$\begin{aligned}\mathcal{I}(\nu|m_*) &= \mathbb{E}\,|U + b_*(Y) - b_*(X)|^2 \\ &\leqslant \mathbb{E}\big(|U| + \sqrt{K}|\Delta|\big)^2 \\ &\leqslant \left(1 + \frac{K}{\lambda}\right)\big(\mathbb{E}\,|U|^2 + \lambda\,\mathbb{E}\,|\Delta|^2\big).\end{aligned}$$

Finally, the frozen ground-state equation and integration by parts against radial cutoffs yield

$$\mathcal{E}(\nu) - \mathcal{E}(m_*) = \mathcal{E}_*(\nu) = \frac{1}{2}\mathcal{I}(\nu|m_*) + F_*(\nu), \tag{59}$$

giving (57) for smooth positive $\nu$ with globally Lipschitz score.

For a general $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ with finite relative Fisher information, approximate $\sqrt{\nu}$ in $\mathcal{H}^1$ and in $L^2(|x|^2\,\mathrm{d}x)$ by compactly supported smooth functions, then add a vanishing Gaussian density and normalize. The resulting smooth positive laws $\nu_n$ have globally Lipschitz scores, converge to $\nu$ in $W_2$, and satisfy $\mathcal{I}(\nu_n|\,\mathrm{d}x) \to \mathcal{I}(\nu|\,\mathrm{d}x)$. The ground-state equation and the quadratic growth of $V_*$ give $\mathcal{I}(\nu_n|m_*) \to \mathcal{I}(\nu|m_*)$; the remaining terms converge by their quadratic growth and continuity. Passing to the limit proves (57). Infinite Fisher information is handled in the extended-value sense. Together with (59), this shows that any minimizer of $\mathcal{E}$ has zero relative Fisher information with respect to $m_*$, hence equals $m_*$. This proves (a).

*(b) Particle bounds.* Set

$$g_N(\boldsymbol{x}) = NF_*(\mu_{\boldsymbol{x}}), \qquad A_N = \big|\mathbb{E}_{m_*^{\otimes N}}\, g_N\big|, \qquad B_N = \mathbb{E}_{m_*^{\otimes N}}\,|\nabla g_N|^2.$$

Write $q = m_*^{\otimes N}$ and $\mathcal{V}_N(\boldsymbol{x}) = NF(\mu_{\boldsymbol{x}})$. The assumptions imply that $\mathcal{V}_N$ is $\lambda$-convex and $K$-semiconcave, whereas $g_N$ has Hessian between $(\lambda - K)\mathrm{Id}$ and $(K - \lambda)\mathrm{Id}$. For a smooth positive $\nu^N$ with globally Lipschitz score, use the stationary synchronous coupling $(X, Y)$ of (a), now with laws $\nu^N$ and $q$. Put $\Delta = X - Y$, $U = \nabla \log \nu^N(X) - \nabla \log q(Y)$ and $\mu_s = \mathcal{L}(Y + s\Delta)$, and write $\mathcal{J}_N(\mu) = \frac{1}{2}\mathcal{I}(\mu|\,\mathrm{d}\boldsymbol{x}) + \int \mathcal{V}_N\,\mathrm{d}\mu$. Combining the stationary Fisher bound with the $\lambda$-convexity of $\mathcal{V}_N$ gives

$$\mathcal{J}_N(\mu_s) - \mathcal{J}_N(q) \leqslant s\big(\mathcal{J}_N(\nu^N) - \mathcal{J}_N(q)\big) - \frac{s(1-s)}{2}\,\mathbb{E}\big(|U|^2 + \lambda|\Delta|^2\big).$$

For the lower bound, the product ground-state identity and $(K - \lambda)$-semiconvexity of $g_N$ give

$$\begin{aligned}\mathcal{J}_N(\mu_s) - \mathcal{J}_N(q) &= \frac{1}{2}\mathcal{I}(\mu_s|q) + \mathbb{E}[g_N(Y + s\Delta) - g_N(Y)] \\ &\geqslant s\,\mathbb{E}[\nabla g_N(Y)\cdot\Delta] - \frac{K - \lambda}{2}s^2\,\mathbb{E}\,|\Delta|^2.\end{aligned}$$

Divide by $s$ and let $s \downarrow 0$. Since $\mathcal{J}_N(\nu^N) - \mathcal{J}_N(q) = \frac{1}{2}\mathcal{I}(\nu^N|q) + \mathbb{E}_{\nu^N}\, g_N - \mathbb{E}_q\, g_N$, Young's inequality yields

$$\frac{1}{2}\mathcal{I}(\nu^N|q) + \mathbb{E}_{\nu^N}\, g_N \geqslant \frac{1}{2}\,\mathbb{E}\,|U|^2 + \frac{\lambda}{4}\,\mathbb{E}\,|\Delta|^2 - A_N - \frac{B_N}{\lambda}.$$

The $\sqrt{K}$-Lipschitz score of $q$ gives $\mathcal{I}(\nu^N|q) \leqslant (1+2K/\lambda)\,\mathbb{E}(|U|^2+\lambda|\Delta|^2/2)$. The approximation used in (a) extends the resulting bound to every $\nu^N \in \mathcal{P}_2(\mathbb{R}^{dN})$:

$$\frac{1}{2}\mathcal{I}(\nu^N|m_*^{\otimes N}) + \mathbb{E}_{\nu^N}\, g_N \geqslant \frac{\lambda}{2(\lambda+2K)}\mathcal{I}(\nu^N|m_*^{\otimes N}) - A_N - \frac{B_N}{\lambda}. \tag{60}$$

For an optimal transport coupling $(X,Y)$ of $\nu^N$ and $q$, $(K-\lambda)$-semiconcavity and Young's inequality give $\mathbb{E}\, g_N(X) \leqslant A_N + B_N/(2\lambda) + K\,\mathbb{E}\,|X-Y|^2/2$. The tensorized Talagrand inequality therefore gives

$$\mathbb{E}_{\nu^N}\, g_N \leqslant \frac{K}{\sqrt{\lambda}}\mathcal{H}(\nu^N|m_*^{\otimes N}) + A_N + \frac{B_N}{2\lambda},$$

so it remains to bound $A_N$ and $B_N$ uniformly in $N$.

