UNIFORM CONTRACTIVITY OF THE FISHER INFINITESIMAL MODEL WITH STRONGLY CONVEX SELECTION

VINCENT CALVEZ, DAVID POYATO, AND FILIPPO SANTAMBROGIO

Abstract. The Fisher infinitesimal model is a classical model of phenotypic trait inheritance in quantitative genetics. Here, we prove that it encompasses a remarkable convexity structure which is compatible with a selection function having a convex shape. It yields uniform contractivity along the flow, as measured by a L^∞ version of the Fisher information. It induces in turn asynchronous exponential growth of solutions, associated with a well-defined, log-concave, equilibrium distribution. Although the equation is non-linear and non-conservative, our result shares some similarities with the Bakry-Emery approach to the exponential convergence of solutions to the Fokker-Planck equation with a convex potential. Indeed, the contraction takes place at the level of the Fisher information. Moreover, the key lemma for proving contraction involves the Wasserstein distance W_∞ between two probability distributions of a (dual) backward-in-time process, and it is inspired by a maximum principle by Caffarelli for the Monge-Ampère equation.

1. Introduction

Let us consider the following nonlinear model

$$F_n = \mathcal{T}[F_{n-1}], \quad n \in \mathbb{N},\, x \in \mathbb{R}, \tag{1.1}$$

describing the evolution of the distribution $F_n = F_n(x)$ of a one-dimensional trait $x \in \mathbb{R}$, subject to sexual reproduction and the effect of selection at each generation. The operator $\mathcal{T}$ above is defined by

$$\mathcal{T}[F](x) := e^{-m(x)}\mathcal{B}[F](x), \quad x \in \mathbb{R}, \tag{1.2}$$

$$\mathcal{B}[F](x) := \iint_{\mathbb{R}^2} G\left(x - \frac{x_1 + x_2}{2}\right) F(x_1)\, \frac{F(x_2)}{\|F\|_{L^1}}\, dx_1\, dx_2, \quad x \in \mathbb{R}, \tag{1.3}$$

for any $F \in L^1_+(\mathbb{R}) \setminus \{0\}$. On the one hand, the operator $\mathcal{B}$ describes the distribution of traits of descendants of the previous generation F_{n-1}, arising as recombination of parental traits in agreement with *Fisher's infinitesimal model*, which is a classical model in quantitative genetics [6, 23]. Accordingly, the mixing kernel G is set to a centered Gaussian distribution with unit *segregation variance* without loss of generality, namely

$$G(x) := \frac{1}{(2\pi)^{1/2}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}. \tag{1.4}$$

On the other hand, the trait-dependent mortality function $m = m(x) \geq 0$ represents the effect of selection on the population, which acts multiplicatively over the descendants. In other words, the multiplicative factor $e^{-m(x)}$ in (1.2) represents the survival probability to the next generation of individuals having the trait x. We note that the time-discrete generations $n \in \mathbb{N}$ are assumed non-overlapping since, altogether, F_n describes the distribution of those offspring of F_{n-1} having survived after the selection step, and then different generations do not get mixed, see [16] for further insight.

2020 *Mathematics Subject Classification.* 35B40; 35P30; 35Q92; 47G20; 92D15.

Key words and phrases. Integro-differential equations, asymptotic behavior, nonlinear spectral theory, quantitative genetics, Monge-Ampère equation, maximum principle.

Acknowledgment. The authors are indebted to Thomas Lepoutre for stimulating discussions and valuable suggestions on the paper's presentation. We also thank Laurent Lafleche for pointing us the identity in Corollary A.2. VC and DP have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 865711). DP has received funding from the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement No 101064402, and partially from Grant C-EXP-265-UGR23 funded by Consejería de Universidad, Investigación e Innovación & ERDF/EU Andalusia Program, from Grant PID2022-137228OB-I00 funded by the Spanish Ministerio de Ciencia, Innovación y Universidades, MICIU/AEI/10.13039/501100011033 & "ERDF/EU A way of making Europe", and from Modeling Nature (MNAT) project QUAL21-011 funded by Junta de Andalucía. VC thanks the France 2030 framework programme Centre Henri Lebesgue ANR-11-LABX-0020-01 for creating an attractive mathematical environment. FS acknowledges the support of the Lagrange Mathematics and Computation Research Center project on Optimal Transportation.

As the model is tracking only one trait distribution, it applies either when individuals are hermaphroditic, or when the traits are equally distributed between male and female individuals within the population. We refer to [6] for a comprehensive presentation of the model, its derivation and its limitations.

The goal of this paper is to extend the studies initiated in [16] to a broader class of selection functions. Specifically, when m is a strongly convex function we prove *asynchronous exponential growth* [45] of solutions to (1.1). In other words, we derive quantitative rates for the relaxation of the solutions $\{F_n\}_{n\in\mathbb{N}}$ of (1.1) to a strongly log-concave quasi-equilibrium of the form $\boldsymbol{\lambda}^n\,\boldsymbol{F}$, where $\boldsymbol{\lambda}>0$ and $\boldsymbol{F}\in L^1(\mathbb{R})\cap\mathcal{P}(\mathbb{R})$ is an appropriate probability density. The fact that the quasi-equilibrium is strongly log-concave is crucial in our approach and will be present all along the paper.

Definition 1.1 (Log-concavity). *Consider any non-negative function $F=e^{-V}:\mathbb{R}^d\longrightarrow\mathbb{R}_+$.*

- *(i) F is said log-concave when V is a convex function.*
- *(ii) F is said strongly log-concave with log-concavity parameter $\gamma>0$ (or γ-log-concave) when V is a strongly convex function with convexity parameter γ (or γ-convex).*

When the potential function $V\in C^2(\mathbb{R}^d)$, we can equivalently formulate log-concavity in terms of second order derivatives. Namely, F is log-concave when $D^2V\geq 0$, and F is γ-log-concave when $D^2V\geq\gamma I_d$.

We remark that in order for an anstaz of the form $F_n(x)=\lambda^n F(x)$ to define a solution to (1.1), we need that the pair (λ,F) solves the following nonlinear eigenproblem:

$$\begin{aligned}&\lambda F=\mathcal{T}[F],\quad x\in\mathbb{R},\\&F\geq 0,\quad \int_{\mathbb{R}}F(x)\,dx=1.\end{aligned}\tag{1.5}$$

Hence, the possible quasi-equilibria are to be found as solutions to (1.5). Note that contrarily to the special quadratic regime treated in [16], the Gaussian structure can no longer be exploited and, in particular, the existence of solutions to (1.5) is unclear. Indeed, the above non-linear integral operator is 1-homogenenous but non-monotone, and therefore the Krein-Rutman theorem [31] cannot be applied as it has been done in other (usually linear) problems in population dynamics [7, 19]. Hence, the study of the non-linear evolution problem (1.1) and the non-linear eigenproblem (1.5) requires innovative ideas.

Along this paper, we address jointly the following two problems: (i) Existence of a strongly log-concave solution $(\boldsymbol{\lambda},\boldsymbol{F})$ to (1.5), and (ii) Quantitative relaxation of the solutions to (1.1) towards the quasi-equilibrium $\boldsymbol{\lambda}^n\,\boldsymbol{F}$. We make the crucial hypothesis that m is a strongly convex function,

$$m''\geq\alpha\quad\text{for some}\quad\alpha>0,\tag{H1}$$

The function m necessarily reaches its minimum value over $\mathbb{R}$. For convenience, we assume the following additional hypothesis without loss of generality,

$$m\geq 0,\quad\text{and}\quad m(0)=0.\tag{H2}$$

The L^∞ relative Fisher information $\mathcal{I}_\infty$ plays a pivotal role in our analysis, as it measures the contractivity along the flow (see methodological notes below). It is defined as follows, for a pair of functions $P,Q\in L^1_+(\mathbb{R})\cap C^1(\mathbb{R})$,

$$\mathcal{I}_\infty(P\|Q):=\left\|\frac{d}{dx}\left(\log\frac{P}{Q}\right)\right\|_{L^\infty}.\tag{1.6}$$

Theorem 1.2. *Let $m\in C^2(\mathbb{R}^d)$ satisfy* (H1)-(H2)*. Then, the following statements hold true:*

- *(i)* (**Existence of quasi-equilibrium**)
  There is at least one solution $(\boldsymbol{\lambda},\boldsymbol{F})$ to (1.5)*. In addition, $\boldsymbol{F}=e^{-\boldsymbol{V}}\in L^1_+(\mathbb{R})\cap C^\infty(\mathbb{R})$ is β-log-concave, where $\beta>\frac{1}{2}$ is uniquely defined by the following relationship*

$$\beta=\alpha+\frac{2\,\beta}{1+2\,\beta}.\tag{1.7}$$

  Moreover, $(\boldsymbol{\lambda},\boldsymbol{F})$ is the unique solution to (1.5) *among all pairs (λ,F) such that*

$$\frac{d}{dx}\left(\log\frac{F}{\boldsymbol{F}}\right)\in L^\infty(\mathbb{R}).\tag{1.8}$$

- *(ii)* (**One-step contraction**)
  Consider any $F_0\in L^1_+(\mathbb{R})\cap C^1(\mathbb{R})$ such that

$$\frac{d}{dx}\left(\log\frac{F_0}{\boldsymbol{F}}\right)\in L^\infty(\mathbb{R}),\tag{H3}$$

and let $\{F_n\}_{n\in\mathbb{N}}$ *be the solution to* (1.1) *issued at* F_0. *Then, we have*

$$\mathcal{I}_\infty\left(F_n\| \boldsymbol{F}\right) \leq \frac{2}{1+2\,\beta}\,\mathcal{I}_\infty\left(F_{n-1}\| \boldsymbol{F}\right), \tag{1.9}$$

for any $n \in \mathbb{N}$.

(iii) (**Asynchronous exponential growth**)
Consider any $F_0 \in L^1_+(\mathbb{R}) \cap C^1(\mathbb{R})$ *verifying the assumption* (H3) *above, and let* $\{F_n\}_{n\in\mathbb{N}}$ *be the solution to* (1.1) *issued at* F_0. *Then, we have*

$$\left|\frac{\|F_n\|_{L^1}}{\|F_{n-1}\|_{L^1}} - \boldsymbol{\lambda}\right| \leq C\left(\frac{2}{1+2\,\beta}\right)^n, \tag{1.10}$$

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1}}\middle\| \boldsymbol{F}\right) \leq C\left(\frac{2}{1+2\,\beta}\right)^{2n}, \tag{1.11}$$

for every $n \in \mathbb{N}$, *where* $C > 0$ *is a explicit constant depending on* F_0, *and* $\mathcal{D}_{\mathrm{KL}}$ *is the Kullback-Leibler divergence (or relative entropy), that is,*

$$\mathcal{D}_{\mathrm{KL}}(P\|Q) := \int_{\mathbb{R}} \log\left(\frac{P(x)}{Q(x)}\right) P(x)\,dx, \quad P, Q \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R}). \tag{1.12}$$

Remark 1.3 (Case of quadratic selection). *For quadratic selection* $m(x) = \frac{\alpha}{2}|x|^2$, *we have that* m *satisfies the hypothesis* (H1)-(H2) *in Theorem 1.2, and then our new result applies. Such a special case was studied in detail in* [16], *where in particular it was proven that there is a unique eigenpair* $(\boldsymbol{\lambda}, \boldsymbol{F})$ *of* (1.5), *which involves a Gaussian eigenfunction* $\boldsymbol{F}(x) = G_{0,\sigma^2}(x)$ *with variance* $\sigma^2 > 0$ *verifying*

$$\frac{1}{\sigma^2} = \alpha + \frac{1}{1+\frac{\sigma^2}{2}}. \tag{1.13}$$

In particular, $\boldsymbol{F}$ *is* $\frac{1}{\sigma^2}$*-log-concave (cf. Definition 1.1), which is compatible with our new result in view of the identity* $\sigma^2 = \beta^{-1}$ *stemming from equations* (1.7) *and* (1.13). *Furthermore, the contraction factor in* (1.9) *predicted by Theorem 1.2 also recovers the one obtained in* [16] *for quadratic selection. Specifically,*

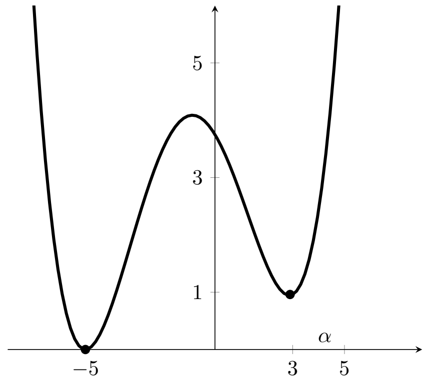
$$\frac{2}{1+2\,\beta} = \frac{(3+2\,\alpha) - \sqrt{(3+2\,\alpha)^2 - 8}}{2},$$

which agrees precisely with the contraction factor found in [16, Lemma 6.3].

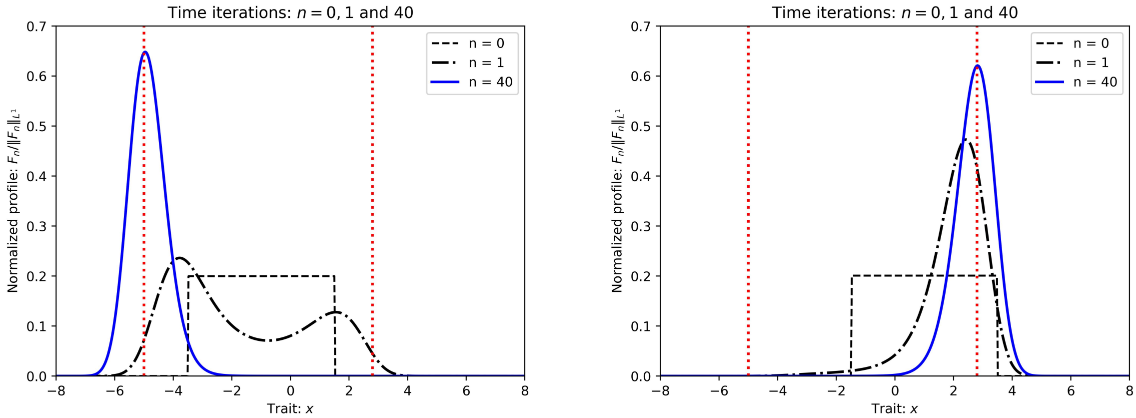
Remark 1.4 (Close-to-equilibrium initial data). *In contrast with* [16], *where the above framework was restricted to* $m(x) = \frac{\alpha}{2}|x|^2$ *but generic* $F_0 \in \mathcal{M}_+(\mathbb{R})$, *Theorem 1.2 applies to a broader class of selection functions verifying* (H1)-(H2) *at the cost of restricting to initial data fulfilling the hypothesis* (H3). *Specifically, such a condition imposes a precise behavior of the tails of* F_0, *which must be very close to those of the eigenfunction* $\boldsymbol{F}$ *(in particular, two Gaussian initial distributions should have the same variance).*

Remark 1.5 (Conditional uniqueness). *Another difference with* [16] *is that the current approach does not guarantee global uniqueness of solutions to the eigenproblem* (1.5), *but only within the class of eigenpairs verifying* (1.8). *Nevertheless, we conjecture that global uniqueness holds true, as in the quadratic case* $m(x) = \frac{\alpha}{2}|x|^2$. *Proving global uniqueness would require a careful control of the behavior at infinity, in the spirit of* [16], *which is beyond the scope of this paper.*

Remark 1.6 (On the convexity assumption). *The convexity assumption* (H1) *ensures that* m *must have a unique minimum. It implies that the quasi-equilibrium* $\boldsymbol{F}$ *obtained in Theorem 1.2 is log-concave, as a consequence of the Prékopa-Leindler inequality. In the presence of multiple local minima of* m, *it was proven in* [15, Corollary 1.5] *that several quasi-equilibria could co-exist in the time-continuous version of* (1.1) *provided that the variance of kernel* (1.4) *is small enough (in original units). That is, in the case of non-convex* m *there is evidence that the generalized eigenproblem* (1.5) *may admit non-unique solutions, in contrast with general conclusions of the Krein–Rutman theory in the linear case. This is illustrated by numerical simulations shown in Figure 1, where two different quasi-equilibria (one of them bimodal) are found numerically if* m *has two minima. A similar behaviour can be observed in a population adapting to a heterogeneous, patchy environment, when each patch is associated with a different optimal trait* [20]. *The same conclusions also hold for the (continuous) time-marching problem in* [40, 37, 27].



(a) Double-well selection function



(b) Non-uniqueness of quasi-equilibria for the double-well selection function

Figure 1. (a) Double-well selection function $m(x) = 0.015\,((x-3)^2+1)(x+5)^2$ used in the simulations. (b) Time-evolution of the normalized profiles $F_n/\|F_n\|_{L^1}$ up to generation $n = 40$ (solid line) for two different choices of initial datum F_0. On the left, $F_0 = \mathbb{1}_{[-3.5,\,1.5]}$ leads to concentration near the left (globally) optimal trait. On the right, $F_0 = \mathbb{1}_{[-1.5,\,3.5]}$ leads to concentration near the right (locally) optimal trait.