Let $\mathrm{Var}(\nu)$ denote the total variance of $\nu$, put $v_* = \mathrm{Var}(m_*)$, and define $J_N(\nu) = \int F(\mu_{\boldsymbol{y}})\,\nu^{\otimes N}(\mathrm{d}\boldsymbol{y})$. For independent $Y_1,\dots,Y_N \sim \nu$, introduce $\nu_j = N^{-1}\sum_{i\leqslant j}\delta_{Y_i} + (1-j/N)\nu$. Taylor-expand $F(\nu_j) - F(\nu_{j-1})$ conditionally on $Y_1,\dots,Y_{j-1}$. The linear term has zero expectation, while (21) bounds the absolute remainder by $\frac{K}{2N^2}\int |Y_j - z|^2\nu(\mathrm{d}z)$. Summing and taking expectations gives $|J_N(\nu) - F(\nu)| \leqslant K\,\mathrm{Var}(\nu)/N$, as in the exchangeability argument of [17]. In particular,

$$|J_N(m_*) - F(m_*)| \leqslant \frac{K}{N}v_*. \tag{61}$$

The same bound for arbitrary $\nu$ shows that the corrected sampling bias

$$\mathcal{B}_N(\nu) = J_N(\nu) - F(\nu) + \frac{K}{N}\,\mathrm{Var}(\nu)$$

is nonnegative. To estimate its derivative, note that the lift of $J_N$ is $K$-semiconcave by applying (20) to $N$ independent copies of any coupling. The lifts of $-F$ and Var are respectively $(-\lambda)$-semiconcave and 2-semiconcave. Thus the lift of $\mathcal{B}_N$ is $(K-\lambda+2K/N)$-semiconcave, hence $3K$-semiconcave. Nonnegativity and the descent inequality give

$$\|D\mathcal{B}_N(m_*)\|^2_{L^2(m_*)} \leqslant 6K\mathcal{B}_N(m_*) \leqslant \frac{12K^2}{N}v_*.$$

Using $D\,\mathrm{Var}(\nu,x) = 2(x - \int y\nu(\mathrm{d}y))$, we obtain

$$\|DJ_N(m_*) - DF(m_*)\|^2_{L^2(m_*)} \leqslant 2\|D\mathcal{B}_N(m_*)\|^2_{L^2(m_*)} + \frac{8K^2}{N^2}v_* \leqslant \frac{32K^2}{N}v_*.$$

For independent $Y_1,\dots,Y_N$ with law $m_*$, $\mathbb{E}[DF(\mu_{\boldsymbol{Y}},Y_i)\mid Y_i] = DJ_N(m_*,Y_i)$. Conditional Efron–Stein and (21) therefore give

$$A_N \leqslant Kv_*, \qquad B_N \leqslant K^2v_* + N\|DJ_N(m_*) - DF(m_*)\|^2_{L^2(m_*)} \leqslant 33K^2v_*.$$

Finally, Poincaré gives $v_* \leqslant d/\sqrt{\lambda}$; since $\lambda \leqslant K$, the stated constants follow. □

# C Finite-entropy diffusions

We identify the path law of least entropy relative to a reference law $S$ among laws with prescribed marginals. For references obtained by tilting Brownian motion,

a drift perturbation has a quadratic entropy cost. We first prove the converse realization result, then use time reversal to express this cost in terms of the relative velocity and Fisher information of the marginal flow. The resulting variational formula is given in Section C.3. We also describe the minimizing evolution when only one endpoint is prescribed.

Fix $s < t$ and let $\mathcal{X} = \mathbb{R}^n$ or $\mathbb{T}^n$. All path laws are defined on $\mathcal{C}([s,t];\mathcal{X})$, and the diffusion coefficient is $\sqrt{2}$. Put $\langle x\rangle_{\mathcal{X}}^2 = 1 + |x|^2$ on $\mathbb{R}^n$ and $\langle x\rangle_{\mathcal{X}} = 1$ on $\mathbb{T}^n$.

## C.1 Brownian tilts

**Assumption Diff** (Diffusion reference)**.** Let $W$ be Brownian motion with generator $\Delta$ and initial law $\nu^W \in \mathcal{P}_2(\mathcal{X})$. The reference law $S \in \mathcal{P}_2(\mathcal{C}([s,t];\mathcal{X}))$ has $\mathcal{H}(S|W) < \infty$ and

$$\frac{\mathrm{d}S}{\mathrm{d}W} = \alpha_s(X_s)\alpha_t(X_t)\exp\Big(-\int_s^t V_r(X_r)\,\mathrm{d}r\Big).$$

The potential $V$ is measurable in time, continuous in space, and satisfies $|V_r(x)| \leqslant C\langle x\rangle_{\mathcal{X}}^2$. The endpoint factors $\alpha_s$, $\alpha_t$ are nonnegative Borel functions, finite and positive $S_e$-almost everywhere, $e = s$, $t$.

Write $\sigma_r = S_r$. Solving the drift equation with drift $b$ means that, for every $\psi \in \mathcal{C}_c^\infty(\mathcal{X})$,

$$\psi(X_r) - \psi(X_s) - \int_s^r (\Delta\psi + b_q\cdot\nabla\psi)(X_q)\,\mathrm{d}q$$

is a local martingale. The reference class is closed under restriction and tensor products, as shown in the following proof.

**Lemma 24** (Reference drifts and path moments)**.** *Under Assumption Diff, the following hold.*

*(i) $S$ and its time reversal are nonexplosive Markov diffusions with noise $\sqrt{2}$ and locally bounded drifts on compact subsets of $(s,t)\times\mathcal{X}$. For $s < r < t$, $\sigma_r$ has a positive $\mathcal{C}^1$ density. If $h^+$ is the backward Feynman–Kac factor with terminal data $h_t^+ = \alpha_t$, the forward drift $b^S$ and current velocity $v^S$ satisfy*

$$b_r^S = 2\nabla\log h_r^+, \qquad v_r^S = b_r^S - \nabla\log\sigma_r,$$
$$\mathcal{H}(S|W) = \mathcal{H}(S_s|\nu^W) + \frac{1}{4}\int_s^t\int |b_r^S|^2\,\mathrm{d}\sigma_r\,\mathrm{d}r.$$

*On compact interior time intervals, the forward and reversed equations stopped on leaving a bounded spatial domain have unique weak laws for each starting point. On the full interval, uniqueness holds among absolutely continuous laws with the same initial law, in either time direction and for reference transition laws from almost every starting point. If $\sigma_r = m$ is stationary and $|\nabla\log m(x)| \leqslant C\langle x\rangle_{\mathcal{X}}$, then $v_r^S = 0$ and $b_r^S = \nabla\log m$ for almost every $r$.*

*(ii) Let $P$ have $\mathcal{H}(P|S) < \infty$, endpoints in $\mathcal{P}_2(\mathcal{X})$, and $\int_s^t\int\langle x\rangle_{\mathcal{X}}^2\,\mathrm{d}P_r\,\mathrm{d}r < \infty$. Then $P \in \mathcal{P}_2(\mathcal{C}([s,t];\mathcal{X}))$. If, in addition, $\log\alpha_e \in L^1(P_e)$ for $e = s$, $t$, then $\mathcal{H}(P|W) < \infty$ and*

$$\mathcal{H}(P|W) = \mathcal{H}(P|S) + \mathbb{E}_P\Big[\log\alpha_s(X_s) + \log\alpha_t(X_t) - \int_s^t V_r(X_r)\,\mathrm{d}r\Big].$$

*Proof of Lemma 24. (i) Reference laws.* At each interior time $r$, the displayed density is a product of a past factor and a future factor. Brownian past and future are independent conditionally on $X_r$; weighting them by separate factors preserves that independence. Thus $S$ is Markov. Its Feynman–Kac factors satisfy

$$\partial_r h_r^+ = -\Delta h_r^+ + V_r h_r^+, \qquad \partial_r h_r^- = \Delta h_r^- - V_r h_r^-, \qquad \sigma_r = h_r^- h_r^+,$$

with boundary data $h_t^+ = \alpha_t$ and $h_s^- = \alpha_s \nu^W$; the latter may be a measure. The forward and reversed drifts, indexed by original time, are $2\nabla \log h_r^+$ and $2\nabla \log h_r^-$, respectively [38]. Positivity of the heat kernel and interior parabolic regularity with locally bounded potential give positive $\mathcal{C}^1$ marginal densities and locally bounded drifts.

Stop the forward equation in bounded domains and where $h^+$ and its inverse are bounded. Undoing the Doob transform reduces stopped uniqueness to Brownian uniqueness. Use $h^-$ for the reversed equation. For almost every starting point, each strict-prefix conditional law in either time direction is equivalent to Brownian motion: its density is the positive ratio of the corresponding Feynman–Kac factors, times the running weight. Equivalent change of measure transfers Brownian predictable representation. The density martingale of an absolutely continuous competitor with the same drift and initial law consequently has zero stochastic differential and is identically one, conditionally on the starting point. This proves uniqueness among absolutely continuous laws. Increasing strict prefixes and using path continuity gives the terminal extension. Finite-entropy Girsanov gives the energy identity [37, Theorems 1 and 2].

For $s \leqslant a < b \leqslant t$, let $W^{\sigma_a}$ be Brownian motion on $[a, b]$ started from $\sigma_a$. The Markov property and the Feynman–Kac formula give

$$\frac{\mathrm{d}S_{[a,b]}}{\mathrm{d}W^{\sigma_a}} = \frac{h_b^+(X_b)}{h_a^+(X_a)} \exp\Big( - \int_a^b V_r(X_r)\,\mathrm{d}r \Big).$$

Here $h_a^+$ and $h_b^+$ are finite and positive at the corresponding reference endpoints. The restricted law has a finite path second moment, and its drift energy is bounded by that on $[s, t]$. Stopping at levels of this energy, applying Girsanov and using weak uniqueness of Brownian motion give

$$\mathcal{H}(S_{[a,b]} | W^{\sigma_a}) \leqslant \frac{1}{4} \int_a^b \int |b_r^S|^2 \,\mathrm{d}\sigma_r \,\mathrm{d}r < \infty$$

by lower semicontinuity as the stopping levels tend to infinity. Thus the restricted law satisfies Assumption Diff, with endpoint factors $1/h_a^+$, $h_b^+$. For tensor products, the factors multiply, the potentials and entropies add, and the path second moments remain finite. This proves both closure properties.

The reversed Brownian drift is $2\nabla \log W_r$ and $\int |\nabla \log W_r| \,\mathrm{d}W_r \leqslant C(r-s)^{-1/2}$. This integrable bound makes $W^{\leftarrow}$ a semimartingale on the closed interval, even for singular $\nu^W$. Since $\mathcal{H}(S^{\leftarrow} | W^{\leftarrow}) = \mathcal{H}(S|W) < \infty$, finite-entropy Girsanov gives a global semimartingale decomposition for $S^{\leftarrow}$, whose interior drift is $2\nabla \log h_r^-$.

For the stationary assertion, put $\psi_r = 2 \log h_r^+ - \log m$. Finite drift energy and the growth bound on $\nabla \log m$ give $\nabla \psi_r \in L^2(m)$ for almost every $r$, while stationarity gives $\nabla \cdot (m \nabla \psi_r) = 0$. On $\mathbb{R}^n$, test with $\zeta_R^2 T_K(\psi_r)$, where $T_K(z) = \max(-K, \min(K, z))$ and $\zeta_R$ is a smooth spatial cutoff with $|\nabla \zeta_R| \leqslant C/R$. Then

$$\int \zeta_R^2 \mathbf{1}_{\{|\psi_r| < K\}} |\nabla \psi_r|^2 \,\mathrm{d}m = -2 \int \zeta_R T_K(\psi_r) \nabla \psi_r \cdot \nabla \zeta_R \,\mathrm{d}m \longrightarrow 0$$

by Cauchy–Schwarz. Letting $R \to \infty$ and then $K \to \infty$ gives $\nabla\psi_r = 0$. On the torus use $T_K(\psi_r)$ directly. Thus $v_r^S = 0$ and $b_r^S = \nabla \log m$.

*(ii) Path moments and change of reference.* On $\mathbb{R}^n$, condition on the endpoints of a subinterval $[a, b]$. All endpoint factors cancel in the conditional density:

$$\frac{\mathrm{d}S^{x,y}_{[a,b]}}{\mathrm{d}W^{x,y}_{[a,b]}} = \frac{\exp(-\int_a^b V_r(X_r)\,\mathrm{d}r)}{\mathbb{E}_{W^{x,y}_{[a,b]}} \exp(-\int_a^b V_r(X_r)\,\mathrm{d}r)},$$

for $S_{a,b}$-almost every $(x, y)$, where $S_{a,b} = \mathcal{L}_S(X_a, X_b)$ and $W^{x,y}_{[a,b]}$ is the Brownian bridge. For some $\delta$, $\eta > 0$ and $A$, $B < \infty$, depending only on the potential bound and dimension, whenever $b - a \leqslant \delta$,

$$\log \mathbb{E}_{S^{x,y}_{[a,b]}} \exp\bigl(\eta \sup_{a\leqslant r\leqslant b} |X_r|^2\bigr) \leqslant A + B(|x|^2 + |y|^2).$$

Indeed, write the Brownian bridge as the linear interpolation between $x$ and $y$ plus a centered bridge. For sufficiently small $\delta$, $\eta$, the Gaussian exponential moment of the centered maximum controls the numerator; Jensen's inequality bounds the denominator below by $\exp[-C'\delta(1 + |x|^2 + |y|^2 + \delta)]$. This proves the asserted exponential moment bound.

Entropy disintegration over the two endpoints and the conditional entropy inequality now give

$$\eta\,\mathbb{E}_P \sup_{a\leqslant r\leqslant b} |X_r|^2 \leqslant \mathcal{H}(P_{[a,b]}|S_{[a,b]}) + A + B\biggl(\int |x|^2\,\mathrm{d}P_a + \int |y|^2\,\mathrm{d}P_b\biggr).$$

Choose a finite partition with mesh at most $\delta$ and partition times having finite marginal second moments. Such a partition exists by the integrated moment assumption and the endpoint moments. Data processing bounds each segment entropy by $\mathcal{H}(P|S)$; summing the estimates proves the path second moment. On the torus it is automatic.