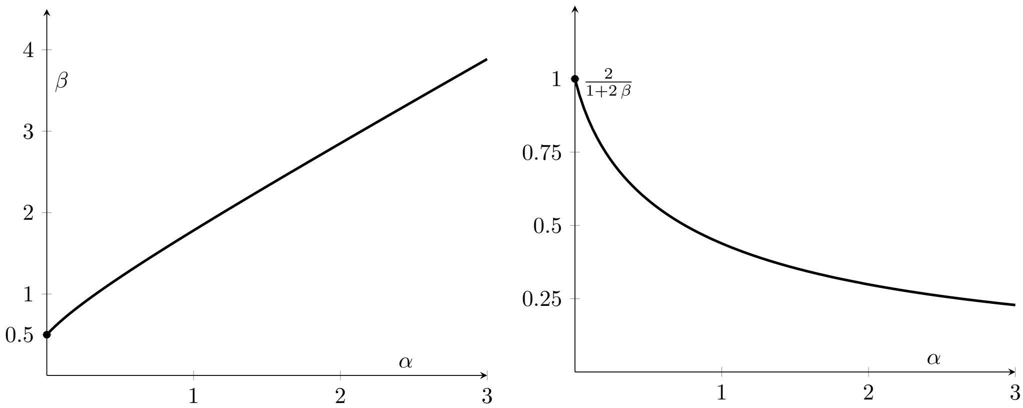
Remark 1.7 (Log-concavity and contraction factor)**.** *For any $\alpha > 0$, we have that the log-concavity parameter β in* (1.7) *and the corresponding contraction factor $\frac{2}{1+2\,\beta}$ in* (1.9) *satisfy the following properties:*

$$\alpha \searrow 0 \quad \Longrightarrow \quad \beta \searrow \frac{1}{2} \quad \text{and} \quad \frac{2}{1+2\,\beta} \nearrow 1,$$
$$\alpha \nearrow \infty \quad \Longrightarrow \quad \beta \nearrow \infty \quad \text{and} \quad \frac{2}{1+2\,\beta} \searrow 0,$$

see Figure 2. In particular, we have genuine contraction in (1.9) *since $0 < \frac{2}{1+2\,\beta} < 1$ for every $\alpha > 0$.*

Remark 1.8 (One-dimensional traits)**.** *In this paper we restrict to one-dimensional traits, but note that an analogous version of* (1.1) *and* (1.5) *makes sense in higher dimensions yet. In fact, these were studied in* [16] *for quadratic selection functions. However, a higher-dimensional version of our result for generic strongly convex selection function would require some non-trivial improvements of the present methods. Just to emphasize some non-trivial obstructions, we remark that our approach exploits a maximum principle for the Monge-Ampère equation in convex but not uniformly-convex domains, as described below. In this setting, it is not even clear why the standard elliptic regularity should hold up to the boundary, as in the seminal works* [12]*. In two-dimensional domains with special symmetries, this theory has been developed recently in* [28]*, but a higher dimensional extension would require further work which goes beyond the scope of this paper. The extension to any dimension was achieved in* [29]*, which was released during the time of revision of the present work.*

Bibliographical notes.



(a) Log-concavity parameter β as a function of α (b) Contraction parameter $\frac{2}{1+2\beta}$ as a function of α

FIGURE 2. Plot of the log-concavity parameter β of the eigenfunction $\boldsymbol{F}$ and the contraction parameter $\frac{2}{1+2\beta}$ in Theorem 1.2 as a function of α

This work can be viewed as another brick to combine optimal transportation tools for non-conservative problems arising in biology. The connection between the Fisher infinitesimal model and the L^2 Wasserstein distance was spotted by G. RAOUL [39] (see also [32] for similar results in a different context of protein exchanges between cells). In fact, when there is no selection (that is, $m \equiv const$), the operator $\mathcal{T}$ is non-expansive for the latter distance. Contraction cannot be expected because of translational invariance. Nevertheless, it is contractive with rate $1/\sqrt{2}$ in the class of distributions having the same center of mass (the latter being preserved by the flow) [39, Theorem 4.1 and Corollary 4.2]. This remarkable structure was further exploited by G. RAOUL [40] in a perturbative setting, when selection is small (in amplitude), and restricted to a compact interval (m is constant beyond a certain range). More precisely, G. RAOUL proved that the dynamics is well captured by some averaged quantities ("moments") of the Gaussian distribution coupled with the selection function, provided that the initial data is well-prepared, in the basin of attraction of the stationary state, and the amplitude of selection is small enough. For that purpose, he carefully established that the contraction issued from the infinitesimal operator was robust enough to dominate detrimental effects due to selection. Note that the later references consider overlapping generations, that is, a continuous-in-time rather than discrete dynamics. However, some fruitful analogy can be drawn between the results and methodology.

In parallel, the regime of small segregation variance (when G (1.4) has variance ε^2 and ε is small enough) was investigated by [15, 37] in another perturbative setting, without exploiting the Wasserstein metric structure. This methodology built upon the seminal works on vanishing viscosity limits associated with linear (asexual) modes of reproduction in quantitative genetics models [21, 38, 5]. Interestingly, it was proven in [15] that the problem (1.5) lacks uniqueness in full generality. More precisely, it was possible to build a solution to (1.5) centered in the vicinity of any local minimum of m, provided that the selection value at the local minimum is close enough to the global minimum. This result entails a clear separation with linear, order-preserving operators (and non-linear extensions [31, 33]) for which (1.5) genuinely admits a unique solution (under standard irreducibility assumptions), see Remark 1.6. The Cauchy problem initialized with some concentrated initial data was further investigated in [37] (in a multiplicative perturbative approach) and more recently in [27] (in a moment-based approach), still in the regime of small segregation variance. The case of zero segregation variance was the subject of the recent [25].

Heuristically, uniqueness of the (non-linear) eigenpair $(\boldsymbol{\lambda}, \boldsymbol{F})$ is rather clear when the selection function m is convex, and [16] was a first contribution in this direction, restricted to $m(x) = \frac{\alpha}{2}|x|^2$. By exploiting the quadratic structure of the operator $\mathcal{T}$ in (1.2) (which involves products and convolutions by Gaussian density functions), it was possible to prove asynchronous exponential growth towards the explicit Gaussian distribution of equilibrium $\boldsymbol{F}$, starting from any initial configuration F_0. This was achieved by a careful study of the binary tree of ancestors, together with explicit change of variables in a high-dimensional integral, to prove a sort of concentration of measure estimates. More precisely, it was shown that the traits of the ancestors decorrelate sufficiently fast, backward in the tree, from the trait of the individual at generation n. This implies that the dependence of the trait distribution F_n at generation n upon the initial distribution F_0 diminishes exponentially fast. Asynchronous exponential growth is a consequence of this observation, which is a backward feature.

Last, but not least, let us mention that both the infinitesimal model (1.2), and the relative information (1.6) (or rather (1.18) below) date back to a couple of seminal works by R. A. FISHER in the same years (circa 1920) on seemingly different purposes, respectively [23] and [24], see [42] for a discussion.

Methodological notes.

In the present study, we push further the observations of [16]. We identify a key mechanism ensuring a one-step contraction for the flow (1.1). This can be summarized roughly as follows:

> *For any two given individuals with traits X and X' respectively, the associated parental traits (X_1, X_2) and (X_1', X_2') are closer to each other than X and X' are, in some sense,*

see also [26, Appendix F.2] for a visual explanation. To make sense of this contraction, we shall work with the L^∞ Wasserstein distance, denoted by W_∞ (in contrast with the L^2 Wasserstein distance). This naturally leads to estimates on the so-called L^∞ relative Fisher information $\mathcal{I}_\infty$ (1.6) (in contrast with the (L^2) relative Fisher information $\mathcal{I}_2$, see (1.18) below). The core estimate (1.9) is forward in time, and it naturally arises as a dual estimate of a backward in time estimate analogous to the work in [16].

A forward-backward argument. We propose a short warm-up to this argument, which may help the reader follow our method (without details of the proofs). Indeed, one complication of our setting is that each individual has two parents, so that the dimension of the distribution doubles at each generation. Nonetheless, the same methodology can be applied to the case of a single parent, which boils down to a *linear operator*. We thus consider, temporarily, the following linear operator:

$$\mathcal{A}[F](x) := e^{-m(x)} \int_{\mathbb{R}} G\,(x-y)\ F(y)\,dy, \quad x \in \mathbb{R}, \tag{1.14}$$

in place of the above non-linear operator $\mathcal{T}$ in (1.2). In this simpler case, the Krein-Rutman theorem can be applied (at least formally), and there exists an eigenpair $(\boldsymbol{\lambda}, \boldsymbol{F})$ of the linear eigenproblem (1.5) with $\mathcal{T}$ replaced by $\mathcal{A}$. Now, consider any solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1.1) with $\mathcal{T}$ replaced again by the linear operator $\mathcal{A}$. We may introduce the associated relative distribution $u_n = \frac{F_n}{\boldsymbol{\lambda}^n \boldsymbol{F}}$ to follow the trend of F_n across generations. It satisfies the following equation:

$$u_n(x) = \frac{\int_{\mathbb{R}} G(x-y)\, u_{n-1}(y)\, \boldsymbol{F}(y)\, dy}{\int_{\mathbb{R}} G(x-z)\, \boldsymbol{F}(z)\, dz} = \int_{\mathbb{R}} \boldsymbol{P}(x;y)\, u_{n-1}(y)\, dy, \quad n \in \mathbb{N},\ x \in \mathbb{R},$$

where the x-dependent probability distribution function $\boldsymbol{P}(x;\cdot)$ is defined as

$$\boldsymbol{P}(x;y) = \frac{G(x-y)\, \boldsymbol{F}(y)}{\int_{\mathbb{R}} G(x-z)\, \boldsymbol{F}(z)\, dz}, \quad x, y \in \mathbb{R}, \tag{1.15}$$

and it can be interpreted as the transition probability from trait y to trait x. The fact that it is a probability distribution function, $\int \boldsymbol{P}(x;y)\, dy = 1$, is immediate by the choice of the normalization, which is such that constant functions $u_n \equiv const$ are invariant by the flow.

Next, it can be proven that, if $\boldsymbol{F}$ is strongly log-concave, then we have

$$W_\infty(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(x',\cdot)) \le \kappa\, |x - x'|, \tag{1.16}$$

where $\kappa \in (0,1)$ is related to the modulus of convexity of $\boldsymbol{V} = -\log \boldsymbol{F}$. By duality, this backward contraction estimate results in the forward estimate below (*cf.* Lemma 2.4)

$$\left\| \frac{d}{dx} (\log u_n) \right\|_{L^\infty} \le \kappa \left\| \frac{d}{dx} (\log u_{n-1}) \right\|_{L^\infty},$$

which by iteration and using the L^∞ relative Fisher information, it can be expressed as follows

$$\mathcal{I}_\infty(F_n \| \boldsymbol{F}) \le \kappa^n\, \mathcal{I}_\infty(F_0 \| \boldsymbol{F}). \tag{1.17}$$

As mentioned in Remark 1.8, the key estimate (1.16) is a consequence of the maximum principle on the Monge-Ampère equation for the optimal transportation plan between $\boldsymbol{P}(x;\cdot)$ and $\boldsymbol{P}(x';\cdot)$. Interestingly, this is an argument borrowed from the theory of conservative equations, whereas our problem is not. The trick is to match an individual to its ancestor, which is obviously a conservative process, backward in time.

Analogy with the Bakry-Emery argument. There is some analogy between our results and the standard Bakry-Emery method for exponential relaxation towards equilibrium for the gradient flow of some displacement convex "entropy", for instance, the Fokker-Planck equation with a convex potential [3, 2, 43, 4]. Indeed, from (1.9) (alternatively (1.17) in the linear case) we obtain exponential convergence on a quantity which is the L^∞ analog of the usual (L^2) relative Fisher information,

$$\mathcal{I}_2(P\|Q) := \int_{\mathbb{R}} \left| \frac{d}{dx}\left(\log \frac{P}{Q}\right)(x)\right|^2 P(x)\, dx. \tag{1.18}$$

Recall that, in the usual Bakry-Emery argument, the exponential convergence is established at the level of the dissipation of entropy, that is, the usual relative Fisher information [43]. In turn, the exponential relaxation of the dissipation is intimately linked with the displacement convexity of the entropy functional (essentially because the gradient flow is differentiated, which leads to the second derivative of the entropy functional). In our argument, it is the convexity of $\boldsymbol{V} = -\log \boldsymbol{F}$ which induces the geometrical relaxation of the uniform relative Fisher information.

Connection with another projective metric. The uniform relative Fisher information (1.6) may also be viewed as a kind of first order version of the *Hilbert's projective distance* associated with the cone of non-negative functions, that is,

$$\mathfrak{H}(P,Q) := \operatorname{osc}\left(\log \frac{P}{Q}\right) \equiv \sup_{x\in\mathbb{R}} \log \frac{P(x)}{Q(x)} - \inf_{x\in\mathbb{R}} \log \frac{P(x)}{Q(x)}.$$

The latter distance is well-suited for the analysis of 1-positively homogeneous, order-preserving, operators [33]. An obvious reason is the projective character of that metric [34], which makes it insensitive to the exponential growth (or decay) $\mathcal{O}(\boldsymbol{\lambda}^n)$. This character is also shared by $\mathcal{I}_\infty$ (in contrast with $\mathcal{I}_2$).

A linear argument, even in the non-linear case. The previous discussion focussed on the linear operator (1.14) for the sake of clarity. Interestingly, the non-linear case under study (1.2) also involves a linear argument when formulated backward in time. Similarly, define the relative distribution $u_n = \frac{F_n}{\boldsymbol{\lambda}^n \boldsymbol{F}}$, where the pair $(\boldsymbol{\lambda}, \boldsymbol{F})$ is the strongly log-concave solution to (1.5) from part (i) in the main Theorem 1.2. Then, u_n satisfies the following forward-in-time non-linear problem:

$$u_n(x) = \frac{1}{\|u_{n-1}\, \boldsymbol{F}\,\|_{L^1}} \iint_{\mathbb{R}^d} \boldsymbol{P}(x; x_1, x_2)\, u_{n-1}(x_1)\, u_{n-1}(x_2)\, dx_1\, dx_2, \quad n \in \mathbb{N}, \quad x \in \mathbb{R}, \tag{1.19}$$

where the function $\boldsymbol{P}(x; x_1, x_2)$ is explicitly defined as

$$\boldsymbol{P}(x; x_1, x_2) = \frac{G\left(x - \frac{x_1+x_2}{2}\right)\, \boldsymbol{F}(x_1)\, \boldsymbol{F}(x_2)}{\iint_{\mathbb{R}^2} G\left(x - \frac{x_1'+x_2'}{2}\right)\, \boldsymbol{F}(x_1')\, \boldsymbol{F}(x_2')\, dx_1'\, dx_2'}, \quad x \in \mathbb{R},\, (x_1, x_2) \in \mathbb{R}^2. \tag{1.20}$$

Since $\boldsymbol{P}$ is normalized with respect to the variables (x_1, x_2), then it can be regarded as a Markov kernel with source $x \in \mathbb{R}$ and target $(x_1, x_2) \in \mathbb{R}^2$ representing the probability of transitioning from the trait of the offspring x to the traits of the parents (x_1, x_2). In Lemma 2.6, we prove the very same contraction estimate as in (1.16) for the family of Markov kernels $\boldsymbol{P}$ indexed by its first variable x. The key difference is that this Markov kernel makes the transition between u_n and $u_{n-1} \otimes u_{n-1}$ due to the joint distribution of parental traits (the non-linearity, in fact). This is rescued by an appropriate tensorization property of the relative Fisher information, which is expressed in Lemma 2.4.

A close-to-optimal result despite a non-optimal argument. The rate of contraction $\frac{2}{1+2\beta}$ coincides with the optimal one in the quadratic case (see Remark 1.3). However, there is some non-optimal step in the proof. Indeed, our key contraction estimate (1.16) is a consequence of the maximum principle on the Monge-Ampère equation satisfied by the Brenier transportation map between the joint distributions of the parental traits (X_1, X_2) and (X_1', X_2'). There is some subtlety here to be noticed, as the contraction is set for the L^∞ Wasserstein distance (maximum of the optimal transportation displacement), whereas the Brenier transportation map used in our argument is optimal for the L^2 Wasserstein distance. Nevertheless, in the quadratic case, the transportation map is simply a translation, so that it comes with the same cost, measured either in (weighted) L^2 or in L^∞.

In the recent contribution [29], the authors used a different approach based on Langevin dynamics to make the connection between the two joint distributions. Hence, they by-pass the use of the Brenier map. Their approach is much simpler, and it enables to extend the result readily to higher dimensions. These results were originally motivated by a computation in a previous version of our paper, where we obtained an upper bound on the displacement $||T(x) - x||_2$ for the Brenier map between a strongly log-concave density and a perturbation of it. In the current version, such an estimate cited by [29] is no more crucial, as the important one concerns the displacement $||T(x)-x||_1$ (see Sections 2.3 and 2.4) and interpolating ℓ_1 estimates from ℓ_2 ones worsens the coefficients (*cf.* Remark 3.1). Then we have moved the ℓ_2 estimates to Appendix C for an easier readability. In [29], the authors by-pass this delicate issue of choosing ℓ_1 rather than ℓ_2 based distances by establishing some fruitful anisotropic version of our Lemma C.2.

Organization of the paper.

In Section 2 we provide a sketch of the proof of the one-step contraction property in Theorem 1.2(ii) under an additional technical condition. In Section 3 we derive the fundamental contraction property of the one-step transition probability of the problem under the $W_{\infty,1}$ Wasserstein distance (see definition below), thus removing the technical condition used in the sketch of proof of Section 2. In Section 4 we analyze a truncated version of the time-marching problem (1.5) to bounded intervals, which will be necessary in next part. Section 5 focuses on proving the existence of strongly log-concave solutions of the nonlinear eigen-problem (1.5) as claimed in Theorem 1.2(i). In Section 6 we prove asymptotic exponential growth of (1.5) for restricted initial data (H3) as in Theorem 1.2(iii). Finally, Appendices (A) and (B) contain some technical results to alleviate the reading of the paper.

Notation.