It remains to change reference. The integrated second moment makes the running potential integrable under $P$. The assumptions $\log \alpha_e \in L^1(P_e)$ therefore make $\log(\mathrm{d}S/\,\mathrm{d}W)$ integrable under $P$. The entropy change-of-reference identity then gives the formula for $\mathcal{H}(P|W)$ and shows that it is finite. □

## C.2 Drift perturbations

The next lemma relates path entropy to forward drift energy and realizes a given Fokker–Planck flow by a Markov law. For a drift perturbation that depends on the whole past, conditioning on the present then lowers the entropy while preserving all marginals.

**Lemma 25** (Drift energy and Markov realization)**.** *Let Assumption Diff hold, with forward drift $b^S$ as in Lemma 24.*

*(i) Every law $P$ with $\mathcal{H}(P|S) < \infty$ has a canonical predictable drift perturbation $u$ relative to $S$ and satisfies*

$$\mathcal{H}(P|S) = \mathcal{H}(P_s|S_s) + \frac{1}{4}\,\mathbb{E}_P \int_s^t |u_r|^2\,\mathrm{d}r. \tag{62}$$

*The right-hand side is finite, and the same identity holds for the reversed laws.*

*(ii) Let $u$ be a Borel state field and $\nu \in \mathcal{C}([s,t];\mathcal{P}_2(\mathcal{X}))$ satisfy*

$$\partial_r \nu_r = \Delta \nu_r - \nabla \cdot (\nu_r (b_r^S + u_r)), \qquad \mathcal{H}(\nu_s | S_s) + \int_s^t \int |u_r|^2 \, \mathrm{d}\nu_r \, \mathrm{d}r < \infty.$$

*Then there is a Markov diffusion law $P \in \mathcal{P}_2(\mathcal{C}([s,t];\mathcal{X}))$ with these marginals, drift $b^S + u$, and entropy (62).*

*Proof of Lemma 25.* Part (i) follows from finite-entropy Girsanov [37, Theorems 1 and 2] and the uniqueness among absolutely continuous laws in Lemma 24(i). For (ii), we first control the full drift on each strict prefix $[s,v]$, $v < t$. We then realize the flow there, estimate its relative entropy and pass to the terminal time. The path moment bound and the Markov property are verified last.

*Step 1: Bounds on the Feynman–Kac factor.* To control $b^S$ under $\nu$, we will test the Fokker–Planck equation with $\psi_r = \log h_r^+$, where $h^+$ is the backward Feynman–Kac factor of the reference. We first establish the growth bounds needed for this test. In the interior, $h^+$ is positive and finite, $b^S = 2\nabla\psi$, and

$$(\partial_r + \Delta)\psi = V - |\nabla\psi|^2 = V - \frac{1}{4}|b^S|^2.$$

These identities can be understood through local parabolic regularization when $V$ is only measurable in time. For every $v < t$,

$$\psi_r(x) \geqslant -C_v \langle x \rangle_{\mathcal{X}}^2, \qquad s < r \leqslant v,$$

and for every $s < a < v < t$,

$$|\psi_r(x)| \leqslant C_{a,v} \langle x \rangle_{\mathcal{X}}^2, \qquad a \leqslant r \leqslant v.$$

We obtain these bounds directly from the Feynman–Kac kernel. There are a bounded set $B$ of positive volume and $c > 0$ such that $\alpha_t \geqslant c$ on $B$. If $K_V$ denotes the Feynman–Kac kernel and $p_\delta$ the heat kernel, Jensen's inequality under the Brownian bridge gives, on $\mathbb{R}^n$,

$$K_V(r,x;t,y) \geqslant p_{t-r}(x,y) \exp\bigl[-C'(t-r)(1+|x|^2+|y|^2+t-r)\bigr].$$

Integrating over $B$ proves the lower growth bound. For the upper bound, fix $x_* \in \mathcal{X}$ and a sufficiently small $\delta > 0$ with $r - \delta > s$ and $r + \delta < t$. The exponential second moment of a short Brownian bridge yields

$$K_V(r,x;r+\delta,y) \leqslant A_\delta p_\delta(x,y) \exp\bigl[C'\delta(1+|x|^2+|y|^2)\bigr].$$

Combine this with the preceding lower bound for $K_V(r-\delta,x_*;r+\delta,y)$. For $\delta$ small enough, comparison of the Gaussian exponents gives

$$K_V(r,x;r+\delta,y) \leqslant A \exp(B\langle x \rangle_{\mathcal{X}}^2) K_V(r-\delta,x_*;r+\delta,y).$$

Integration against $h_{r+\delta}^+(y)$ proves the upper bound uniformly on compact interior time intervals. On the torus, the same argument uses the positive heat kernel on the compact state space.

*Step 2: Reference drift energy under $\nu$.* The assumption gives finite energy for $u$ under $\nu$. We now use Step 1 to control $b^S$ on each strict prefix $[s,v]$, $v < t$. The

marginal second moments are uniformly bounded by $W_2$ continuity. We first justify testing with $\psi$ on $[a,v] \subset (s,t)$. On $\mathbb{R}^n$, choose cutoffs $0 \leqslant \theta_L \leqslant 1$, equal to one on $B_L$ and zero outside $B_{2L}$, with

$$|\Delta\theta_L| \leqslant C/L^2, \qquad \frac{|\nabla\theta_L|^2}{\theta_L} \leqslant C/L^2.$$

Testing first with local smooth approximations of $\theta_L\psi$ and passing to the limit is justified by local boundedness of $\nabla\psi$ and approximation of $(\partial_r + \Delta)\psi = V - |\nabla\psi|^2$ in $L^1_r C(K)$ on each spatial compact set $K$.

Young's inequality absorbs the terms containing $b^S \cdot u$ and $(1+\psi)b^S \cdot \nabla\theta_L$ into a fixed fraction of $\int_a^v \int \theta_L |b^S|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r$. By the interior growth bound on $\psi$, the remaining cutoff terms are bounded by

$$C_{a,v} \int_a^v \int_{\{L \leqslant |x| \leqslant 2L\}} \big(\langle x\rangle_{\mathcal{X}}^2 + |u_r|^2\big) \,\mathrm{d}\nu_r \,\mathrm{d}r.$$

The endpoint terms, the potential term and the perturbation energy are finite. We obtain an energy bound independent of $L$, and hence $b^S \in L^2(\nu_r \,\mathrm{d}r)$ on $[a,v]$. The cutoff errors now vanish by Cauchy–Schwarz and the second-moment tails. On the torus, test directly with $\psi$. Thus

$$\int \psi_v \,\mathrm{d}\nu_v - \int \psi_a \,\mathrm{d}\nu_a = \int_a^v \int \Big(V + \frac{1}{4}|b^S|^2 + \frac{1}{2} b^S \cdot u\Big) \,\mathrm{d}\nu_r \,\mathrm{d}r.$$