• **(Vector norms)** Along the paper, $\mathbb{R}^d$ will be endowed with the various ℓ_q norms, namely, for any $z=(z_1,\ldots,z_d)\in\mathbb{R}^d$ and any $1\leq q\leq\infty$ we denote

$$\|z\|_q := \left\{ \begin{array}{ll} \left(\displaystyle\sum_{i=1}^d |z_i|^q\right)^{1/q}, & \text{if } 1\leq q<\infty,\\ \displaystyle\max_{1\leq i\leq d}|z_i|, & \text{if } q=\infty. \end{array}\right. \tag{1.21}$$

The associated ℓ_2 and ℓ_∞ open balls centered at 0 with radius $R>0$ are respectively denoted by

$$B_R:=\{z\in\mathbb{R}^d:\ \|z\|_2<R\}\quad\text{and}\quad Q_R:=\{z\in\mathbb{R}^d:\ \|z\|_\infty<R\}. \tag{1.22}$$

• **(Characteristic function)** Given any set $A\subset\mathbb{R}^d$, we will denote the associated characteristic function of convex analysis by $\chi_A:\mathbb{R}^d\longrightarrow(-\infty,+\infty]$, which is the mapping defined by

$$\chi_A(z):=\left\{\begin{array}{lll} 0, & \text{if} & z\in A,\\ +\infty, & \text{if} & z\in\mathbb{R}^d\setminus A. \end{array}\right. \tag{1.23}$$

• **(Measure spaces)** We denote by $\mathcal{M}(\mathbb{R}^d)$ the space of finite Radon measures, endowed with the total variation norm, and $\mathcal{M}^+(\mathbb{R}^d)$ represents the cone of non-negative finite Radon measures. Similarly, $\mathcal{P}(\mathbb{R}^d)$ is the subspace of probability measures, endowed with the narrow topology except otherwise specified.

• **(Wasserstein metrics)** For any $1\leq p\leq\infty$, we define the L^p Wasserstein space

$$\begin{aligned} \mathcal{P}_p(\mathbb{R}^d)&:=\left\{P\in\mathcal{P}(\mathbb{R}^d):\ \int_{\mathbb{R}^d}|z|^p\,P(dz)<\infty\right\},\quad \text{if } 1\leq p<\infty,\\ \mathcal{P}_\infty(\mathbb{R}^d)&:=\left\{P\in\mathcal{P}(\mathbb{R}^d):\ \operatorname{supp}P \text{ is compact}\right\}. \end{aligned}$$

Similarly, we consider the L^p Wasserstein metric associated with the ℓ_q vector norm of $\mathbb{R}^d$. Specifically, for any $P,Q\in\mathcal{P}(\mathbb{R}^d)$ and any $1\leq p,q\leq\infty$ we denote

$$\begin{aligned} W_{p,q}(P,Q)&:=\left(\inf_{\gamma\in\Gamma(P,Q)}\int_{\mathbb{R}^{2d}}\|z-\tilde z\|_q^p\,\gamma(dz,d\tilde z)\right)^{1/p},\quad \text{if } 1\leq p<\infty,\\ W_{\infty,q}(P,Q)&:=\inf_{\gamma\in\Gamma(P,Q)}\ \gamma\text{-}\operatorname*{ess\,sup}_{z,\tilde z\in\mathbb{R}^d}\|z-\tilde z\|_q, \end{aligned} \tag{1.24}$$

where $\Gamma(P,Q)$ is the family of transference plan $\gamma\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R})^d$ with marginals P and Q. Whilst the L^p Wasserstein distances could be infinitely-valued over $\mathcal{P}(\mathbb{R}^d)$, note that they take finite values over $\mathcal{P}_p(\mathbb{R}^d)$ at least, although not exclusively. In particular, note that the L^∞ Wasserstein distances take finite values over distributions P and Q that only differ on a space translation independently on their supports being compact or not. For this reason, along the paper we shall not restrict to compactly supported distributions, but anyway in all our computations the involved L^∞ Wasserstein distances will take finite values as it will become clear later in the proofs.

2. Proof of the one-step contraction property

For the reader convenience, we provide first the main ingredients behind the proof of the fundamental one-step contraction property in Theorem 1.2(ii). Here, we shall assume that Theorem 1.2(i) holds true, *i.e.*, there exists a β-log-concave solution $(\boldsymbol{\lambda}, \boldsymbol{F})$ to (1.5) with β given by (1.7) (recall the precise notion of strong log-concavity in Definition 1.1). We remark that its use will be crucial in our following argument, but its proof is not apparent with regards to classical approaches based on the application of the Krein-Rutman theorem. For this reason, a major part of this paper is devoted to rigorously address this question, which will be introduced in full detail in Section 5 of this paper.

2.1. Sharp log-concavity parameter.

2.1. **Sharp log-concavity parameter.** First, we elaborate on the precise value of β given in (1.7). Specifically, we prove that it amounts to the sharpest possible log-concavity parameter of a generic solution (λ, F) to (1.5). To this end, it is worthwhile to note that the nonlinear operator $\mathcal{T}$ in (1.2) can be restated as the composition of a multiplicative operator and a double convolution operator, namely,

$$\mathcal{T}[F] = \frac{e^{-m}}{\|F\|_{L^1}}(G * \bar{F} * \bar{F}), \tag{2.1}$$

for every $F \in L^1_+(\mathbb{R}) \setminus \{0\}$, where we define $\bar{F}(x) := 2F(2x)$ for $x \in \mathbb{R}$. The starting point is to realize that strong log-concavity is stable under convolutions. This is a classical corollary of the celebrated Prékopa-Leindler inequality, which reads as follows (see [41, Proposition 7.1] for futher details).

Lemma 2.1 (Stability of log-concavity under convolutions). *Assume that $F_1, F_2 \in L^1_+(\mathbb{R})$ verify that F_i are γ_i-log-concave for some $\gamma_1, \gamma_2 > 0$. Then $F_1 * F_2$ is also γ-log-concave for $\gamma > 0$ given by*

$$\frac{1}{\gamma} = \frac{1}{\gamma_1} + \frac{1}{\gamma_2}.$$

Let us remark that the above result could be applied to any couple of Gaussian distributions F_1 and F_2 with respective variances σ_1^2 and σ_2^2 since they are in particular γ_i-log-concave with parameters $\gamma_i = \frac{1}{\sigma_i^2}$ for $i = 1, 2$. In doing so one finds that the above result is consistent with the classical fact that the convolution $F_1 * F_2$ of two Gaussian distributions is again Gaussian with variance $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

In addition, note that the mortality function m has been chosen α-convex by the hypothesis (H1) in Theorem 1.2, and then e^{-m} is α-log-concave. Since strong log-concavity is also preserved under multiplication, and $\bar{F}$ is 4γ-log-concave whenever F is γ-log-concave, then we obtain that log-concavity must also be preserved under the full operator $\mathcal{T}$.

Lemma 2.2 (Stability of log-concavity under $\mathcal{T}$). *Assume that $F \in L^1_+(\mathbb{R}) \setminus \{0\}$ is γ-log-concave for some $\gamma > 0$. Then, $\mathcal{T}[F]$ is also δ-log-concave for $\delta > 0$ given by*

$$\delta = \alpha + \frac{2\gamma}{1 + 2\gamma}.$$

Thereby, log-concavity is preserved by the dynamics in (1.1), and we also obtain that the sharpest log-concavity coefficient of the eigenfunction $\boldsymbol{F}$ must be the one given in (1.7).

Lemma 2.3 (Propagation of log-concavity).

(i) Assume that $F_0 \in L^1_+(\mathbb{R}) \setminus \{0\}$ is β_0-log-concave for some $\beta_0 > 0$. Then, the solution $\{F_n\}_{n\in\mathbb{N}}$ to the evolution problem (1.1) *verifies that F_n is β_n-log-concave for $\beta_n > 0$ verifying the recurrence*

$$\beta_n = \alpha + \frac{2\,\beta_{n-1}}{1 + 2\,\beta_{n-1}}, \quad n \in \mathbb{N}. \tag{2.2}$$

(ii) Assume that (λ, F) is any solution to the nonlinear eigenproblem (1.5) *and that F is strongly log-concave. Then, F is β-log-concave with β given by* (1.7), *that is,*

$$\beta = \alpha + \frac{2\,\beta}{1 + 2\,\beta}.$$

Proof. Since (i) is clear by Lemma 2.2, we just prove (ii). Recall that for any solution (λ, F) of (1.5) with γ-log-concave F, we can build $F_n(x) = \lambda^n F(x)$, which solves the evolution problem (1.1). Therefore, the above applied to $\{F_n\}_{n\in\mathbb{N}}$ shows that F is β_n log-concave for any $n \in \mathbb{N}$ with $\{\beta\}_{n\in\mathbb{N}}$ verifying the recurrence (2.2) above and $\beta_0 = \gamma$. Since $\beta_n \to \beta$, then F is also β-log-concave. ☐

2.2. **The renormalized problem.** We introduce a renormalized version of the evolution problem (1.1). Specifically, for any solution $\{F_n\}_{n\in\mathbb{N}}$ to (1.1) we renormalize by the strongly log-concave quasi-equilibrium $\boldsymbol{\lambda}^n\,\boldsymbol{F}$ granted in Theorem 1.2(i). Namely, we set

$$u_n(x) := \frac{F_n(x)}{\boldsymbol{\lambda}^n\,\boldsymbol{F}(x)},\quad n\in\mathbb{N},\, x\in\mathbb{R}. \tag{2.3}$$

By inspection, we obtain that $\{u_n\}_{n\in\mathbb{N}}$ must solve the evolution problem

$$u_n(x) = \frac{1}{\|u_{n-1}\,\boldsymbol{F}\|_{L^1}}\iint_{\mathbb{R}^2}\boldsymbol{P}(x;x_1,x_2)\,u_{n-1}(x_1)\,u_{n-1}(x_2)\,dx_1\,dx_2, \tag{2.4}$$

for any $x\in\mathbb{R}$, where $\boldsymbol{P}(x;x_1,x_2)$ is the *one-step transition probability* of transitioning from the parental traits (x_1,x_2) to the descendant trait x. More, specifically, $\boldsymbol{P}(x;\cdot)\in L^1_+(\mathbb{R}^2)\cap\mathcal{P}(\mathbb{R}^2)$ is a probability density on two variables (x_1,x_2) depending on the parameter $x\in\mathbb{R}$ which takes the form (recall the notation $\boldsymbol{F}=e^{-\boldsymbol{V}}$),

$$\begin{aligned}
\boldsymbol{P}(x;x_1,x_2) &:= \frac{1}{\boldsymbol{Z}(x)}e^{-\boldsymbol{W}(x;x_1,x_2)},\quad x\in\mathbb{R},\quad (x_1,x_2)\in\mathbb{R}^2,\\
\boldsymbol{W}(x;x_1,x_2) &:= \frac{1}{2}\left|x-\frac{x_1+x_2}{2}\right|^2+\boldsymbol{V}(x_1)+\boldsymbol{V}(x_2),\\
\boldsymbol{Z}(x) &:= \iint_{\mathbb{R}^2}e^{-\boldsymbol{W}(x;x_1,x_2)}\,dx_1\,dx_2.
\end{aligned} \tag{2.5}$$

Inspired by our method in [16], we plan to study the relaxation to zero of $\left\|\frac{d}{dx}(\log u_n)\right\|_{L^\infty}$ as n grows. Nevertheless, contrarily to the aforementioned paper, we do not need to accumulate a large enough amount of generations in order to observe some ergodic behavior, but we rather find a precise contraction of such a quantity after a single step.

2.3. **A nonlinear Kantorovich-type duality.** Our new approach exploits a nice nonlinear version of a Kantorovich-type duality which relates the L^∞ transport distance to the Lipschitz norm of the log of test functions. This nonlinear extension is reminiscent of the usual Kantorovich duality theorem, which relates the L^1 transport distance to the Lipschitz norm of test functions, see [1, Theorem 6.1.1]. More specifically, we remark that the usual Kantorovich duality is fundamental in the linear setting to establish a general equivalence between the contraction of a forward semigroup under the Lipschitz norm, and the contraction of its backward (or dual) semigroup under the L^1 transport distance. We refer to [30] for further extensions, yet in a linear setting. In our case, our nonlinear relation provides a method to derive contraction of a forward semigroup under the Lipchitz norms of the log of tests functions, once we know that there is contraction of the backward semigroup under a suitable L^∞ transport distance. Interestingly, our nonlinear relation does not only apply to the linear setting, but also to our nonlinear setting. To the best of our knowledge, this relation appears to be new. Moreover, it does not represent an isolated example but there is a full family of related inequalities interpolating between the (classical) L^1 result and the (seemingly new) L^∞ result, and which further adapt to L^p transport distances, see Appendix A.

Lemma 2.4 (L^∞-type Kantorovich duality)**.** *Consider the one-step transition from u_0 to u_1 in* (2.4)*, where it is assumed that $u_0\in C^1(\mathbb{R})$ with $u_0>0$ and $\frac{d}{dx}(\log u_0)\in L^\infty(\mathbb{R})$. Then, we have*

$$|\log u_1(x)-\log u_1(\tilde{x})|\le\left\|\frac{d}{dx}(\log u_0)\right\|_{L^\infty}W_{\infty,1}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot)), \tag{2.6}$$

for any $x,\tilde{x}\in\mathbb{R}$. Here, the metric $W_{\infty,1}$ represents the L^∞ Wasserstein distance associated with the ℓ_1 norm, cf. (1.24)*.*

Proof. Set $x,\tilde{x}\in\mathbb{R}$ and assume that $W_{\infty,1}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot))<\infty$ (otherwise the inequality is obvious). Indeed, this will always be the case as we prove later in Section 3. Then, consider any $\gamma\in\Gamma(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot))$

minimizing the $W_{\infty,1}$ transport distance (1.24) and note that

$$\begin{aligned}
u_1(x) &= \frac{1}{\|u_0\,\boldsymbol{F}\,\|_{L^1}}\iint_{\mathbb{R}^2} u_0(x_1)\,u_0(x_2)\,\gamma(dx_1,dx_2,d\tilde{x}_1,d\tilde{x}_2)\\
&= \frac{1}{\|u_0\,\boldsymbol{F}\,\|_{L^1}}\iint_{\mathbb{R}^2} \exp\Big(\log u_0(x_1)-\log u_0(\tilde{x}_1)+\log u_0(x_2)-\log u_0(\tilde{x}_2)\Big)\\
&\qquad\qquad\times u_0(\tilde{x}_1)\,u_0(\tilde{x}_2)\,\gamma(dx_1,dx_2,d\tilde{x}_1,d\tilde{x}_2)\\
&\leq \frac{1}{\|u_0\,\boldsymbol{F}\,\|_{L^1}}\iint_{\mathbb{R}^2} \exp\left(\left\|\frac{d}{dx}(\log u_0)\right\|_{L^\infty}\|(x_1,x_2)-(\tilde{x}_1,\tilde{x}_2)\|_1\right)\\
&\qquad\qquad\times\, u_0(\tilde{x}_1)\,u_0(\tilde{x}_2)\,\gamma(dx_1,dx_2,d\tilde{x}_1,d\tilde{x}_2)\\
&\leq \exp\left(\left\|\frac{d}{dx}(\log u_0)\right\|_{L^\infty} W_{\infty,1}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x},\cdot))\right)u_1(\tilde{x}),
\end{aligned}$$

where in the next-to-last line we have used the mean value theorem and in the last one we have exploited the fact that γ is minimizer. Then, taking logarithm on each side of the above inequality ends the proof. ☐

Remark 2.5 (The choice of ℓ_1 norm)**.** *We note that Lemma 2.4 is a particular instance of Proposition A.1 in Appendix A which can be recovered by setting $d_1 = 1$, $d_2 = 2$, $q = 1$ and*

$$u(x_1,x_2) := u_0(x_1)\,u_0(x_2),\quad (x_1,x_2)\in\mathbb{R}^2.$$

However, the special choice $q = 1$ (that is ℓ_1 norms) is apparently less clear at this stage since in fact choosing any other $1\leq q\leq\infty$ would be possible in Proposition A.1 and it would yield more generally

$$|\log u_1(x)-\log u_1(\tilde{x})|\leq 2^{1/q'}\left\|\frac{d}{dx}(\log u_0)\right\|_{L^\infty} W_{\infty,q}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot)), \tag{2.7}$$

for every $x,\tilde{x}\in\mathbb{R}$. Here, the metric $W_{\infty,q}$ represents the L^∞ Wasserstein distance associated with the ℓ_q norm, cf. (1.24)*. By the natural relation between ℓ_1 and ℓ_q vector norms, we infer that the above estimate* (2.6) *is sharper than* (2.7)*, namely*

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot))\leq 2^{1/q'}W_{\infty,q}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot)).$$

Therefore, it is clear that whenever $q>1$ the additional factor $2^{1/q'}$ makes the one-step contraction factor in next section non-optimal as compared to the explicit one-step contraction for quadratic selection $m(x)=\frac{\alpha}{2}|x|^2$, as illustrated in Remark 2.7 and more detailed later in Remark 3.1.

2.4. **Contraction of the one-step transition probability.** The last step of our argument requires showing that the mapping $x\in\mathbb{R}\mapsto\boldsymbol{P}(x;\cdot)\in L^1_+(\mathbb{R}^2)\cap\mathcal{P}(\mathbb{R}^2)$ is a contraction when the space $\mathcal{P}(\mathbb{R}^2)$ is endowed with the $W_{\infty,1}$ Wasserstein distance in (1.24). Specifically, in the following result we quantify the exact Lipschitz constant, which will account for the precise contraction factor in Theorem 1.2(ii).

Lemma 2.6 ($W_{\infty,1}$-contraction)**.** *Consider the one-step transition probability $\boldsymbol{P}=\boldsymbol{P}(x;x_1,x_2)$ defined in* (2.5) *in terms of the potential $\boldsymbol{V}$ of the β-log-concave quasi-equilibrium $\boldsymbol{F}=e^{-\boldsymbol{V}}$ in Theorem 1.2(i). Then, the following inequality holds true*

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot))\leq\frac{2}{1+2\,\beta}|x-\tilde{x}|,$$

for every $x,\tilde{x}\in\mathbb{R}$.

A similar contraction property, with respect to W_1 distances instead of W_∞, appeared previously in [35, 36] leading to the definition of coarse Ricci curvature of a Markov kernel $\boldsymbol{P}(x;\cdot)$:

$$\kappa(x,\tilde{x}) = 1-\frac{W_1(\boldsymbol{P}(x;\cdot),\boldsymbol{P}(\tilde{x};\cdot))}{|x-\tilde{x}|},\quad x,\,\tilde{x}\in\mathbb{R}.$$

Specifically, the above references proved that a positive lower bound on the coarse Ricci curvature amounts to the aforementioned contraction of the forward semigroup under the Lipschitz norm (or equivalently, the contraction of the backward semigroup under the L^1 transport distance [30]). For heat kernels in a linear setting, this hypothesis on the coarse Ricci curvature is compatible with the Bakry-Emery convexity condition and it proved equivalent to the contraction of the backward semigroup in all W_p transport distances [44], including W_∞. However, the decay of the L^∞ relative Fisher information has not been addressed in those works, and a nonlinear adaptation of them does not seem straightforward.

Before entering into the details of the proof of the Lemma 2.6, let us note that putting Lemmas 2.4 and 2.6 together automatically implies the following one-step contraction estimate

$$\left\| \frac{d}{dx}(\log u_1) \right\|_{L^\infty} \leq \frac{2}{1+2\beta} \left\| \frac{d}{dx}(\log u_0) \right\|_{L^\infty}, \tag{2.8}$$

which can be iterated and propagated into (1.9) in Theorem 1.2(ii) (at generation n), thus concluding this section. Nevertheless, we remark that Lemma 2.6 is far from straightforward as one typically cannot even ensure that the above $W_{\infty,1}$ distance must be finite because the probability densities $\boldsymbol{P}(x;\cdot)$ and $\boldsymbol{P}(\tilde{x};\cdot)$ are supported on the full plane $\mathbb{R}^2$.

Remark 2.7 (Quadratic selection)**.** *In the case of quadratic selection $m(x) = \frac{\alpha}{2}|x|^2$ studied in* [16]*, we recall from Remark 1.3 that the unique eigenfunction of* (1.5) *is the Gaussian $\boldsymbol{F} = G_{0,\sigma^2}$ with variance $\sigma^2 = \beta^{-1}$. Therefore, one easily obtains from* (2.5) *that*

$$\boldsymbol{P}(x, x_1, x_2) \propto \exp\left(-\frac{1}{2}\left| x - \frac{x_1+x_2}{2} \right|^2 - \frac{\beta}{2}|x_1|^2 - \frac{\beta}{2}|x_2|^2 \right).$$

Completing squares with respect to the variables (x_1, x_2) we readily find that $\boldsymbol{P}(x;\cdot) = G_{\mu_x,\Sigma}$ is the density of a bivariate normal distribution with mean and covariance matrix determined by

$$\mu_x := \frac{1}{1+2\beta}(x,x), \qquad \Sigma^{-1} := \begin{pmatrix} \frac{1}{4}+\beta & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4}+\beta \end{pmatrix}.$$

Since Σ is independent of x, then any couple of Gaussians $\boldsymbol{P}(x;\cdot)$ and $\boldsymbol{P}(\tilde{x};\cdot)$ must agree up to a translation in the direction joining their means. Hence, the transport cost reduces to moving the center μ_x of $\boldsymbol{P}(x;\cdot)$ to the center $\mu_{\tilde{x}}$ of $\boldsymbol{P}(\tilde{x};\cdot)$, which yields Lemma 2.6 (with identity indeed):

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) = \|\mu_x - \mu_{\tilde{x}}\|_1 = \frac{2}{1+2\beta}|x-\tilde{x}|.$$

The goal of this section is to prove Lemma 2.6. To alleviate the notation, along this section we name $z := (x_1, x_2) \in \mathbb{R}^2$, we fix $x, \tilde{x} \in \mathbb{R}$ with $x \neq \tilde{x}$ and then we simplify the notation on the one-step transition probability in (2.5) by setting $\boldsymbol{p}(z) := \boldsymbol{P}(x; x_1, x_2)$ and $\tilde{\boldsymbol{p}}(z) := \boldsymbol{P}(\tilde{x}; x_1, x_2)$, that is,

$$\boldsymbol{p}(z) = \frac{1}{\boldsymbol{Z}} e^{-\boldsymbol{W}(z)}, \quad \tilde{\boldsymbol{p}}(z) = \frac{1}{\tilde{\boldsymbol{Z}}} e^{-\tilde{\boldsymbol{W}}(z)}, \tag{2.9}$$

where the potentials $\boldsymbol{W}$ and $\tilde{\boldsymbol{W}}$, and the normalizing constants $\boldsymbol{Z}$ and $\tilde{\boldsymbol{Z}}$ are then given by

$$\begin{aligned} \boldsymbol{W}(z) &:= \boldsymbol{W}(x; x_1, x_2) = \frac{1}{2}\left| x - \frac{x_1+x_2}{2} \right|^2 + \boldsymbol{V}(x_1) + \boldsymbol{V}(x_2), \\ \tilde{\boldsymbol{W}}(z) &:= \boldsymbol{W}(\tilde{x}; x_1, x_2) = \frac{1}{2}\left| \tilde{x} - \frac{x_1+x_2}{2} \right|^2 + \boldsymbol{V}(x_1) + \boldsymbol{V}(x_2), \\ \boldsymbol{Z} &:= \boldsymbol{Z}(x) = \iint_{\mathbb{R}^2} e^{-\boldsymbol{W}(z)}\,dz, \quad \tilde{\boldsymbol{Z}} := \boldsymbol{Z}(\tilde{x}) = \iint_{\mathbb{R}^2} e^{-\tilde{\boldsymbol{W}}(z)}\,dz. \end{aligned} \tag{2.10}$$

For any transport map $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ with $T_\# \boldsymbol{p} = \tilde{\boldsymbol{p}}$, note that a possible strategy in order to estimate the $W_{\infty,1}$ distance is to compute an L^∞ bound for the ℓ_1 associated displacement, namely,

$$W_{\infty,1}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \leq \| \|T - I\|_1 \|_{L^\infty}. \tag{2.11}$$

Whilst the choice of T is somehow arbitrary at this point, a comfortable one is usually the Brenier map $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ from the density $\boldsymbol{p}$ to the density $\tilde{\boldsymbol{p}}$, which is characterized as the unique transport map verifying $T_\# \boldsymbol{p} = \tilde{\boldsymbol{p}}$ and solving the Monge problem [9]

$$\iint_{\mathbb{R}^2} \|T(z) - z\|_2^2 \, \boldsymbol{p}(z)\,dz = W_{2,2}^2(\boldsymbol{p}, \tilde{\boldsymbol{p}}),$$

where $W_{2,2}$ is the L^2 Wasserstein distance associated with the ℓ_2 norm of $\mathbb{R}^2$, *cf.* (1.24). As we anticipated in the **Methodological notes** in Section 1, in many cases this non-optimal argument leads to no loss of generality since the $W_{\infty,1}$ and the uniform bound of the ℓ_1 displacement of the Brenier map have the same order. This was further depicted in the example of the Gaussians from Remark 2.7, where the Brenier map is a translation, and therefore the transport cost is indeed identical to the displacement.

Our proof of Lemma 2.6 is based on the derivation of a novel L^∞ bound of the ℓ_1 displacement $\|T - I\|_1$ associated with the Brenier map T between the densities $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$. We derive those bounds by reformulating such a Brenier map as a solution to a Monge-Ampère equation and using a version of

Caffarelli's maximum principle along with the strong log-concavity of our densities. Indeed, by the strong log-concavity of $\boldsymbol{F}$ in Theorem 1.2(i) we have

$$-D^2_{(x_1,x_2)}\log \boldsymbol{p} = -D^2_{(x_1,x_2)}\log \tilde{\boldsymbol{p}} \geq \begin{pmatrix} \frac{1}{4}+\beta & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4}+\beta \end{pmatrix} \geq \beta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and then $\boldsymbol{p}, \tilde{\boldsymbol{p}}$ are β-log-concave. The aforementioned strategy recalls the one applied in *Caffarelli's contraction principle* [13, 14, 17, 18] to find Lipschitz bounds of the Brenier map between strongly log-concave probability densities. Yet, in order to obtain Lipschitz bounds on the map (*i.e.* bounds on the Hessian of the potential), it is necessary to differentiate twice the Monge-Ampère equation; here we only require bounds on the displacement, and we need to differentiate only once. This recalls more what was done in [22], where the goal was to obtain Lipschitz bounds on the logarithm of the solution of a JKO scheme or, equivalently, L^∞ bounds of the displacement associated with the Brenier map between two subsequent measures in the same JKO scheme. Among the important differences, [22] was not concerned with log-concave measures, but required one of the two to be obtained from the other via the JKO scheme. As another important difference, [22] was concerned with ℓ_2 displacement bounds, and the choice of the Euclidean ball played a special role. In our setting, in view of the definition (1.24) of $W_{\infty,1}$, the choice of ℓ_2 is not suitable and we focus on ℓ_1. For the ℓ_1 norm, we obtain new bounds on the Monge-Ampère equation, which are able to find the sharp contraction factor, and which cannot be recovered by interpolation from known ℓ_2 estimates, see Remark 3.1.

For the reader's convenience, we provide below a formal proof of Lemma 2.6 under the strong additional assumption that the maximal ℓ_1 displacement associated with the Brenier map is attained. Whilst true in particular situations (*cf.* Remark 2.7), unfortunately this hypothesis is not necessarily always true, and thus the rigorous derivation requires further work which we provide in detail in Section 3.

Formal proof of Lemma 2.6. It is well known that the Brenier map $T:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ from $\boldsymbol{p}$ to $\tilde{\boldsymbol{p}}$ takes the form $T=\nabla\phi$ for some convex function $\phi:\mathbb{R}^2\longrightarrow\mathbb{R}$. Since $\boldsymbol{p},\tilde{\boldsymbol{p}}>0$ and $\boldsymbol{p},\tilde{\boldsymbol{p}}\in C^\infty(\mathbb{R}^2)$, then the regularity results in [11] imply that $\phi\in C^\infty(\mathbb{R}^2)$. Moreover, the change of variable formula implies

$$\det(D^2\phi) = \frac{\boldsymbol{p}}{\tilde{\boldsymbol{p}}\circ\nabla\phi}, \quad z\in\mathbb{R}^2. \tag{2.12}$$

As usual we make the change of variables through the displacement potential

$$\psi(z) := \phi(z) - \frac{1}{2}\|z\|_2^2, \quad z\in\mathbb{R}^2. \tag{2.13}$$

In view of the relation (2.11), we note that the core of the proof then reduces to obtaining L^∞ bounds for the ℓ_1 norm of the displacement of the Brenier map, that is,

$$H(z) := \|T(z)-z\|_1 = \|\nabla\psi(z)\|_1 = |\partial_{x_1}\psi(z)| + |\partial_{x_2}\psi(z)|, \quad z\in\mathbb{R}^2. \tag{2.14}$$

We start by restating the Monge-Ampère equation (2.12) by taking its logarithm,

$$\log\det(D^2\psi(z)+I) = \tilde{\boldsymbol{W}}(\nabla\psi(z)+z) - \boldsymbol{W}(z) + \log\frac{\tilde{\boldsymbol{Z}}}{\boldsymbol{Z}}, \quad z\in\mathbb{R}^2. \tag{2.15}$$

Taking partial derivatives ∂_{x_k} in (2.15) we have

$$\operatorname{tr}\left((D^2\phi)^{-1}\partial_{x_k}D^2\psi\right) = \nabla\tilde{\boldsymbol{W}}(\nabla\psi+z)\cdot\partial_{x_k}\nabla\psi + (\nabla\tilde{\boldsymbol{W}}(\nabla\psi+z) - \nabla\,\boldsymbol{W})\cdot e_k, \quad z\in\mathbb{R}^2, \tag{2.16}$$

for $k=1,2$. Let us assume that H attains its maximum at some $z^*=(x_1^*,x_2^*)\in\mathbb{R}^2$ (for the general case where the maximum is not attained we refer to Section 3) and let us also define the auxiliary function

$$\tilde{H}(z) := \operatorname{sgn}(\partial_{x_1}\psi(z^*))\,\partial_{x_1}\psi(z) + \operatorname{sgn}(\partial_{x_2}\psi(z^*))\,\partial_{x_2}\psi(z), \quad z\in\mathbb{R}^2. \tag{2.17}$$

Then, $\tilde{H}$ must also attain its maximum at z^* and it agrees with the maximum of H. In particular, we have the necessary optimality conditions

$$\nabla\tilde{H}(z^*) = 0, \quad D^2\tilde{H}(z^*) \leq 0. \tag{2.18}$$

Now, we perform an appropriate convex combination of (2.16) depending on the signs of $\partial_{x_1}\psi(z^*)$ and $\partial_{x_2}\psi(z^*)$ in order to make the auxiliary function $\tilde{H}$ in (2.14) appear.

$\diamond$ Case 1: $\partial_{x_1}\psi(z^*)\geq 0$ and $\partial_{x_2}\psi(z^*)\geq 0$.
In this case we have $\tilde{H}:=\partial_{x_1}\psi+\partial_{x_2}\psi$. Evaluating (2.16) at z^* and summing over $k\in\{1,2\}$ we have

$$\begin{aligned}\operatorname{tr}((D^2\phi(z^*))^{-1}D^2\tilde{H}(z^*)) &= \nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)\cdot\nabla\tilde{H}(z^*) \\ &\quad + (\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*) - \nabla\,\boldsymbol{W}(z^*))\cdot(1,1).\end{aligned}$$

By the optimality conditions (2.18) and since $D^2\phi(z^*)^{-1}$ is positive definite, the term in the left hand side above is non-positive, and we obtain

$$(\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)-\nabla\tilde{\boldsymbol{W}}(z^*))\cdot(1,1)\leq\nabla(\boldsymbol{W}-\tilde{\boldsymbol{W}})(z^*)\cdot(1,1)=\tilde{x}-x.$$

By expanding the left hand side we obtain

$$\begin{aligned}&(\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)-\nabla\tilde{\boldsymbol{W}}(z^*))\cdot(1,1)\\&\quad=\frac{\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*)}{2}+\boldsymbol{V}'(\partial_{x_1}\psi(z^*)+x_1^*)-\boldsymbol{V}'(x_1^*)+\boldsymbol{V}'(\partial_{x_2}\psi(z^*)+x_2^*)-\boldsymbol{V}'(x_2^*)\\&\quad\geq\frac{\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*)}{2}+\beta(\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*))=\frac{1+2\,\beta}{2}\tilde{H}(z^*),\end{aligned}$$

where we have used that in this case $\partial_{x_1}\psi(z^*)\geq 0$ and $\partial_{x_2}\psi(z^*)\geq 0$, along with the β-convexity of $\boldsymbol{V}$. Therefore, we conclude that $\tilde{x}>x$ and

$$\|H\|_{L^\infty}=H(z^*)=\tilde{H}(z^*)\leq\frac{2}{1+2\,\beta}|x-\tilde{x}|.$$

⋄ CASE 2: $\partial_{x_1}\psi(z^*)<0$ and $\partial_{x_2}\psi(z^*)<0$.
This case follows the same argument as CASE 1. Indeed, note now that $\tilde{H}=-\partial_{x_1}\psi-\partial_{x_2}\psi$. Then, we sum over $k\in 1,2$, multiply by -1 on (2.16) and we obtain

$$\frac{1+2\,\beta}{2}\tilde{H}(z^*)\leq x-\tilde{x}.$$

Hence, in this case we obtain $x>\tilde{x}$ and we recover

$$\|H\|_{L^\infty}=H(z^*)=\tilde{H}(z^*)\leq\frac{2}{1+2\,\beta}|x-\tilde{x}|.$$

We show below that the other two cases (namely, $\partial_{x_1}\psi(z^*)\geq 0$ and $\partial_{x_2}\psi(z^*)<0$, or $\partial_{x_1}\psi(z^*)<0$ and $\partial_{x_2}\psi(z^*)\geq 0$) cannot happen.

⋄ CASE 3: $\partial_{x_1}\psi(z^*)\geq 0$ and $\partial_{x_2}\psi(z^*)<0$.
Our goal is to show that this case cannot take place. In this case, we have $\tilde{H}:=\partial_{x_1}\psi-\partial_{x_2}\psi$. Taking the difference of (2.16) with $k=1$ and $k=2$ we obtain

$$\begin{aligned}\operatorname{tr}((D^2\phi(z^*))^{-1}D^2\tilde{H}(z^*))&=\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)\cdot\nabla\tilde{H}(z^*)\\&\quad+(\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)-\nabla\,\boldsymbol{W}(z^*))\cdot(1,-1).\end{aligned}$$

Since z^* is a maximizer of $\tilde{H}$ we have

$$(\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)-\nabla\tilde{\boldsymbol{W}}(z^*))\cdot(1,-1)\leq\nabla(\boldsymbol{W}-\tilde{\boldsymbol{W}})(z^*)\cdot(1,-1)=0$$

The expansion on the left hand side is now radically different because the above factor $\frac{\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*)}{2}$ cancels and now we obtain

$$\begin{aligned}&(\nabla\tilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)-\nabla\tilde{\boldsymbol{W}}(z^*))\cdot(1,-1)\\&\quad=\boldsymbol{V}'(\partial_{x_1}\psi(z^*)+x_1^*)-\boldsymbol{V}'(x_1^*)-\boldsymbol{V}'(\partial_{x_2}\psi(z^*)+x_2^*)+\boldsymbol{V}'(x_2^*)\\&\quad\geq\beta(\partial_{x_1}\psi(z^*)-\partial_{x_2}\psi(z^*))=\beta\,\tilde{H}(z^*),\end{aligned}$$

which implies $\|H\|_{L^\infty}=H(z^*)=\tilde{H}(z^*)=0$. This is clearly impossible since otherwise $T(z)=z$ for all $z\in\mathbb{R}^2$, that is, $x=\tilde{x}$.

⋄ CASE 4: $\partial_{x_1}\psi(z^*)<0$ and $\partial_{x_2}\psi(z^*)\geq 0$.
This case cannot happen either thanks to the same argument as in CASE 3 with $\tilde{H}$ replaced by $\tilde{H}=-\partial_{x_1}\psi+\partial_{x_2}\psi$. Then, we omit the proof. □

2.5. **Proof of the one-step contraction property.** With all the above machinery in hand, we are finally in position to prove the one-step contraction property (1.9) in Theorem 1.2.