Since

$$\frac{1}{4}|b^S|^2 + \frac{1}{2} b^S \cdot u \geqslant \frac{1}{8}|b^S|^2 - \frac{1}{2}|u|^2,$$

the lower growth bound, the upper bound at the fixed time $v$, and the bound on $V$ give

$$\frac{1}{8}\int_a^v \int |b^S_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r \leqslant C_v \sup_{s \leqslant r \leqslant t} \int \langle x\rangle_{\mathcal{X}}^2 \,\mathrm{d}\nu_r + \frac{1}{2}\int_s^v \int |u_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r.$$

The constant is independent of $a \downarrow s$. Consequently,

$$\int_s^v \int |b^S_r + u_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r < \infty, \qquad v < t.$$

*Step 3: Realization on strict prefixes.* Fix $v < t$. The Brownian superposition principle [47, Theorem 2.5], together with the preceding finite-energy bound, gives a diffusion law $Q^v$ on $[s,v]$ with marginals $\nu$ and drift $b^S + u$. Let $W^{\nu_s}$ be Brownian motion with initial law $\nu_s$. To estimate $\mathcal{H}(Q^v|W^{\nu_s})$, stop when $\int_s^r |(b^S_q + u_q)(X_q)|^2 \,\mathrm{d}q$ reaches $k$, and continue with Brownian motion. The inverse Girsanov exponential has bounded energy, hence is a true martingale; full weak uniqueness of Brownian motion identifies the transformed law. The stopped laws have entropy at most one quarter of the original energy and converge to $Q^v$. Lower semicontinuity gives finite entropy, and finite-entropy Brownian Girsanov then yields

$$\mathcal{H}(Q^v|W^{\nu_s}) = \frac{1}{4}\int_s^v \int |b^S_r + u_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r < \infty.$$

Disintegrate $Q^v$ and $S_{[s,v]}$ with respect to $X_s = x$. For $\sigma_s$-almost every $x$, $0 < h^+_s(x) < \infty$, and the conditional reference law satisfies

$$\frac{\mathrm{d}S^x_{[s,v]}}{\mathrm{d}W^x} = \frac{h^+_v(X_v)}{h^+_s(x)} \exp\Big(-\int_s^v V_r(X_r) \,\mathrm{d}r\Big).$$

This density is strictly positive. Since $\nu_s \ll \sigma_s$, the formula holds for $\nu_s$-almost every $x$. Disintegrating $\mathcal{H}(Q^v|W^{\nu_s})$ over $X_s$ gives $\mathcal{H}(Q^{v,x}|W^x) < \infty$; the conditional second moments and the conditional energy are also finite almost everywhere. The displayed log density is integrable under $Q^{v,x}$, by the interior growth bound on $\psi$, the quadratic bound on $V$, and the finiteness of the number $\log h_s^+(x)$. Changing reference therefore proves $\mathcal{H}(Q^{v,x}|S^x_{[s,v]}) < \infty$.

We can now apply the finite-entropy Girsanov identity [37, Theorems 1 and 2] to these conditional laws. Disintegrating $\mathcal{H}(S|W) < \infty$ also gives $\mathcal{H}(S^x|W^x) < \infty$ for $\sigma_s$-almost every $x$. The required uniqueness among absolutely continuous laws is inherited from Brownian motion, as in Lemma 24(i). Since both initial laws are $\delta_x$, it gives

$$\mathcal{H}(Q^{v,x}|S^x_{[s,v]}) = \frac{1}{4}\,\mathbb{E}_{Q^{v,x}} \int_s^v |u_r(X_r)|^2 \,\mathrm{d}r.$$

The entropy disintegration formula then yields

$$\mathcal{H}(Q^v|S_{[s,v]}) = \mathcal{H}(\nu_s|\sigma_s) + \frac{1}{4}\int_s^v \int |u_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r.$$

*Step 4: Extension to the terminal time.* Extend $Q^v$ after $v$ using the reference transition law of $S$, and call the resulting full-path law $\hat{Q}^v$. This is possible because $\nu_v \ll \sigma_v$. The entropy chain rule and the preceding entropy identity give

$$\mathcal{H}(\hat{Q}^v|S) = \mathcal{H}(Q^v|S_{[s,v]}) \leqslant \mathcal{H}(\nu_s|\sigma_s) + \frac{1}{4}\int_s^t \int |u_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r.$$

Thus these laws are tight on $\mathcal{C}([s,t];\mathcal{X})$. Take a weak limit $P$ as $v \uparrow t$. For every $r < t$, eventually $v > r$, so $P_r = \nu_r$. Continuity of paths and of $\nu$ also gives $P_t = \nu_t$. Lower semicontinuity gives $\mathcal{H}(P|S) < \infty$.

On each $[s,w]$ with $w < t$, the marginals of all sufficiently late approximants are exactly $\nu$, and $b^S + u \in L^2(\nu\,\mathrm{d}r)$. Approximation of $b^S + u$ by bounded continuous fields in $L^1(\nu\,\mathrm{d}r)$ passes the martingale problem to the limit. Hence $P$ has drift $b^S + u$ on $[s,t)$. Apply finite-entropy Girsanov directly relative to $S$. It gives a global drift perturbation with finite squared energy; the interior martingale problem identifies it with $u$. Moreover, $\int_s^t |b^S_r(X_r)|^2 \,\mathrm{d}r < \infty$ holds $S$-almost surely, because $\mathcal{H}(S|W) < \infty$, and hence also $P$-almost surely. Thus the drift equation extends to the closed interval, and the same Girsanov identity gives (62).

*Step 5: Path moments and the Markov property.* The marginal moments are uniformly bounded by $W_2$ continuity. The path moment bound in Lemma 24(ii) therefore gives $P \in \mathcal{P}_2(\mathcal{C}([s,t];\mathcal{X}))$.

For two finite-entropy laws with the same initial law and state drift $b^S + u$, their mixture solves the same drift equation. Part (i) makes its entropy the average of their entropies, since its right-hand side is affine in the path law. Strict convexity of entropy forces the laws to coincide. Conditional future laws, given either the whole past or just the present, have finite entropy relative to the corresponding reference transition law, almost surely. They solve the same state-drift equation; the preceding uniqueness identifies them. Thus $P$ is Markov. □

**Corollary 26** (Markovization at fixed marginals)**.** *Let Assumption Diff hold and let $Q \in \mathcal{P}_2(\mathcal{C}([s,t];\mathcal{X}))$ satisfy $\mathcal{H}(Q|S) < \infty$. There is a Markov law $\bar{Q}$ with the same*

*marginals as $Q$ and $\mathcal{H}(\bar{Q}|S) \leqslant \mathcal{H}(Q|S)$. If $\alpha_r$ is the predictable drift perturbation of $Q$ relative to $S$, the drift of $\bar{Q}$ is*

$$b_r^S(x) + \bar{\alpha}_r(x), \qquad \bar{\alpha}_r(x) = \mathbb{E}_Q[\alpha_r \mid X_r = x].$$

*Proof of Corollary 26.* The marginals $\nu_r = Q_r$ are $W_2$ continuous and solve the Fokker–Planck equation with state drift $b^S + \bar{\alpha}$. Conditional Jensen gives

$$\int_s^t \int |\bar{\alpha}_r|^2 \, \mathrm{d}\nu_r \, \mathrm{d}r \leqslant \mathbb{E}_Q \int_s^t |\alpha_r|^2 \, \mathrm{d}r.$$

Lemma 25(ii) realizes this flow by a Markov law. The entropy identity in part (i) then gives the claimed entropy bound. □

## C.3 Relative velocities and entropy

Adding the forward and reversed entropy formulas gives the cost in terms of the relative velocity and score of the marginal flow. Subtracting them gives the entropy chain rule. We first prove these identities, including their endpoint values, and then minimize the velocity energy for a prescribed marginal curve.