Proof of Theorem 1.2(ii). Combining Lemmas 2.4 and 2.6 applied to the solution (2.3) of (2.4) we obtain

$$\left\|\frac{d}{dx}\left(\log\frac{F_n}{\boldsymbol{F}}\right)\right\|_{L^\infty}\leq\frac{2}{1+2\,\beta}\left\|\frac{d}{dx}\left(\log\frac{F_{n-1}}{\boldsymbol{F}}\right)\right\|_{L^\infty},$$

for every $n\in\mathbb{N}$, and this amounts to (1.9). □

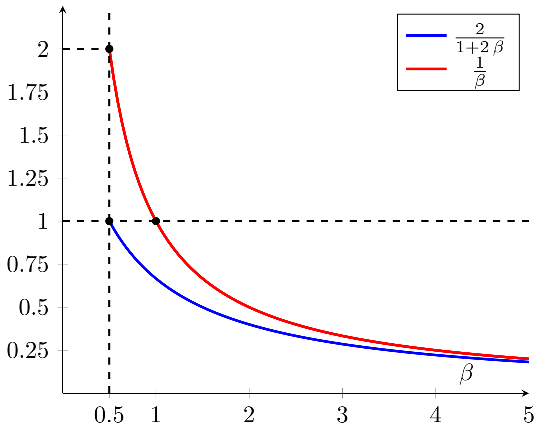


FIGURE 3. Comparison of the theoretical contraction factor $\frac{1}{1+2\beta}$ in Lemma 2.6, and the contraction factor $\frac{1}{\beta}$ obtained by estimating the ℓ_1 norm with the ℓ_2 norm in $\mathbb{R}^2$.

3. Main contractivity lemma

In this section, we provide a rigorous proof of Lemma 2.6, where the *a priori* assumption that the maximal displacement associated with the Brenier map must be attained is no longer required. To do so, we shall argue by deriving a local version of the Lemma valid for more general strongly log-concave densities f and g compactly supported on an appropriate domain and bounded away from zero on it. More specifically, we propose to adapt the contribution of the maximum principle to the formal argument above (Section 2.4) to compact domains. However, since the maximum may be attained at the boundary, the boundary information is crucial in order to infer information from the non-linear elliptic PDE (2.12) and therefore the choice of the domain cannot be made arbitrarily.

We refer to Appendix C for a bound on the maximum of $\|T-I\|_2$ (in ℓ_2 norm) for the Brenier map $T:\bar{B}_R\longrightarrow\bar{B}_R$ between two generic strongly log-concave probability densities $f=e^{-W}$ and $g=e^{-\tilde{W}}$, supported and strictly positive on an Euclidean ball $\bar{B}_R$. Specifically, we obtain

$$W_{\infty,2}(f,g)\le \|\|T-I\|_2\|_{L^\infty(\bar{B}_R)}\le\frac{1}{\gamma}\|\|\nabla(W-\tilde{W})\|_2\|_{L^\infty(\bar{B}_R)},\tag{3.1}$$

where $\gamma>0$ is the log-concavity parameter of f and g.

Remark 3.1 (Inaccuracy of controlling ℓ_1 by ℓ_2 norms)**.** *We may be tempted to apply this ℓ_2 estimate to our setting by setting f and g as truncations of $\boldsymbol{p}\propto e^{-\boldsymbol{W}}$ and $\tilde{\boldsymbol{p}}\propto e^{-\tilde{\boldsymbol{W}}}$ (see* (2.9)-(2.10)*) to ℓ_2 balls and using the Cauchy-Schwarz inequality to get ℓ_1 estimates. Specifically, consider an increasing sequence of balls B_R and set f and g in* (3.1) *to be the truncation of $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ on such balls. First, recall that*

$$D^2\,\boldsymbol{W}(x_1,x_2)=D^2\tilde{\boldsymbol{W}}(x_1,x_2)=\begin{pmatrix}\frac14+\boldsymbol{V}''(x_1) & \frac14\\ \frac14 & \frac14+\boldsymbol{V}''(x_2)\end{pmatrix}\ge\begin{pmatrix}\beta & 0\\ 0 & \beta\end{pmatrix},$$

because $\boldsymbol{V}''\ge 0$, and therefore we can set $\gamma=\beta$ in (3.1)*. Also note that*

$$\nabla(\boldsymbol{W}-\tilde{\boldsymbol{W}})(x_1,x_2)=\frac12(\tilde{x}-x,\tilde{x}-x).$$

Altogether implies the ℓ_2 estimate

$$W_{\infty,2}(f,g)\le\|\|T-I\|_2\|_{L^\infty(\bar{B}_R)}\le\frac{1}{\beta}\|\,\|\nabla(\boldsymbol{W}-\tilde{\boldsymbol{W}})\|_2\|_{L^\infty}=\frac{1}{\beta}\left\|\frac12(\tilde{x}-x,\tilde{x}-x)\right\|_2=\frac{1}{\sqrt{2}\,\beta}|x-\tilde{x}|,$$

and by the Cauchy-Schwarz inequality we also have the ℓ_1 estimate

$$W_{\infty,1}(f,g)\le\sqrt{2}\,W_{\infty,2}(f,g)\le\frac{1}{\beta}|x-\tilde{x}|.$$

In particular, we note that such an estimate only provides contraction as long as $\beta>1$ and, in addition, the contraction factor is worse than the one claimed in Lemma 2.6 as depicted in Figure 3.

We refer to [29] *for a nice and fruitful anisotropic version of 3.1 which enables to obtain directly the claimed contraction factor.*

Thus, we need to improve our proof and avoid using the ℓ_2 norm. This was done, formally, in the previous section, but we need a rigorous proof which also takes care of the boundary. Let us focus on the observation made in [22, Lemma 3.1] that, for generic f and g smooth on a ℓ_2 ball and bounded away from zero on it, the maximal ℓ_2 displacement of the Brenier map must be attained at some interior point in the ball. Apparently, the use of ℓ_2 norms to quantify the size of the displacement proved extremely well suited in order to control the boundary information on ℓ_2 balls. Interestingly, in the sequel we show that in order to find precise information about the maximizers for the ℓ_1 displacement, we need densities f and g to be supported over ℓ_∞ balls $\bar{B}_R$ (*cf.* (1.22)). This is the content of the following

Lemma 3.2 (Maximizers in the ℓ_1 setting)**.** *Consider two densities $f, g \in L^1_+(\mathbb{R}^2) \cap \mathcal{P}(\mathbb{R}^2)$, assume that,*

$$\{z \in \mathbb{R}^2 : f(z) > 0\} = \{z \in \mathbb{R}^2 : g(z) > 0\} = \bar{Q}_R,$$

where Q_R is the ℓ_∞ ball (cf. (1.22)), and suppose that $f, g \in C^{1,\delta}(\bar{Q}_R)$ for some $\delta > 0$. Let $T = \nabla\phi : \bar{Q}_R \longrightarrow \bar{Q}_R$ be the Brenier map from f to g, define the displacement potential $\psi(z) := \phi(z) - \frac{1}{2}\|z\|_2^2$ and the displacement function quantified in ℓ_1 norm

$$H(z) := \|T(z) - z\|_1 = |\partial_{x_1}\psi(z)| + |\partial_{x_2}\psi(z)|, \quad z \in \bar{Q}_R, \tag{3.2}$$

Then, $T \in C^{2,\delta}(\bar{Q}_R)$ and we have the optimality conditions

$$\nabla\tilde{H}(z^*) = 0, \quad D^2\tilde{H}(z^*) \leq 0, \tag{3.3}$$

for any maximizer $z^* = (z_1^*, z_2^*) \in \bar{Q}_R$ of H, where $\tilde{H}$ is the auxiliary function

$$\tilde{H}(z) := \operatorname{sgn}(\partial_{x_1}\psi(z^*))\, \partial_{x_1}\psi(z) + \operatorname{sgn}(\partial_{x_2}\psi(z^*))\, \partial_{x_2}\psi(z), \quad z \in \bar{Q}_R. \tag{3.4}$$

In contrast with the standard regularity theory for optimal transport, Q_R is not uniformly convex. Then, the regularity theory of the Monge-Ampère equation is not directly applicable in full generality. Specifically, since $f, g \in C^{1,\delta}(\bar{Q}_R)$ are bounded away from zero on $\bar{Q}_R$, then $T \in C^{0,\delta}(\bar{Q}_R)$ by [10]. However, the lack of uniform convexity may prevent the full elliptic regularity [12], which claims that T is a diffeomorphism of class $C^{2,\delta}(\bar{Q}_R)$. Fortunately, we can proceed as in [28, Theorem 3.3] which, thanks to a clever symmetrization argument around each corner of Q_R and the classical interior regularity in [11], shows that T is indeed a diffeomorphism of class $C^{2,\delta}(\bar{Q}_R)$. Moreover, it fixes the corners and sends each segment of the boundary to itself. This guarantees in particular that $\tilde{H} \in C^2(\bar{Q}_R)$ and the optimality conditions above make sense, as shown below.

Proof of Lemma 3.2. We remark that $z^* \in \bar{Q}_R$ must also be a maximizer of $\tilde{H}$ since we have

$$\tilde{H}(z) \leq H(z) \leq H(z^*) = \tilde{H}(z^*),$$

for every $z^* \in \bar{Q}_R$ by the definition of H and $\tilde{H}$ in (3.2) and (3.4). Since the maximizer z^* may lie in principle in all $\bar{Q}_R$, two possible options arise, either $z^* \in Q_R$ or $z^* \in \partial Q_R$. In the first case, the usual optimality conditions at interior points yield (3.3). In the second case, namely $z^* \in \partial Q_R$, note that the result is trivial if z^* is one of the four corners since those are fixed points of T and therefore $\tilde{H} \equiv 0$. Hence, here on we will assume that $z^* \in \partial Q_R$ is not at a corner, but it lies in the interior of some of the four segments. Note that at those points we only have to prove that $\nabla\tilde{H}(z^*) = 0$. In fact, we remark that those z^* can be approached by interior points from any direction, and then the above readily implies the second order optimality condition $D^2\tilde{H}(z^*) \leq 0$. To show that $\nabla\tilde{H}(z^*) = 0$, note that the boundary ∂Q_R contains four segments:

$$\begin{aligned}
S_1^+ &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = R,\, x_2 \in [-R, R]\},\\
S_1^- &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = -R,\, x_2 \in [-R, R]\},\\
S_2^+ &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [-R, R],\, x_2 = R\},\\
S_2^- &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [-R, R],\, x_2 = -R\}.
\end{aligned}$$

Since $T(\partial Q_R) = \partial Q_R$ and each segment is mapped to itself, then we have the following information

$$\partial_{x_1}\psi(z) = 0, \qquad \text{if } z \in S_1^+ \cup S_1^-, \tag{3.5}$$

$$\partial_{x_2}\psi(z) = 0, \qquad \text{if } z \in S_2^+ \cup S_2^-. \tag{3.6}$$

By differentiation it is clear that we also have

$$\partial_{x_1 x_2}\psi(z) = 0, \qquad \text{if } z \in \partial Q_R. \tag{3.7}$$

Now, we argue according to the four possible segments of ∂Q_R that z^* may belong to.

$\diamond$ CASE 1: $z^* \in S_1^+ \cup S_1^-$.
In this case, by (3.5) we have $\partial_{x_1}\psi(z^*) = 0$ and therefore we have

$$\tilde{H}(z) = \operatorname{sgn}(\partial_{x_2}\psi(z^*))\,\partial_{x_2}\psi(z), \quad z \in \bar{Q}_R.$$

Since z^* is a maximizer of $\tilde{H}$, then there exist $\lambda \in \mathbb{R}$ (indeed $\lambda \geq 0$ if $z^* \in S_1^+$ and $\lambda \leq 0$ if $z^* \in S_1^-$) such that its gradient at z^* equals the multiple $\lambda(1,0)$ of the outer normal vector, that is,

$$\nabla\tilde{H}(z^*) = \operatorname{sgn}(\partial_{x_2}\psi(z^*)) \begin{pmatrix} \partial_{x_1x_2}\psi(z^*) \\ \partial_{x_2x_2}\psi(z^*) \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}.$$

This implies that the second component of the gradient must vanish, but the first one also vanishes by the condition (3.7) on the crossed derivative. Then, we have $\nabla\tilde{H}(z^*) = 0$.

$\diamond$ CASE 2: $z^* \in S_2^+ \cup S_2^-$.
In this case, by (3.6) we have $\partial_{x_2}\psi(z^*) = 0$ and therefore we have

$$\tilde{H}(z) = \operatorname{sgn}(\partial_{x_1}\psi(z^*))\,\partial_{x_1}\psi(z), \quad z \in \bar{Q}_R.$$

Since z^* is a maximizer of $\tilde{H}$, then there exist $\lambda \in \mathbb{R}$ (indeed $\lambda \geq 0$ if $z^* \in S_2^+$ and $\lambda \leq 0$ if $z^* \in S_2^-$) such that its gradient at z^* equals the multiple $\lambda(0,1)$ of the outer normal vector, that is,

$$\nabla\tilde{H}(z^*) = \operatorname{sgn}(\partial_{x_1}\psi(z^*)) \begin{pmatrix} \partial_{x_1x_1}\psi(z^*) \\ \partial_{x_1x_2}\psi(z^*) \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda \end{pmatrix}.$$

This implies that the first component of the gradient must vanish, but the second one also vanishes by the condition (3.7) on the crossed derivative. Then, we have $\nabla\tilde{H}(z^*) = 0$. □

We remark that the unique formal point of the sketch of the proof of Lemma 2.6 in Section 2 which could break down is the fact that for the global densities $f = \boldsymbol{p}$ and $g = \tilde{\boldsymbol{p}}$ in (2.9)-(2.10) the ℓ_1 displacement of their Brenier map does not attain its maximum necessarily. In particular, we may be deprived from the optimality condition (2.18), which was crucially used throughout the maximum-type principle sketched in Section 2. However, Lemma 3.2 does guarantee that the maximum must be attained and the optimality conditions (3.3) must hold in particular when f and g are set to be the truncation of the densities $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ on ℓ_∞ balls. In fact, the result does not exploit the special potential $\boldsymbol{V}$ in the definition (2.9)-(2.10) of $\boldsymbol{p}, \tilde{\boldsymbol{p}}$, which corresponds to the potential of the eigenfunction $\boldsymbol{F} = e^{-\boldsymbol{V}}$ in Theorem 1.2(i), but it can actually be replaced by any strongly convex function supported on $\bar{Q}_R$. Since we shall use this more general version later in Section 4, we state in full generality below.

Lemma 3.3 (Maximum principle on ℓ_∞ balls)**.** *For any γ-convex potential $V \in C^{1,\delta}_{\mathrm{loc}}(\mathbb{R})$ with $\gamma > 0$, any $x, \tilde{x} \in \mathbb{R}$ with $x \neq \tilde{x}$, and any $R > 0$ we define $f, g \in L^1_+(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^2)$ given by*

$$f(z) = \frac{1}{Z}e^{-W(z)}, \quad g(z) = \frac{1}{\tilde{Z}}e^{-\tilde{W}(z)}, \quad z \in \mathbb{R}^2,$$

where the potentials W and $\tilde{W}$, and the normalizing constants Z and $\tilde{Z}$ are set as follows

$$\begin{aligned}
W(z) &:= \frac{1}{2}\left| x - \frac{x_1+x_2}{2}\right|^2 + V(x_1) + V(x_2) + \chi_{\bar{Q}_R}(z),\\
\tilde{W}(z) &:= \frac{1}{2}\left| \tilde{x} - \frac{x_1+x_2}{2}\right|^2 + V(x_1) + V(x_2) + \chi_{\bar{Q}_R}(z),\\
Z &:= \iint_{\mathbb{R}^2} e^{-W(z)}\,dz, \quad \tilde{Z} := \iint_{\mathbb{R}^2} e^{-\tilde{W}(z)}\,dz,
\end{aligned}$$

and $\chi_{\bar{Q}_R}$ is the characteristic function associated to the ℓ_∞ ball $\bar{Q}_R$ (cf. (1.23)). Then, the Brenier map $T = \nabla\phi : \bar{Q}_R \longrightarrow \bar{Q}_R$ from f to g verifies

$$W_{\infty,1}(f,g) \leq \|\,\|T - I\|_1\|_{L^\infty(\bar{Q}_R)} \leq \frac{2}{1+2\gamma}|x - \tilde{x}|.$$

As explained above, we omit the proof since it follows the formal proof of Lemma 2.6 in Section 2 and the optimality conditions in Lemma 3.2. In particular, by setting $V = \boldsymbol{V}$ (and therefore $\gamma = \beta$) we have that Lemma 3.3 is directly applicable to the truncations to $\bar{Q}_R$ of the densities $\boldsymbol{p}, \tilde{\boldsymbol{p}}$ in (2.9)-(2.10).

Definition 3.4 (Truncation to $\bar{Q}_R$). *For the probability densities $\boldsymbol{p},\tilde{\boldsymbol{p}}\in L^1_+(\mathbb{R}^2)\cap\mathcal{P}(\mathbb{R}^2)$ given in* (2.9)-(2.10)*, we define their truncations to the ℓ_∞ ball $\bar{Q}_R$ (cf.* (1.22)*) as follows*

$$\begin{aligned}\boldsymbol{p}_R(z) &:= \frac{1}{\boldsymbol{Z}_R}e^{-\boldsymbol{W}_R(z)}, &\quad \tilde{\boldsymbol{p}}_R(z) &:= \frac{1}{\tilde{\boldsymbol{Z}}_R}e^{-\tilde{\boldsymbol{W}}_R(z)},\\ \boldsymbol{W}_R(z) &:= \boldsymbol{W}(z)+\chi_{\bar{Q}_R}(z), &\quad \tilde{\boldsymbol{W}}_R(z) &:= \tilde{\boldsymbol{W}}(z)+\chi_{\bar{Q}_R}(z),\\ \boldsymbol{Z}_R &:= \int_{\mathbb{R}^2} e^{-\boldsymbol{W}_R(z)}\,dz, &\quad \tilde{\boldsymbol{Z}}_R &:= \int_{\mathbb{R}^2} e^{-\tilde{\boldsymbol{W}}_R(z)}\,dz,\end{aligned}$$

for any $R>0$, where $\chi_{\bar{Q}_R}$ is the characteristic function associated to the ℓ_∞ ball $\bar{Q}_R$ (cf. (1.23)*).*

Then, we are in position to rigorously prove Lemma 2.6 by taking limits $R\to\infty$ and noting that Lemma 3.3 yields a uniform bound of the displacement independent on R.

Rigorous proof of Lemma 2.6. Consider $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ given in (2.9)-(2.10) and set the associated Brenier map $T:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ from $\boldsymbol{p}$ to $\tilde{\boldsymbol{p}}$. Similarly, we consider the family of truncations $\boldsymbol{p}_R$ and $\tilde{\boldsymbol{p}}_R$ in Definition 3.4 and we set the associated Brenier maps $T_R:\mathbb{R}^2\longrightarrow\mathbb{R}^2$. By the above Lemma 3.3 we have

$$\|\,\|T_R-I\|_1\|_{L^\infty(\bar{Q}_R)}\leq\frac{2}{1+2\beta}|x-\tilde{x}|, \tag{3.8}$$

for every $R>0$. We set the optimal transference plans $\gamma\in\Gamma_o(\boldsymbol{p},\tilde{\boldsymbol{p}})$ and $\gamma_R\in\Gamma_o(\boldsymbol{p}_R,\tilde{\boldsymbol{p}}_R)$ associated with the $W_{2,2}$ distance, which are known to be supported on the graph of the above Brenier maps, *i.e.*,

$$\gamma:=(I,T)_{\#}\,\boldsymbol{p},\quad \gamma_R:=(I,T_R)_{\#}\,\boldsymbol{p}_R\,.$$

Since the involved potentials $\boldsymbol{W}$ and $\tilde{\boldsymbol{W}}$ are β-convex, we have the enough integrability on $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ to ensure that $\boldsymbol{p},\tilde{\boldsymbol{p}}\in\mathcal{P}_2(\mathbb{R}^2)$. Hence, the dominated convergence theorem applies and we have indeed

$$\boldsymbol{p}_R\to\boldsymbol{p},\quad \tilde{\boldsymbol{p}}_R\to\tilde{\boldsymbol{p}}\quad\text{in}\quad(\mathcal{P}_2(\mathbb{R}^2),W_{2,2}).$$

By stability of optimal transference plans, the sequence γ_R must converge narrowly to some optimal transference plan (up to a subsequence), see [1, Proposition 7.1.3]. Since the unique optimal transference plan between $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ is precisely the above γ supported on the graph of T, then we obtain

$$\gamma_R\to\gamma\quad\text{narrowly in }\mathcal{P}(\mathbb{R}^2).$$

Now we use the Kuratowski convergence of the supports under the narrow convergence of measures, see [1, Proposition 5.1.8]. Namely, consider any $z\in\mathbb{R}^2$. Since $(z,T(z))\in\operatorname{supp}\gamma$, then there exists $(z^R,w^R)\in\operatorname{supp}\gamma_R$ such that $(z^R,w^R)\to(z,T(z))$. Since γ_R is supported on the graph of T_R then $z^R\in\bar{Q}_R$ and $w^R=T_R(z^R)$. In particular, we have $T_R(z^R)-z^R\to T(z)-z$ as $R\to\infty$ and by the above uniform bound (3.8) the same bound is preserved in the limit, that is,

$$W_{\infty,1}(\boldsymbol{p},\tilde{\boldsymbol{p}})\leq\|\,\|T-I\|_1\|_{L^\infty}\leq\frac{2}{1+2\beta}|x-\tilde{x}|.$$

□

Remark 3.5 (Replacing ℓ_∞ balls by ℓ_1 balls). *We note that in Lemmas 3.2 and 3.3 the choice of ℓ_∞ is crucial. However, this is not the only possible choice and a similar proof could be obtained if replacing ℓ_∞ balls with ℓ_1 balls. It is clear anyway that the shape of the boundary and the norm to be optimized should satisfy some form of compatibility conditions.*

4. Analysis of a truncated problem

In this part, we study an auxiliary version of the original time marching problem (1.1) restricted to the bounded interval $I_R:=(-R,R)$ with $R>0$, namely,

$$F_n^R=\mathcal{T}_R[F_{n-1}^R],\quad n\in\mathbb{N},\ x\in\mathbb{R}. \tag{4.1}$$

Here, we truncate the selection function m_R as follows

$$m_R(x):=m(x)+\chi_{\bar{I}_R}(x),\quad x\in\mathbb{R}, \tag{4.2}$$

where $\chi_{\bar{I}_R}$ is the characteristic function associated to the interval $\bar{I}_R$ (*cf.* (1.23)), so that the truncated integral operator $\mathcal{T}_R$ takes the form

$$\mathcal{T}_R[F](x):=e^{-m_R(x)}\iint_{\mathbb{R}^2}G\left(x-\frac{x_1+x_2}{2}\right)F(x_1)\,\frac{F(x_2)}{\|F\|_{L^1}}\,dx_1\,dx_2,\quad x\in\mathbb{R}. \tag{4.3}$$

Again, solutions of the form $F_n^R(x) = (\lambda^R)^n\, F^R(x)$ come as eigenpairs of the non-linear eigenproblem

$$\begin{aligned} &\lambda^R F^R = \mathcal{T}_R[F^R], \quad x \in \mathbb{R},\\ &F^R \geq 0, \quad \int_{\mathbb{R}} F^R(x)\,dx = 1. \end{aligned} \tag{4.4}$$

The goal of this section is to derive an analogous truncated version of Theorem 1.2. More specifically, we study: (i) Existence of a unique strongly log-concave solution $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ to (4.4), and (ii) Quantitative relaxation of the solutions to (4.1) towards the quasi-equilibrium $(\boldsymbol{\lambda}^R)^n\, \boldsymbol{F}^R$.

Theorem 4.1 (Truncated problem)**.** *Consider any $m \in C^2(\mathbb{R})$ verifying* (H1)-(H2) *in Theorem 1.2. Set any $R > 0$ and define the truncation m_R according to* (4.2)*. Then, the following statements hold true:*

- *(i)* (**Existence of quasi-equilibrium**)
  There is a unique solution $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ to (4.4)*. In addition, $\boldsymbol{F}^R = e^{-\boldsymbol{V}^R} \in L^1_+(\mathbb{R}) \cap C^\infty(\bar{I}_R)$ is compactly supported on $\bar{I}_R$ and bounded away from zero on it and β-log-concave with parameter $\beta > 0$ given in* (1.7) *in Theorem 1.2.*
- *(ii)* (**One-step contraction**)
  Consider any $F_0^R \in L^1_+(\mathbb{R}) \cap C^1(\bar{I}_R)$ compactly supported on $\bar{I}_R$ and bounded away from zero on it, and let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to (4.1) *issued at F_0^R. Then, we have*
  $$\left\| \frac{d}{dx}\left(\log \frac{F_n^R}{\boldsymbol{F}^R}\right)\right\|_{L^\infty(\bar{I}_R)} \leq \frac{2}{1+2\,\beta} \left\| \frac{d}{dx}\left(\log \frac{F_{n-1}^R}{\boldsymbol{F}^R}\right)\right\|_{L^\infty(\bar{I}_R)},$$
  for any $n \in \mathbb{N}$.
- *(iii)* (**Asynchronous exponential growth**)
  Consider any $F_0^R \in L^1_+(\mathbb{R}) \cap C^1(\bar{I}_R)$ compactly supported on $\bar{I}_R$ and bounded away from zero on it, and let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to (4.1) *issued at F_0^R. Then, we have*
  $$\begin{aligned} \left| \frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}} - \boldsymbol{\lambda}^R \right| &\leq C_R \left(\frac{2}{1+2\,\beta}\right)^n,\\ \left\| \frac{F_n^R}{\|F_n^R\|_{L^1}} - \boldsymbol{F}^R \right\|_{C^1} &\leq C_R' \left(\frac{2}{1+2\,\beta}\right)^n, \end{aligned}$$
  for any $n \in \mathbb{N}$ and some constants C_R, C_R' depending on R and F_0^R.

As we show below, our proof exploits the overarching local contraction Lemma 3.3 to answer simultaneously both questions. More specifically, our main observation is the following type of contraction which holds true providing that the initial data F_0^R is strongly log-concave.

Lemma 4.2 (Cauchy-type property)**.** *Let $m \in C^2(\mathbb{R})$ satisfy* (H1)-(H2) *in Theorem 1.2. Consider a β_0 log-concave density $F_0^R \in L^1_+(\mathbb{R}) \cap C^{1,\delta}(\bar{I}_R)$ with $\beta_0 > 0$ and $0 < \delta < 1$, compactly supported on $\bar{I}_R$ and bounded away from zero on it. Let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to* (4.1) *issued at F_0^R. Then, we have*

$$\left\| \frac{d}{dx}\left(\log \frac{F_n^R}{F_{n-1}^R}\right)\right\|_{L^\infty(\bar{I}_R)} \leq \frac{2}{1+2\,\beta_{n-2}} \left\| \frac{d}{dx}\left(\log \frac{F_{n-1}^R}{F_{n-2}^R}\right)\right\|_{L^\infty(\bar{I}_R)}, \quad n \geq 2,$$

where the sequence $\{\beta_n\}_{n\in\mathbb{N}}$ is defined by recurrence like in (2.2)*.*

Proof. For any $n \in \mathbb{N}$, we define

$$u_n^R(x) := \frac{F_n^R(x)}{F_{n-1}^R(x)}, \quad x \in \bar{I}_R,$$

and note that, arguing as in (2.3), we have that $\{u_n\}_{n\in\mathbb{N}}$ must solve the following analogue of (2.4):

$$u_n^R(x) = \frac{\|F_{n-2}^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}} \iint_{\bar{Q}_R} P_n^R(x; x_1, x_2)\, u_{n-1}^R(x_1)\, u_{n-1}^R(x_2)\, dx_1\, dx_2,$$

for any $x \in \bar{I}_R$ and $n \geq 2$. We remark that the system above holds only on $\bar{I}_R$ and the one-step transition probability $P_n^R(x;\cdot) \in L^1_+(\bar{Q}_R) \cap \mathcal{P}(\bar{Q}_R)$ is not time-homogeneous but it depends explicitly on n, namely

$$\begin{aligned}
P_n^R(x;x_1,x_2) &:= \frac{1}{Z_n^R(x)} e^{-W_n^R(x;x_1,x_2)}, \quad x \in \bar{I}_R, \quad (x_1,x_2) \in \bar{Q}_R,\\
W_n^R(x;x_1,x_2) &:= \frac{1}{2}\left|x - \frac{x_1+x_2}{2}\right|^2 + V_{n-2}^R(x_1) + V_{n-2}^R(x_2),\\
Z_n^R(x) &:= \iint_{\bar{Q}_R} e^{-W_n^R(x;x_1,x_2)}\,dx_1\,dx_2,
\end{aligned}$$

where we denote $V_n^R : \bar{I}_R \longrightarrow \mathbb{R}$ so that $F_n^R = e^{-V_n^R}$. By Lemma 2.2, V_{n-2}^R is β_{n-2}-convex and therefore the contractivity Lemma 3.3 applies to $f = P_n^R(x;\cdot)$ and $g = P_n^R(\tilde{x};\cdot)$ with $x, \tilde{x} \in \bar{I}_R$ leading to

$$W_{\infty,1}(P_n^R(x;\cdot), P_n^R(\tilde{x};\cdot)) \leq \frac{2}{1+2\,\beta_{n-2}}|x-\tilde{x}|.$$

Therefore, arguing as in Lemma 2.4 we end the proof. □

Proof of Theorem 4.1.

◇ Step 1: Proof of (i).

Under appropriate assumptions on F_0^R we shall prove that $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$ and $F_n^R/\|F_n^R\|_{L^1}$ must converge as in (iii), and their limit $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ solves (4.4). We set a β_0-log-concave density $F_0^R \in L^1_+(\mathbb{R}) \cap C^{1,\delta}(\bar{I}_R)$ with $\beta_0 > \beta$ and $0 < \delta < 1$, compactly supported on $\bar{I}_R$ and bounded away from zero on it. Let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to (4.1). Since the initial datum has been chosen strongly log-concave, Lemma 4.2 implies

$$\left\|\frac{d}{dx}\left(\log\frac{F_n^R}{F_{n-1}^R}\right)\right\|_{L^\infty(\bar{I}_R)} \leq \left(\frac{2}{1+2\,\beta}\right)^{n-1}\left\|\frac{d}{dx}\left(\log\frac{F_1^R}{F_0^R}\right)\right\|_{L^\infty(\bar{I}_R)},$$

for all $n \geq 1$ because F_n^R are β_n-log-concave with $\beta_n > \beta$ for all $n \in \mathbb{N}$ by Lemma 2.2. Setting $V_n^R : \bar{I}_R \longrightarrow \mathbb{R}$ as before so that $F_n^R = e^{-V_n^R}$ we obtain

$$\left\|\frac{d}{dx}(V_n^R - V_m^R)\right\|_{L^\infty(\bar{I}_R)} \leq \sum_{k=m+1}^n \left\|\frac{d}{dx}(V_k^R - V_{k-1}^R)\right\|_{L^\infty(\bar{I}_R)} \leq \sum_{k=m}^{n-1}\left(\frac{2}{1+2\,\beta}\right)^k \left\|\frac{d}{dx}(V_1^R - V_0^R)\right\|_{L^\infty(\bar{I}_R)},$$

for all $n \geq m \geq 1$. Since $\frac{2}{1+2\,\beta} < 1$ by Remark 1.7, then $\left\{\frac{d}{dx}(V_n^R)\right\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $C(\bar{I}_R)$ and therefore it must converge uniformly to some limit $D^R \in C(\bar{I}_R)$. In particular, we have

$$\frac{d}{dx}\left(\log F_n^R\right) \to D^R \quad \text{in} \quad C(\bar{I}_R). \tag{4.5}$$

Now, we show that $F_n^R/\|F_n^R\|_{L^1}$ must also converge when evaluated at least at one point, and we choose $x = 0$ for instance. To this purpose, we note that $F_n^R(0)/\|F_n^R\|_{L^1}$ can be restated as follows

$$\frac{\displaystyle\iint_{\bar{Q}_R} G\left(\frac{x_1+x_2}{2}\right)\exp\left(-(V_{n-1}^R(x_1) - V_{n-1}^R(0)) - (V_{n-1}^R(x_2) - V_{n-1}^R(0))\right)dx_1\,dx_1}{\displaystyle\int_{\bar{I}_R}\iint_{\bar{Q}_R} G\left(x' - \frac{x_1+x_2}{2}\right)\exp\left(-m(x') - (V_{n-1}^R(x_1) - V_{n-1}^R(0)) - (V_{n-1}^R(x_2) - V_{n-1}^R(0))\right)dx'\,dx_1\,dx_2},$$

and, by the fundamental theory of calculus, $V_{n-1}^R(x) - V_{n-1}^R(0)$ in the integrand can be represented by

$$V_{n-1}^R(x) - V_{n-1}^R(0) = \int_0^1 \frac{dV_{n-1}^R}{dx}(\theta x)\,x\,d\theta, \quad x \in \bar{I}_R,$$

which converges uniformly to some limit. Therefore, there exists $L^R \in \mathbb{R}$ such that

$$\log\frac{F_n^R(0)}{\|F_n^R\|_{L^1}} \to L^R. \tag{4.6}$$

Putting (4.5)-(4.6) together and using the fundamental theorem of calculus entail

$$\log\frac{F_n^R(x)}{\|F_n^R\|_{L^1}} = \log\frac{F_n^R(0)}{\|F_n^R\|_{L^1}} + \int_0^1 \frac{d}{dx}\left(\log F_n^R\right)(\theta x)\,x\,d\theta \to L^R + \int_0^1 D^R(\theta x)\,x\,d\theta \quad \text{in} \quad C^1(\bar{I}_R).$$

We define $\boldsymbol{F}^R(x) := \exp(L^R + \int_0^1 D^R(\theta x)\, x\, d\theta + \chi_{\bar{I}_R}(x)) \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R})$ and therefore we achieve

$$\frac{F_n^R}{\|F_n^R\|_{L^1}} \to \boldsymbol{F}^R \quad \text{in} \quad C^1(\bar{I}_R). \tag{4.7}$$

Our second step is to prove the convergence of $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$. Note that we have

$$\frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}} = \iint_{\mathbb{R}^2} H_R(x_1,x_2)\, \frac{F_{n-1}^R(x_1)}{\|F_{n-1}^R\|_{L^1}}\, \frac{F_{n-1}^R(x_2)}{\|F_{n-1}^R\|_{L^1}}\, dx_1\, dx_2, \tag{4.8}$$

where we have defined

$$H_R(x_1,x_2) := \int_{\bar{I}_R} e^{-m(x)} G\left(x - \frac{x_1+x_2}{2}\right) dx, \quad (x_1,x_2) \in \mathbb{R}^2.$$

Since H_R is a bounded function, therefore $H_R \in L^1(\bar{Q}_R)$ and, consequently, the above uniform convergence (4.7) of the normalized profiles along with (4.8) imply that there must exists $\boldsymbol{\lambda}^R$ with

$$\frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}} \to \boldsymbol{\lambda}^R\,. \tag{4.9}$$

The last step is to show that $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ must solve (4.4). This is actually clear because we have

$$\frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}}\, \frac{F_n^R}{\|F_n^R\|_{L^1}} = \mathcal{T}_R\left[\frac{F_{n-1}^R}{\|F_{n-1}^R\|_{L^1}}\right],$$

for all $n \in \mathbb{N}$, and $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$ and $F_n^R/\|F_n^R\|_{L^1}$ converge in the above sense (4.7)-(4.9). We note that $\boldsymbol{F}^R$ must be β-log-concave because so is F_n^R for all $n \in \mathbb{N}$. The uniqueness of solution to (4.4) will not be analyzed here, but it will hold as a consequence of the next contraction property in Step 2.