**Lemma 27** (Time-symmetric entropy identity)**.** *Let Assumption Diff hold, with current velocity $v^S$ as in Lemma 24.*

*(i) For a Markov law $P$ with $\mathcal{H}(P|S) < \infty$, set*

$$\nu_r = P_r, \qquad h_r = \mathcal{H}(\nu_r|\sigma_r), \qquad g_r = \nabla \log(\nu_r/\sigma_r), \qquad w_r = b_r^P - b_r^S - g_r.$$

*Then $g$, $w \in L^2(\nu \, \mathrm{d}r)$, the marginals satisfy*

$$\partial_r \nu_r + \nabla \cdot \big(\nu_r (v_r^S + w_r)\big) = 0,$$

*and*

$$2\mathcal{H}(P|S) = h_s + h_t + \frac{1}{2} \int_s^t \int (|w_r|^2 + |g_r|^2) \, \mathrm{d}\nu_r \, \mathrm{d}r.$$

*This identity is invariant under simultaneous time reversal of $P$ and $S$. Moreover, $h \in AC([s,t])$ and*

$$h_v - h_u = \int_u^v \int g_r \cdot w_r \, \mathrm{d}\nu_r \, \mathrm{d}r, \qquad s \leqslant u \leqslant v \leqslant t.$$

*In particular,*

$$\int_s^t \left( \int |w_r|^2 \, \mathrm{d}\nu_r + \mathcal{I}(\nu_r|\sigma_r) \right) \mathrm{d}r \leqslant 4\mathcal{H}(P|S). \tag{63}$$

*(ii) Conversely, let $\nu \in \mathcal{C}([s,t]; \mathcal{P}_2(\mathcal{X}))$ have $\mathcal{H}(\nu_s|\sigma_s) < \infty$, and let $w$ be a Borel field. Suppose its weak score $g_r = \nabla \log(\nu_r/\sigma_r)$ exists, the continuity equation above holds, and $\int_s^t \int (|w_r|^2 + |g_r|^2) \, \mathrm{d}\nu_r \, \mathrm{d}r < \infty$. Then there is a Markov law $P$ with marginals $\nu_r$, drift $b^S + w + g$, and $\mathcal{H}(P|S) < \infty$. The entropy identity and chain rule in* (i) *hold for this law.*

*Proof of Lemma 27. (i) Interior identities.* Let $u^+$ and $u^-$ be the forward and reversed drift perturbations, indexed by original time. Entropy is invariant under path reversal, so Lemma 25(i) gives finite energy in both directions. On compact interior intervals, the time-reversal formula [14, Theorem 1.16] gives the weak score and

$$u^- = -u^+ + 2g, \qquad u^\pm = g \pm w.$$

Local boundedness of the reference drifts permits that theorem to be applied after stopping; reference uniqueness identifies its forward and reversed diffusion laws. In particular $g$, $w \in L^2(\nu\,\mathrm{d}r)$ on the whole interval. The two Girsanov formulas are

$$\begin{aligned}\mathcal{H}(P_{[u,v]}|S_{[u,v]}) &= h_u + \frac{1}{4}\int_u^v \int |w_r + g_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r,\\ &= h_v + \frac{1}{4}\int_u^v \int |w_r - g_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r.\end{aligned}$$

Adding the two formulas gives

$$2\mathcal{H}(P_{[u,v]}|S_{[u,v]}) = h_u + h_v + \frac{1}{2}\int_u^v \int (|w_r|^2 + |g_r|^2)\,\mathrm{d}\nu_r \,\mathrm{d}r.$$

Subtracting them gives $h_v - h_u = \int_u^v \int g_r \cdot w_r \,\mathrm{d}\nu_r \,\mathrm{d}r$. These identities initially hold on interior intervals. The continuity equation follows by writing the Fokker–Planck equation with current velocity $b^P - \nabla \log \nu = v^S + w$.

*Identifying the endpoint entropies.* Cauchy–Schwarz gives $g \cdot w \in L^1(\nu\,\mathrm{d}r)$. The interior chain rule therefore gives finite one-sided limits of $h$ at $s$ and $t$. These limits still have to be identified with the actual endpoint entropies. Data processing and the forward Girsanov formula give

$$h_r \leqslant \mathcal{H}(P_{[s,r]}|S_{[s,r]}) = h_s + \frac{1}{4}\int_s^r \int |w_q + g_q|^2 \,\mathrm{d}\nu_q \,\mathrm{d}q.$$

The energy tail vanishes as $r \downarrow s$, so $\limsup_{r\downarrow s} h_r \leqslant h_s$. Lower semicontinuity of relative entropy gives $h_s \leqslant \liminf_{r\downarrow s} h_r$. The backward formula gives the two corresponding inequalities at $t$. Thus the limits are $h_s$, $h_t$, the identities extend to $[s,t]$, and $h$ is absolutely continuous on the closed interval. Nonnegativity of the endpoint entropies gives (63).

*(ii) From relative velocities to a path law.* The given continuity equation is equivalent to

$$\partial_r \nu_r = \Delta \nu_r - \nabla \cdot \big(\nu_r (b_r^S + w_r + g_r)\big).$$

Since $w + g \in L^2(\nu\,\mathrm{d}r)$, Lemma 25(ii) gives a Markov realization with finite relative entropy. Its relative velocity is $w$, so part (i) gives both identities. □

We can now minimize the entropy at fixed marginals. For a prescribed $\nu \in \mathcal{C}([s,t];\mathcal{P}_2(\mathcal{X}))$, write $h_r = \mathcal{H}(\nu_r|\sigma_r)$ and $i_r = \mathcal{I}(\nu_r|\sigma_r)$. Assume $h_s + \int_s^t i_r \,\mathrm{d}r < \infty$, and let $\mathcal{W}(\nu)$ be the nonempty set of $w \in L^2(\nu_r\,\mathrm{d}r)$ satisfying the continuity equation in Lemma 27(i). The least path entropy for these marginals is given by

$$2 \min_{P:\, P_r = \nu_r \ \forall r} \mathcal{H}(P|S) = h_s + h_t + \frac{1}{2}\int_s^t i_r \,\mathrm{d}r + \frac{1}{2}\min_{w \in \mathcal{W}(\nu)} \int_s^t \int |w_r|^2 \,\mathrm{d}\nu_r \,\mathrm{d}r.$$

Indeed, the distributional continuity equation makes $\mathcal{W}(\nu)$ a closed affine subset of $L^2(\nu_r\,\mathrm{d}r)$. It has a unique element of least norm. Lemma 27(ii) realizes this velocity by a Markov law attaining the right-hand side. Any competing law of finite entropy has a finite path second moment by Lemma 24(ii). Corollary 26 and the time-symmetric identity give the matching lower bound. Strict convexity of entropy on the convex set of laws with these marginals gives uniqueness.