$\diamond$ Step 2: Proof of *(ii)*.
Once a strongly log-concave solution $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ of the truncated nonlinear eigenproblem (4.4) exists, the one-step contraction property follows the same ideas as in the global version in Theorem 1.2(ii) sketched in Section 2. More specifically, we shall argue like in the proof of Lemma 4.2 where again we replace u_n by the normalization of F_n^R by the quasi-equilibrium $(\boldsymbol{\lambda}^R)^n\, \boldsymbol{F}^R$. That is, for any $n \in \mathbb{N}$, we define

$$u_n^R(x) := \frac{F_n^R(x)}{(\boldsymbol{\lambda}^R)^n\, \boldsymbol{F}^R}, \quad x \in \bar{I}_R,$$

which must solve

$$u_n^R(x) = \frac{1}{\|u_{n-1}^R\, \boldsymbol{F}^R\|_{L^1}} \iint_{\bar{Q}_R} \boldsymbol{P}^R(x;x_1,x_2)\, u_{n-1}^R(x_1)\, u_{n-1}^R(x_2)\, dx_1\, dx_2,$$

for any $x \in \bar{I}_R$ and $n \in \mathbb{N}$, where $\boldsymbol{P}^R(x;\cdot) \in L^1_+(\bar{Q}_R) \cap \mathcal{P}(\bar{Q}_R)$ is the one-step transition probability

$$\begin{aligned}
\boldsymbol{P}^R(x;x_1,x_2) &:= \frac{1}{\boldsymbol{Z}^R(x)} e^{-\boldsymbol{W}^R(x;x_1,x_2)}, \quad x \in \bar{I}_R, \quad (x_1,x_2) \in \bar{Q}_R,\\
\boldsymbol{W}^R(x;x_1,x_2) &:= \frac{1}{2}\left|x - \frac{x_1+x_2}{2}\right|^2 + \boldsymbol{V}^R(x_1) + \boldsymbol{V}^R(x_2),\\
\boldsymbol{Z}^R(x) &:= \iint_{\bar{Q}_R} e^{-\boldsymbol{W}^R(x;x_1,x_2)}\, dx_1\, dx_2.
\end{aligned}$$

Again, we denote $\boldsymbol{V}^R : \bar{I}_R \longrightarrow \mathbb{R}$ so that $\boldsymbol{F}^R = e^{-\boldsymbol{V}^R}$. By Step 1 we have that $\boldsymbol{V}^R$ is β-convex and therefore the contractivity Lemma 3.3 applies to $\boldsymbol{P}^R(x;\cdot)$ and $\boldsymbol{P}^R(\tilde{x};\cdot)$ with $x, \tilde{x} \in \bar{I}_R$ leading to

$$W_{\infty,1}(\boldsymbol{P}^R(x;\cdot), \boldsymbol{P}^R(\tilde{x};\cdot)) \leq \frac{2}{1+2\,\beta}|x - \tilde{x}|.$$

Therefore, arguing as in Lemma 2.4 we end the proof.

In particular, the above implies that $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ must be the unique solution to the truncated nonlinear eigenproblem 4.4. Indeed, if a second solution (λ^R, F^R) exists, one can always define the special solution $F_n^R(x) = (\lambda^R)^n F^R(x)$ of (4.1) and therefore the above one-step contraction implies

$$\left\|\frac{d}{dx}\left(\log\frac{F^R}{\boldsymbol{F}^R}\right)\right\|_{L^\infty(\bar{I}_R)} \leq \frac{2}{1+2\,\beta}\left\|\frac{d}{dx}\left(\log\frac{F^R}{\boldsymbol{F}^R}\right)\right\|_{L^\infty(\bar{I}_R)}.$$

Since $\frac{2}{1+2\,\beta} < 1$ by Remark 1.7, then we have $F^R = \boldsymbol{F}^R$ (and therefore $\lambda^R = \boldsymbol{\lambda}^R$) because both F^R and $\boldsymbol{F}^R$ are probability densities by definition.

⋄ Step 3: Proof of (*iii*).
We prove that the convergence in Step 1 holds for generic initial data $F_0^R \in L^1_+(\mathbb{R}) \cap C^1(\bar{I}_R)$ compactly supported on $\bar{I}_R$ and bounded away from zero on it, and not necessarily strongly log-concave. Note that by the above one-step contractivity property we have again

$$\left\| \frac{d}{dx}(V_n^R - \boldsymbol{V}^R) \right\|_{L^\infty(\bar{I}_R)} \leq \left(\frac{2}{1+2\beta} \right)^n \left\| \frac{d}{dx}(V_0^R - \boldsymbol{V}^R) \right\|_{L^\infty(\bar{I}_R)},$$

for all $n \in \mathbb{N}$. Then, the same argument as in Step 1 can be applied with explicit convergence rates and equal to $\left(\frac{2}{1+2\beta}\right)^n$ at each step: first $\frac{d}{dx}(\log F_n^R)$, second $\log\left(F_n^R(0)/\|F_n^R\|_{L^1}\right)$, hence $\log\left(F_n^R/\|F_n^R\|_{L^1}\right)$, and finally also $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$. Therefore, we readily obtain the claimed convergence rates for the rates of growth and the normalized profiles. □

5. Existence and uniqueness of strongly log-concave quasi-equilibria

In this section, we employ the truncated quasi-equilibria in the above Theorem 4.1 to build a globally defined quasi-equilibrium of the non-truncated model (1.1), thus proving Theorem 1.2(i). In the following, we show that the family of probability densities $\{\boldsymbol{F}^R\}_{R>0}$ are uniformly tight, and therefore weak limits cannot lose mass at infinity, which will be useful in the sequel in order to pass to the limit with $R \to \infty$.

Proposition 5.1 (Bounded second-order moments)**.** *Under the assumptions in Theorem 4.1, let us consider the unique eigenpair* $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ *of* (4.4) *for any* $R > 0$ *according to Theorem 4.1(i). Then,*

$$\sup_{R>0} \int_{\mathbb{R}} x^2\, \boldsymbol{F}^R(x)\, dx < \infty. \tag{5.1}$$

We recall that a similar result was necessary in [16]. Indeed, a general strategy was developed therein to propagate second-order moments along any solution $\{F_n\}_{n\in\mathbb{N}}$ under the *a priori* knowledge that the centers of mass stay uniformly bounded. However, such a condition proved difficult to verify unless the initial datum F_0 is centered at the origin, and m is an even function, which would leave the center of mass fixed at the origin (and thus bounded) for all times. To overcome this problem, an alternative approach was developed in [16, Lemma 4.5] in order to control the convergence to zero of the center of mass in the case of quadratic selection. Unfortunately, the proof exploits the Gaussian structure in a crucial way and cannot be easily adapted to more general selection functions. Here, we propose an alternative strategy based on the extra knowledge that $\boldsymbol{F}^R$ are β-log-concave.

Proof of Proposition 5.1.

⋄ Step 1: Uniform bound of the variance.
Let us define the center of mass and the variance

$$\boldsymbol{\mu}_R := \int_{\mathbb{R}} x\, \boldsymbol{F}^R(x)\, dx,$$
$$\boldsymbol{\sigma}_R^2 := \int_{\mathbb{R}} (x - \boldsymbol{\mu}_R)^2\, \boldsymbol{F}^R(x)\, dx,$$

for any $R > 0$. Since each eigenfunction $\boldsymbol{F}^R$ is β-log-concave, then a straightforward application of the Brascamp-Lieb inequality shows that variances $\boldsymbol{\sigma}_R^2$ verify

$$\boldsymbol{\sigma}_R^2 \leq \frac{1}{\beta}, \tag{5.2}$$

for any $R > 0$, see [8, Theorem 4.1]. Then, in order to control the (non-centered) second order moments, we actually need to find a bound of the center of mass $\boldsymbol{\mu}_R$.

⋄ Step 2: Uniform bound of the center of mass.
Assume that $\{\boldsymbol{\mu}_R\}_{R>0}$ is unbounded by contradiction. Changing variables x with $-x$ if necessary, we may assume without loss of generality that $\boldsymbol{\mu}_R \nearrow +\infty$ as $R \nearrow +\infty$ up to an appropriate subsequence, which we denote in the same way for simplicity of notation. Note that integrating (4.4) against $e^{m_R(x)}$ and remarking that $\int_{\mathbb{R}} \mathcal{B}[\boldsymbol{F}^R](x)\, dx = \int_{\mathbb{R}} \boldsymbol{F}^R(x)\, dx = 1$ (where $\mathcal{B}$ is given in (1.3)) we obtain

$$A_R\, B_R = 1, \tag{5.3}$$

for every $R > 0$, where each factor reads

$$A_R := \int_{\mathbb{R}} e^{m_R(x)}\ \boldsymbol{F}^R(x)\,dx,$$

$$B_R := \int_{\mathbb{R}^2} \phi^R\left(\frac{x_1+x_2}{2}\right) \boldsymbol{F}^R(x_1)\ \boldsymbol{F}^R(x_2)\,dx_1\,dx_2,$$

and $\phi^R := G * e^{-m_R}$. By Chebyshev's inequality we know that

$$\int_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} \boldsymbol{F}^R(x)\,dx \geq \frac{1}{2}, \tag{5.4}$$

for all $R > 0$. Therefore, noting that m is non-decreasing in $\mathbb{R}_+$ by virtue of the hypothesis (H1)-(H2) we obtain the following lower bound

$$\begin{aligned} A_R &\geq \int_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} e^{m_R(x)}\ \boldsymbol{F}^R(x)\,dx \\ &\geq \frac{1}{2}\min_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} e^{m(x)} = \frac{1}{2}e^{m(\boldsymbol{\mu}_R-\sqrt{2}\,\boldsymbol{\sigma}_R)}, \end{aligned} \tag{5.5}$$

for large enough $R > 0$ so that $[\boldsymbol{\mu}_R - \sqrt{2}\,\boldsymbol{\sigma}_R, \boldsymbol{\mu}_R + \sqrt{2}\,\boldsymbol{\sigma}_R] \subset \mathbb{R}_+$. Similarly, using (5.4) and noting that ϕ^R is non-increasing at the right of its maximizer (by strong log-concavity, *cf.* Lemma 2.2) we obtain

$$\begin{aligned} B_R &\geq \iint_{|x_i-\boldsymbol{\mu}_R|\leq\sqrt{2}\sigma_R} \phi^R\left(\frac{x_1+x_2}{2}\right) \boldsymbol{F}^R(x_1)\ \boldsymbol{F}^R(x_2)\,dx_1\,dx_2 \\ &\geq \frac{1}{4}\min_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} \phi^R(x) \geq \frac{1}{4}\phi^R(\boldsymbol{\mu}_R+\sqrt{2}\,\boldsymbol{\sigma}_R), \end{aligned} \tag{5.6}$$

for large enough $R > 0$ so that $[\boldsymbol{\mu}_R - \sqrt{2}\,\boldsymbol{\sigma}_R, \boldsymbol{\mu}_R + \sqrt{2}\,\boldsymbol{\sigma}_R]$ lies in that region of the domain. Note that the above can be obtained if $R > 0$ is large enough since $\boldsymbol{\mu}^R - \sqrt{2}\,\boldsymbol{\sigma}_R \to \infty$ by assumptions, but however the maximizers of ϕ^R must converge to the maximizer of ϕ, which is a fixed number in the real line. Multiplying (5.5) and (5.6) yields the lower bound

$$A_R B_R \geq \frac{1}{8}e^{m_R(\boldsymbol{\mu}_R-\sqrt{2}\,\boldsymbol{\sigma}_R)}\,(G * e^{-m_R})(\boldsymbol{\mu}_R+\sqrt{2}\,\boldsymbol{\sigma}_R), \tag{5.7}$$

for large enough $R > 0$. Lemma B.2 provides a explicit lower bound (B.6) on Gaussian convolutions. Therefore, applying it to the second factor in (5.7) with the choices

$$f = e^{-m}, \quad \gamma = \alpha, \quad x_0 = \boldsymbol{\mu}_R, \quad \delta = \sqrt{2}\,\boldsymbol{\sigma}_R,$$

implies the following lower bound

$$\begin{aligned} A_R B_R &\geq G(2\sqrt{2}\,\boldsymbol{\sigma}_R)\int_0^{\frac{\alpha}{\alpha+1}\boldsymbol{\mu}_R-\frac{\sqrt{2}\,\boldsymbol{\sigma}_R}{\alpha+1}} \exp\left(\frac{\alpha+1}{2}z^2\right)dz \\ &\geq G\left(\frac{2\sqrt{2}}{\sqrt{\beta}}\right)\int_0^{\frac{\alpha}{\alpha+1}\boldsymbol{\mu}_R-\frac{\sqrt{2}}{\sqrt{\beta}(\alpha+1)}} \exp\left(\frac{\alpha+1}{2}z^2\right)dz, \end{aligned} \tag{5.8}$$

where in the last line we have used the bound (5.2) of variances. Since the left hand side in (5.8) diverges as $R \to \infty$ because $\boldsymbol{\mu}_R \to +\infty$, then we reach a contradiction with (5.3), and this ends the proof. □

Theorem 5.2 (Existence of quasi-equilibria)**.** *Under the assumptions in Theorem 4.1, let us consider the unique eigenpair* $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ *of* (4.4) *for any* $R > 0$*. Then, there exist* $\boldsymbol{\lambda} \in \mathbb{R}$ *and* $\boldsymbol{F} \in L^1_+(\mathbb{R}) \cap C^\infty(\mathbb{R})$ *which is* β*-log-concave (with* β *given in* (1.7)*) such that*

$$\boldsymbol{\lambda}^R \to \boldsymbol{\lambda}, \quad \boldsymbol{F}^R \to \boldsymbol{F}, \quad \text{as} \quad R \to \infty,$$

up to subsequence, both pointwise and in any space $(\mathcal{P}_p(\mathbb{R}), W_p)$ *with* $1 \leq p < 2$*. Moreover, the pair* $(\boldsymbol{\lambda}, \boldsymbol{F})$ *is the unique solution to* (1.5) *among all pairs* (λ, F) *verifying* (1.8)*.*

Proof.

◇ Step 1: Existence via limit as $R \to \infty$.
Let us notice that by (5.1) in Proposition 5.1 we have that $\{\boldsymbol{F}^R\}_{R>0}$ is a uniformly tight sequence of probability measures. Therefore, by Prokhorov's theorem there must exist $R_n \nearrow \infty$ and some limiting probability measure $\boldsymbol{F} \in \mathcal{P}(\mathbb{R})$ such that

$$\boldsymbol{F}^{R_n} \to \boldsymbol{F} \quad \text{narrowly in } \mathcal{P}(\mathbb{R}). \tag{5.9}$$

By integration on (4.4) we also obtain that

$$\boldsymbol{\lambda}^{R_n} = \iint_{\mathbb{R}^2} (e^{-m_{R_n}} * G)\left(\frac{x_1+x_2}{2}\right) \boldsymbol{F}^{R_n}(x_1)\ \boldsymbol{F}^{R_n}(x_2)\,dx_1\,dx_2,$$

and then we can pass to the limit as $n\to\infty$ in the eigenvalues too. Specifically, since $e^{-m_R}\to e^{-m}$ in $L^\infty(\mathbb{R})$, then $e^{-m_R}*G\to e^{-m}*G$ in $C_b(\mathbb{R})$, and therefore by (5.9) we obtain

$$\boldsymbol{\lambda}^{R_n}\to\boldsymbol{\lambda}, \tag{5.10}$$

as $n\to\infty$, where $\boldsymbol{\lambda}$ is given by

$$\boldsymbol{\lambda} := \iint_{\mathbb{R}^2} (e^{-m} * G)\left(\frac{x_1+x_2}{2}\right) \boldsymbol{F}(x_1)\ \boldsymbol{F}(x_2)\,dx_1\,dx_2 = \int_{\mathbb{R}} \mathcal{T}[\boldsymbol{F}](x)\,dx. \tag{5.11}$$

Putting (5.9) and (5.10) together and taking limits as $n\to\infty$ in (4.4) implies that $\{\boldsymbol{F}^{R_n}\}_{n\in\mathbb{N}}$ must also converge pointwise to some other limit $\tilde{\boldsymbol{F}}\in L^1_+(\mathbb{R})$ by Fatou's lemma. Note that since $\boldsymbol{F}^R$ are all β-log-concave, then so must also be their pointwise limit $\tilde{\boldsymbol{F}}$. Indeed, note that we further have

$$\boldsymbol{\lambda}\,\tilde{\boldsymbol{F}}(x) = \mathcal{T}[\boldsymbol{F}](x),\quad x\in\mathbb{R}, \tag{5.12}$$

and therefore, $\tilde{\boldsymbol{F}}\in L^1_+(\mathbb{R})\cap\mathcal{P}(\mathbb{R})$, in view of (5.11). Then, we actually have $\boldsymbol{F}^{R_n}\to\tilde{\boldsymbol{F}}$ in $L^1(\mathbb{R})$ (thus narrowly in $\mathcal{P}(\mathbb{R})$) by Scheffé's lemma. Since $\boldsymbol{F}$ is a narrow limit of the same sequence, then we have $\tilde{\boldsymbol{F}}=\boldsymbol{F}$ and by (5.12) we obtain that $(\boldsymbol{\lambda},\boldsymbol{F})$ must verify the initial problem (1.5). Let us also emphasize that, we indeed have convergence in any L^p Wasserstein space with $1\leq p<2$ because all the p-th order moment with $1\leq p<2$ are uniformly integrable by (5.1), see [1, Proposition 7.1.5].