When only an endpoint marginal is prescribed, entropy disintegration selects the reference conditional path law given that endpoint.

**Corollary 28** (Evolution from a prescribed endpoint)**.** *Let Assumption Diff hold, fix $e \in \{s,t\}$, and let $\mu$ be a probability with $\mathcal{H}(\mu|\sigma_e) < \infty$. Set $p = \mathrm{d}\mu/\,\mathrm{d}\sigma_e$. The unique minimizer of $\mathcal{H}(Q|S)$ subject to $Q_e = \mu$ is $\mathrm{d}Q = p(X_e)\,\mathrm{d}S$, with $\mathcal{H}(Q|S) = \mathcal{H}(\mu|\sigma_e)$. It is Markov and has the reference transition kernels in the time direction starting from $e$. Write $\nu_r = Q_r$, $h_r = \mathcal{H}(\nu_r|\sigma_r)$, $i_r = \mathcal{I}(\nu_r|\sigma_r)$ and $g_r = \nabla \log(\nu_r/\sigma_r)$.*

*For $e = s$ the marginal flow solves the reference equation $\partial_r \nu_r = \Delta\nu_r - \nabla\cdot(\nu_r b_r^S)$. In original time, for every $s \leqslant a \leqslant b \leqslant t$,*

$$\begin{aligned} e = s: &\quad w_r = -g_r, \quad h_b + \int_a^b i_r\,\mathrm{d}r = h_a, \\ e = t: &\quad w_r = g_r, \quad\ \ h_a + \int_a^b i_r\,\mathrm{d}r = h_b. \end{aligned} \tag{64}$$

*In both cases the flow is $W_2$ continuous on the torus. On $\mathbb{R}^n$ the same holds if, for some $c > 0$,*

$$\sup_{s\leqslant r\leqslant t} \int e^{c|x|^2}\,\mathrm{d}\sigma_r < \infty.$$

*Proof of Corollary 28.* Disintegrating entropy at $X_e$ gives the endpoint entropy plus the mean conditional path entropy. The latter vanishes exactly when the conditional path laws agree with those of $S$, $\mu$-almost everywhere. This gives the density $p(X_e)$ and preserves the transition kernels starting from $e$, hence the Markov property. For $e = s$ the forward drift perturbation vanishes, so $w = -g$. For $e = t$ the reversed drift perturbation vanishes, giving $w = g$ in original time by the drift reversal formula in Lemma 27. Substituting $w = -g$ or $w = g$ in its chain rule gives (64), including the endpoint values.

The marginals are weakly continuous by path continuity. On the torus this gives $W_2$ continuity. On $\mathbb{R}^n$, data processing gives $h_r \leqslant \mathcal{H}(Q|S) = h_e$. The entropy inequality yields

$$c\int |x|^2\,\mathrm{d}\nu_r \leqslant h_e + \log\int e^{c|x|^2}\,\mathrm{d}\sigma_r.$$

The uniform bound gives the endpoint and integrated second moments required by Lemma 24(ii). Thus $Q$ has a finite path second moment, and the coupling $(X_a, X_b)$ under $Q$ gives $W_2$ continuity by dominated convergence. □

# References


[1] Manuel Arnese and Daniel Lacker. Sharp propagation of chaos for mean field Langevin dynamics, control, and games. 2026. arXiv: 2603.10988 [math.PR].

[2] Donald G. Aronson. Non-negative solutions of linear parabolic equations. *Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser.*, 22(4):607–694, 1968. URL: https://www.numdam.org/item/ASNSP_1968_3_22_4_607_0/.

[3] Julio Backhoff, Giovanni Conforti, Ivan Gentil and Christian Léonard. The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities. *Probab. Theory Relat. Fields*, 178(1–2):475–530, 2020. DOI: 10.1007/s00440-020-00977-8.

[4] Dominique Bakry, Ivan Gentil and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren Math. Wiss.* Springer, Cham, 2014. DOI: 10.1007/978-3-319-00227-9.

[5] Herm Jan Brascamp and Elliott H. Lieb. On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.*, 22(4):366–389, 1976. DOI: 10.1016/0022-1236(76)90004-5.

[6] Didier Bresch, Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.*, 172(13):2591–2641, 2023. DOI: 10.1215/00127094-2022-0088.

[7] Didier Bresch, Pierre-Emmanuel Jabin and Zhenfu Wang. Modulated free energy and mean field limit. In *Sémin. Laurent Schwartz, EDP Appl., Année 2019-2020*, Exposé n° II. DOI: 10.5802/slsedp.135.

[8] Pierre Cardaliaguet, Joe Jackson, Nikiforos Mimikos-Stamatopoulos and Panagiotis E. Souganidis. Sharp convergence rates for mean field control in the region of strong regularity. *Ann. Probab.*, 54(4):1839–1876, 2026. DOI: 10.1214/25-AOP1793.

[9] Pierre Cardaliaguet, Jean-Michel Lasry, Pierre-Louis Lions and Alessio Porretta. Long time average of mean field games. *Netw. Heterog. Media*, 7(2):279–301, 2012. DOI: 10.3934/nhm.2012.7.279.

[10] Pierre Cardaliaguet, Jean-Michel Lasry, Pierre-Louis Lions and Alessio Porretta. Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.*, 51(5):3558–3591, 2013. DOI: 10.1137/120904184.

[11] Pierre Cardaliaguet and Alessio Porretta. Long time behavior of the master equation in mean field game theory. *Anal. PDE*, 12(6):1397–1453, 2019. DOI: 10.2140/apde.2019.12.1397.

[12] Pierre Cardaliaguet and Panagiotis E. Souganidis. Regularity of the value function and quantitative propagation of chaos for mean field control problems. *NoDEA, Nonlinear Differ. Equ. Appl.*, 30(2), 25, 2023. DOI: 10.1007/s00030-022-00823-x.

[13] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications I. Mean field FBSDEs, control, and games*, volume 83 of *Probab. Theory Stoch. Model.* Springer, Cham, 2018. DOI: 10.1007/978-3-319-58920-6.

[14] Patrick Cattiaux, Giovanni Conforti, Ivan Gentil and Christian Léonard. Time Reversal of diffusion processes under a finite entropy condition. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 59(4):1844–1881, 2023. DOI: 10.1214/22-AIHP1320.

[15] Alekos Cecchin, Giovanni Conforti, Alain Durmus and Katharina Eichinger. The exponential turnpike phenomenon for mean field game systems: weakly monotone drifts and small interactions. *Electron. J. Probab.*, 31, 31, 2026. DOI: 10.1214/26-EJP1484.

[16] Louis-Pierre Chaintron, Giovanni Conforti and Katharina Eichinger. Propagation of weak log-concavity along generalised heat flows via Hamilton–Jacobi equations. 2025. arXiv: 2508.07931 [math.AP].

[17] Jean-François Chassagneux, Łukasz Szpruch and Alvin Tse. Weak quantitative propagation of chaos via differential calculus on the space of measures. *Ann. Appl. Probab.*, 32(3):1929–1969, 2022. DOI: 10.1214/21-AAP1725.

[18] Fan Chen, Zhenjie Ren and Songbo Wang. Uniform-in-time propagation of chaos for mean field Langevin dynamics. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 61(4):2357–2404, 2025. DOI: 10.1214/24-AIHP1499.

[19] Yongxin Chen, Tryphon T. Georgiou and Michele Pavon. On the relation between optimal transport and Schrödinger bridges: a stochastic control viewpoint. *J. Optim. Theory Appl.*, 169(2):671–691, 2016. DOI: 10.1007/s10957-015-0803-z.

[20] Marco Cirant and Nicolò De Bernardi. The local turnpike property in mean field control and games with quadratic hamiltonian. 2025. arXiv: 2511.18923 [math.AP]. URL: https://arxiv.org/abs/2511.18923.

[21] Marco Cirant and Alessio Porretta. Long time behavior and turnpike solutions in mildly non-monotone mean field games. *ESAIM, Control Optim. Calc. Var.*, 27, 86, 2021. DOI: 10.1051/cocv/2021077.

[22] Gauthier Clerc, Giovanni Conforti and Ivan Gentil. Long-time behaviour of entropic interpolations. *Potential Anal.*, 59(1):65–95, 2023. DOI: 10.1007/s11118-021-09961-w.

[23] Giovanni Conforti. A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost. *Probab. Theory Relat. Fields*, 174(1–2):1–47, 2019. DOI: 10.1007/s00440-018-0856-7.

[24] Giovanni Conforti. Coupling by reflection for controlled diffusion processes: turnpike property and large time behavior of Hamilton–Jacobi–Bellman equations. *Ann. Appl. Probab.*, 33(6A):4608–4644, 2023. DOI: 10.1214/22-AAP1927.

[25] Giovanni Conforti and Luca Tamanini. A formula for the time derivative of the entropic cost and applications. *J. Funct. Anal.*, 280(11), 108964, 2021. DOI: 10.1016/j.jfa.2021.108964.

[26] Samuel Daudin, François Delarue and Joe Jackson. On the optimal rate for the convergence problem in mean field control. *J. Funct. Anal.*, 287(12), 110660, 2024. DOI: 10.1016/j.jfa.2024.110660.

[27] Ivan Gentil, Christian Léonard and Luigia Ripani. Dynamical aspects of the generalized Schrödinger problem via Otto calculus – a heuristic point of view. *Rev. Mat. Iberoam.*, 36(4):1071–1112, 2020. DOI: 10.4171/rmi/1159.

[28] Maximilien Germain, Huyên Pham and Xavier Warin. Rate of convergence for particle approximation of PDEs in Wasserstein space. *J. Appl. Probab.*, 59(4):992–1008, 2022. DOI: 10.1017/jpr.2021.102.

[29] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Class. Math. Springer, Berlin, 2nd edition, 2001. DOI: 10.1007/978-3-642-61798-0. Reprint of the 1998 edition.

[30] Camilo Hernández and Ludovic Tangpi. Optimal transportation under McKean–Vlasov dynamics. 2026. arXiv: 2610.17626 [math.OC].

[31] Camilo Hernández and Ludovic Tangpi. Propagation of chaos for mean field Schrödinger problems. *SIAM J. Control Optim.*, 63(1):112–150, 2025. DOI: 10.1137/23M1566716.

[32] Richard Holley and Daniel Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46(5–6):1159–1194, 1987. DOI: 10.1007/BF01011161.

[33] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018. DOI: 10.1007/s00222-018-0808-y.

[34] Mark Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley, 1954–1955), volume 3, pages 171–197, Berkeley and Los Angeles, CA. University of California Press, 1956.

[35] Daniel Lacker. Limit theory for controlled McKean–Vlasov dynamics. *SIAM J. Control Optim.*, 55(3):1641–1672, 2017. DOI: 10.1137/16M1095895.

[36] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014. DOI: 10.3934/dcds.2014.34.1533.

[37] Christian Léonard. Girsanov theory under a finite entropy condition. In *Séminaire de probabilités XLIV*. Volume 2046, Lect. Notes Math. Pages 429–465. Springer, Berlin, 2012. DOI: 10.1007/978-3-642-27461-9_20.

[38] Christian Léonard. Stochastic derivatives and generalized $h$-transforms of Markov processes. 2011. arXiv: 1102.3172 [math.PR].

[39] Pierre Monmarché, Zhenjie Ren and Songbo Wang. Time-uniform log-Sobolev inequalities and applications to propagation of chaos. *Electron. J. Probab.*, 29, 154, 2024. DOI: 10.1214/24-EJP1217.

[40] Alessio Porretta. On the turnpike property for mean field games. *Minimax Theory Appl.*, 3(2):285–312, 2018.

[41] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press, New York etc., 1975, pages xv+361.

[42] Erwin Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. French. *Ann. Inst. Henri Poincaré*, 2(4):269–310, 1932. URL: https://www.numdam.org/item/AIHP_1932__2_4_269_0/.

[43] Erwin Schrödinger. Über die Umkehrung der Naturgesetze. German. *Sitzungsber. Preuß. Akad. Wiss., Phys.-Math. Kl.*, 1931(8–9):144–153, 1931.

[44] Sylvia Serfaty. Mean field limit for Coulomb-type flows. *Duke Math. J.*, 169(15):2887–2935, 2020. DOI: 10.1215/00127094-2020-0019. Appendix by Mitia Duerinckx and Sylvia Serfaty.

[45] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*. Volume 1464, Lect. Notes Math. Pages 165–251. Springer, Berlin, 1991. DOI: 10.1007/BFb0085169.

[46] Emmanuel Trélat and Enrique Zuazua. The turnpike property in finite-dimensional nonlinear optimal control. *J. Differ. Equations*, 258(1):81–114, 2015. DOI: 10.1016/j.jde.2014.09.005.

[47] Dario Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.*, 21, 22, 2016. DOI: 10.1214/16-EJP4453.

[48] Songbo Wang. Large-scale concentration and relaxation for mean-field Langevin particle systems. 2025. arXiv: 2508.16428 [math.PR].

[49] Songbo Wang and Wenbin Yan. Turnpike-type chaoticity for mean-field control problems. In preparation.