⋄ Step 2: Uniqueness of quasi-equilibria.
Note that several different convergent subsequences of $\{\boldsymbol{F}^R\}_{R>0}$ in Step 1 could give rise to various eigenpairs $(\boldsymbol{\lambda},\boldsymbol{F})$ of (1.5). Whilst the global uniqueness is unclear with this method, we prove that there can only exist one solution to (1.5) among the pairs (λ,F) verifying (1.8). For, we exploit the one-step contraction property in Theorem 1.2(ii). Specifically, assume that (λ,F) is any other solution to (1.5) and define $F_n(x)=\lambda^n F(x)$, which is clearly a solution to the evolution problem (1.1) with initial datum $F_0\in L^1_+(\mathbb{R})\cap C^1(\mathbb{R})$ verifying the hypothesis (H3) by virtue of the assumption (1.8). Then, (1.9) implies

$$\left\|\frac{d}{dx}\left(\log\frac{F}{\boldsymbol{F}}\right)\right\|_{L^\infty}\leq\frac{2}{1+2\,\beta}\left\|\frac{d}{dx}\left(\log\frac{F}{\boldsymbol{F}}\right)\right\|_{L^\infty}.$$

Again, since $\frac{2}{1+2\,\beta}<1$ by Remark 1.7, then we obtain that $F/\boldsymbol{F}$ must be constant. Since both F and $\boldsymbol{F}$ are normalized probability densities, then we necessarily have that $F=\boldsymbol{F}$ (and therefore $\lambda=\boldsymbol{\lambda}$). □

6. Convergence to equilibrium for restricted initial data

In this section, we prove asynchronous exponential as claimed in Theorem 1.2(iii). More specifically, we show that for restricted initial the asymptotic behavior of the rate of growth of mass $\|F_n\|_{L^1}/\|F_{n-1}\|_{L^1}$ and the normalized profiles $F_n/\|F_n\|_{L^1}$ is dictated by the solution $(\boldsymbol{\lambda},\boldsymbol{F})$ of the eigenproblem (1.5) obtained in Theorem 1.2(i). We derive the relaxation of the normalized profiles under the relative entropy metric. Our starting point is the one-step contraction property of the L^∞ relative Fisher information in Theorem 1.2(ii) and the following version of the logarithmic-Sobolev inequality with respect to strongly log-concave densities, which relate the (L^2) relative Fisher information and the relative entropy.

Proposition 6.1 (Logarithmic-Sobolev inequality)**.** *Consider any couple $P,Q\in L^1_+(\mathbb{R})\cap\mathcal{P}(\mathbb{R})$ such that Q is γ-log-concave for some $\gamma>0$. Then, we have*

$$\mathcal{D}_{KL}(P\|Q)\leq\frac{1}{2\gamma}\,\mathcal{I}_2(P\|Q)\leq\frac{1}{2\gamma}\,\mathcal{I}^2_\infty(P\|Q), \tag{6.1}$$

where $\mathcal{D}_{\rm KL}$ is the relative entropy (1.12)*, $\mathcal{I}_2$ is the usual (or L^2) relative Fisher information* (1.18)*, and $\mathcal{I}_\infty$ is the L^∞ relative Fisher information* (1.6)*.*

On the one hand, the first part of the inequality (6.1) amounts to the usual logarithmic-Sobolev inequality with respect to a strongly log-concave measure, see Corollary 5.7.2 and Section 9.3.1 in [4] for details. On the other hand, the second part of the inequality readily holds by definition. Therefore, putting Theorem 1.2(ii) and Proposition (6.1) together, we end the proof of Theorem 1.2(iii).

Proof of Theorem 1.2(iii). Notice that by iterating n times the one-step contraction property in Theorem 1.2(ii) and using the logarithmic-Sobolev inequality (6.1) in Proposition 6.1 we obtain

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1}}\,\middle\|\,\boldsymbol{F}\right) \leq C_1 \left(\frac{2}{1+2\,\beta}\right)^{2n}, \tag{6.2}$$

for every $n \in \mathbb{N}$, where the constant C_1 reads

$$C_1 := \frac{1}{2\gamma}\,\mathcal{I}^2_\infty\left(F_0 \| \boldsymbol{F}\right),$$

and it is finite by the assumption (H3). This proves the relaxation of the normalized profiles towards $\boldsymbol{F}$ in the relative entropy sense. Regarding the rate of growth, we note that

$$\frac{\|F_n\|_{L^1}}{\|F_{n-1}\|_{L^1}} = \iint_{\mathbb{R}^2} \phi\left(\frac{x_1+x_2}{2}\right) \frac{F_{n-1}(x_1)}{\|F_{n-1}\|_{L^1}}\,\frac{F_{n-1}(x_2)}{\|F_{n-1}\|_{L^1}}\,dx_1\,dx_2, \tag{6.3}$$

$$\boldsymbol{\lambda} = \iint_{\mathbb{R}^2} \phi\left(\frac{x_1+x_2}{2}\right) \boldsymbol{F}(x_1)\ \boldsymbol{F}(x_2)\,dx_1\,dx_2. \tag{6.4}$$

where $(\boldsymbol{\lambda}, \boldsymbol{F})$ is the solution to (1.5) in Theorem 1.2(i), and $\phi := G * e^{-m}$ again. Taking the difference of the two identities (6.3) and (6.4) above, we achieve

$$\begin{aligned}
\left|\frac{\|F_n\|_{L^1}}{\|F_{n-1}\|_{L^1}} - \boldsymbol{\lambda}\right| &\leq \|\phi\|_{L^\infty} \left\|\frac{F_{n-1}}{\|F_{n-1}\|_{L^1}} \otimes \frac{F_{n-1}}{\|F_{n-1}\|_{L^1}} - \boldsymbol{F}\otimes\boldsymbol{F}\right\|_{L^1}\\
&\leq \|\phi\|_{L^\infty} \sqrt{\frac{1}{2}\mathcal{D}_{\mathrm{KL}}\left(\frac{F_{n-1}}{\|F_{n-1}\|_{L^1}} \otimes \frac{F_{n-1}}{\|F_{n-1}\|_{L^1}}\,\middle\|\,\boldsymbol{F}\otimes\boldsymbol{F}\right)}\\
&= \|\phi\|_{L^\infty} \sqrt{\mathcal{D}_{\mathrm{KL}}\left(\frac{F_{n-1}}{\|F_{n-1}\|_{L^1}}\,\middle\|\,\boldsymbol{F}\right)}\\
&\leq C_2 \left(\frac{2}{1+2\,\beta}\right)^{n},
\end{aligned}$$

with a explicit constant $C_2 > 0$ taking the form

$$C_2 := \|\phi\|_{L^\infty} \sqrt{C_1}.$$

Note that above, we have used successively Hölder's inequality, Pinsker's inequality, the tensorization property of the relative entropy, and (6.2) to reach the conclusion. ☐

Appendix A. Intermediate dualities

For simplicity of the discussion, we do not present here the intermediate Kantorovich-type dualities in the case of non-linear transition semigroups like in (2.4), but we rather focus on linear semigroups. More specifically, we have the following intermediate result which is reminiscent of the natural interpolation of Kantorovich duality for L^1 Wasserstein distance, and Lemma 2.4 for L^∞ Wassestein metric.

Proposition A.1. *Consider any $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ for some $1 \leq p \leq \infty$, and set any function $u \in C^1(\mathbb{R}^d)$ such that $u > 0$ and $\nabla(u^{1/p}) \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$. Then, the following inequality holds true*

$$\left|\left(\int_{\mathbb{R}^d} u(x)\,\mu(dx)\right)^{1/p} - \left(\int_{\mathbb{R}^d} u(x)\,\nu(dx)\right)^{1/p}\right| \leq \left\|\,\|\nabla(u^{1/p})\|_{q'}\right\|_{L^\infty} W_{p,q}(\mu,\nu),$$

for any $1 \leq q \leq \infty$, and q' given by $\frac{1}{q} + \frac{1}{q'} = 1$. Here, $W_{p,q}$ denotes the L^p Wasserstein distance associated with ℓ_q norm of $\mathbb{R}^d$, cf. (1.24)*, and we admit the convention that $u^{1/\infty} = \log u$ for all $u > 0$.*

Proof. Let us consider any constant-speed geodesic $t \in [0,1] \longmapsto \rho_t \in \mathcal{P}_p(\mathbb{R}^d)$ in the Wasserstein space $(\mathcal{P}_p(\mathbb{R}^d), W_{p,q})$ joining μ to ν. Specifically, ρ verifies the continuity equation

$$\begin{aligned}
&\partial_t \rho_t + \operatorname{div}(\rho_t v_t) = 0, \quad t \in [0,1],\ x \in \mathbb{R}^d,\\
&\rho_0 = \mu, \quad \rho_1 = \nu,
\end{aligned} \tag{A.1}$$

in distributional sense and, in addition, we have

$$\|\,\|v_t\|_q\|_{L^p(\rho_t)} = W_{p,q}(\mu,\nu), \quad t \in [0,1]. \tag{A.2}$$

Let us also define the function

$$E(t) := \int_{\mathbb{R}^d} u(y)\, \rho_t(dy), \quad t \in [0,1].$$

Since $\rho \in \mathrm{Lip}([0,1], \mathcal{P}_p(\mathbb{R}^d))$, then $E \in AC([0,1])$ and by the continuity equation (A.1) we have

$$\frac{dE}{dt}(t) = \int_{\mathbb{R}^d} \nabla u(y) \cdot v_t(y)\, \rho_t(dy) = p \int_{\mathbb{R}^d} \nabla (u^{1/p})(y) \cdot v_t(y)\, u^{1/p'}(y)\, \rho_t(dy), \tag{A.3}$$

for *a.e.* $t \in [0,1]$, where we have used the identity $\nabla u = p\, \nabla(u^{1/p})\, u^{1/p'}$. Therefore, we obtain

$$\begin{aligned}\left|\frac{dE}{dt}(t)\right| &\leq p \int_{\mathbb{R}^d} \|\nabla(u^{1/p})(y)\|_{q'}\, \|v_t(y)\|_q\, u^{1/p'}(y)\, \rho_t(dy)\\ &\leq p \left\| \|\nabla(u^{1/p})\|_{q'} \right\|_{L^\infty} \int_{\mathbb{R}^d} \|v_t(y)\|_q\, u^{1/p'}(y)\, \rho_t(dy)\\ &\leq p \left\| \|\nabla(u^{1/p})\|_{q'} \right\|_{L^\infty} \|\, \|v_t\|_q \|_{L^p(\rho_t)}\, \|u^{1/p'}\|_{L^{p'}(\rho_t)},\end{aligned}$$

for *a.e.* $t \in [0,1]$, where in the first step we have used Hölder's inequality with exponent q applied to the inner product in the integrand of (A.3), and in the last step we have used Hölder's inequality with exponent p applied to the integral of the second line. Using the constant-speed condition (A.2) in the second factor, and $\|u^{1/p'}\|_{L^{p'}(\rho_t)} = E(t)^{1/p'}$ in the last one, we obtain the relation

$$\left|\frac{dE}{dt}(t)\right| \leq p \left\| \|\nabla(u^{1/p})\|_{q'} \right\|_{L^\infty} W_{p,q}(\mu,\nu)\, E(t)^{1/p'},$$

for *a.e.* $t \in [0,1]$, which amounts to

$$\left|\frac{dE^{1/p}}{dt}(t)\right| \leq \left\| \|\nabla(u^{1/p})\|_{q'} \right\|_{L^\infty} W_{p,q}(\mu,\nu),$$

for *a.e.* $t \in [0,1]$. Integrating between 0 and 1 implies

$$\left|E(0)^{1/p} - E(1)^{1/p}\right| \leq \left\| \|\nabla(u^{1/p})\|_{q'} \right\|_{L^\infty} W_{p,q}(\mu,\nu).$$

Then, noting that $E(0) = \int_{\mathbb{R}^d} u(x)\, \mu(dx)$ and $E(1) = \int_{\mathbb{R}^d} u(x)\, \nu(dx)$ ends the proof. □

As a consequence, we obtain the following result, which allows identifying the Lipschitz constant of a function with the Lipschitz constant of an associated nonlinear functional over $\mathcal{P}_p(\mathbb{R}^d)$.

Corollary A.2. *Consider any $1 \leq p \leq \infty$, set any $v \in C^1(\mathbb{R}^d)$ with $\nabla v \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$, and assume that $v > 0$ when $p < \infty$ but not necessarily when $p = \infty$. Define the functional $\Phi_{p,v} : \mathcal{P}_p(\mathbb{R}^d) \longrightarrow \mathbb{R}$ by*

$$\Phi_{p,v}[\mu] := \begin{cases} \left(\int_{\mathbb{R}^d} v(x)^p \mu(dx)\right)^{1/p}, & \text{if } p < \infty,\\ \log\left(\int_{\mathbb{R}^d} e^{v(x)} \mu(dx)\right), & \text{if } p = \infty,\end{cases}$$

for any $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Then, for any $1 \leq q \leq \infty$ the following identify holds true

$$\|\, \|\nabla v\|_{q'} \|_{L^\infty} = \sup_{\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)} \frac{\Phi_{p,v}[\mu] - \Phi_{p,v}[\nu]}{W_{p,q}(\mu,\nu)}.$$

Proof. First, note that the change of variables $v = u^{1/p}$ and Proposition A.1 readily implies

$$\|\, \|\nabla v\|_{q'} \|_{L^\infty} \geq \sup_{\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)} \frac{\Phi_{p,v}[\mu] - \Phi_{p,v}[\nu]}{W_{p,q}(\mu,\nu)}.$$

On the other hand, also note that by particularizing the measures $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ to be Dirac masses at respective points $x, x' \in \mathbb{R}^d$ we obtain

$$\sup_{\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)} \frac{\Phi_{p,v}[\mu] - \Phi_{p,v}[\nu]}{W_{p,q}(\mu,\nu)} \geq \sup_{x,x' \in \mathbb{R}^d} \frac{\Phi_{p,v}[\delta_x] - \Phi_{p,v}[\delta_{x'}]}{W_{p,q}(\delta_x,\delta_{x'})} = \sup_{x,x' \in \mathbb{R}^d} \frac{v(x) - v(x')}{\|x - x'\|_q} = \|\, \|\nabla v\|_{q'} \|_{L^\infty}.$$

This proves the converse inequality and then the above identity holds. □

Appendix B. Lower bound of Gaussian convolution of log-concave densities

We present a technical result which computes an explicit lower bound on the convolution of a Gaussian density and any strongly log-concave probability density.

Lemma B.1 (Lower bound I)**.** *Consider any $f = e^{-V} \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R})$, such that $V \in C^1(\mathbb{R})$ with $V'(0) = 0$, and f is γ-log-concave for some $\gamma > 0$. Then, we have*

$$(G * f)(x_0 + \delta) \geq G(2\delta)\, f(x_0 - \delta) \int_0^{\frac{\gamma}{\gamma+1}x_0 - \frac{\delta}{\gamma+1}} \exp\left(\frac{\gamma+1}{2} z^2\right) dz, \tag{B.1}$$

for any $\delta > 0$ and each $x_0 > \frac{\gamma+2}{\gamma}\delta$, where G denotes the standard Gaussian distribution (1.4)*.*

Proof. For simplicity of notation, we define $x_\pm := x_0 \pm \delta$ and we note that we can write

$$(G * f)(x_+) = \frac{1}{(2\pi)^{1/2}}\, f(x_-) \int_{\mathbb{R}} e^{V(x_-) - U(x)}\, dx, \tag{B.2}$$

where the function $U : \mathbb{R} \longrightarrow \mathbb{R}$ is defined by

$$U(x) := V(x) + \frac{1}{2}(x - x_+)^2, \quad x \in \mathbb{R}.$$

Since the potential V is γ convex, then we have that the potential U is $(\gamma+1)$-convex. By the convexity inequality applied to the pair of points (x, x_-) we then obtain

$$U(x_-) \geq U(x) + U'(x)(x_- - x) + \frac{\gamma+1}{2}(x_- - x)^2, \tag{B.3}$$

for any $x \in \mathbb{R}$. Consider the unique minimizer $x_* \in \mathbb{R}$ of the potential U. Since in particular x_* is a critical point of U, then we have

$$0 = U'(x_*) = V'(x_*) + (x_* - x_+).$$

Multiplying above by x^*, using that $V'(0) = 0$ by hypothesis along with the convexity inequality of V applied at the pair $(x_*, 0)$, we infer $\gamma\, x_*^2 \leq (x_+ - x_*)x_*$, and therefore,

$$|x_*| \leq \frac{1}{\gamma+1} x_+. \tag{B.4}$$

Since $U'(x) > 0$ for $x > x_*$ and $x_- - x > 0$ for $x < x_-$, then (B.3) implies

$$U(x_-) \geq U(x) + \frac{\gamma+1}{2}(x_- - x)^2,$$

for any $x \in (x_*, x_-)$. Let us note that indeed we have the appropriate ordering $x_* < x_-$ since by (B.4) and the assumption $x_0 > \frac{\gamma+2}{\gamma}\delta$ we obtain

$$x_* \leq \frac{1}{\gamma+1} x_+ = \frac{1}{\gamma+1}(x_0 + \delta) \leq x_0 - \delta = x_-.$$

Writing everything in terms of V implies

$$V(x_-) - U(x) \geq -\frac{1}{2}(x_- - x_+)^2 + \frac{\gamma+1}{2}(x_- - x)^2, \tag{B.5}$$

for any $x \in (x_*, x_-)$. Injecting (B.5) into (B.2) we obtain

$$(G * f)(x_+) \geq G(x_+ - x_-)\, f(x_-) \int_{x_*}^{x_-} \exp\left(\frac{\gamma+1}{2}(x_- - x)^2\right) dx.$$

Of course, the above implies (B.1) by a simple change of variables $z = x_- - x$, and noting again that

$$x_- - x_* \geq x_- - \frac{1}{\gamma+1} x_+ = (x_0 - \delta) - \frac{1}{\gamma+1}(x_0 + \delta) = \frac{\gamma}{\gamma+1} x_0 - \frac{\gamma+2}{\gamma+1}\delta,$$

thanks to (B.4), which yields again positive a positive upper bound by the assumption $x_0 > \frac{\gamma+2}{\gamma}\delta$. □

Note that arguing along the same lines, we can prove an analogous result where the above positive strongly log-concave density f is replaced by its truncation f_R to intervals $I_R := (-R, R)$. Specifically, anything that we need to guarantee is that $[x_*, x_-] \subset I_R$. First, note that $x_- < R$ amounts to the condition $x_0 < R + \delta$. Second, by (B.4) we obtain that $x_* > -R$ as long as $\frac{1}{\gamma+1} x_+ < R$, which amounts to the condition $x_0 < (\gamma+1)R - \delta$. If we take R large enough (namely $R > 2\delta/\gamma$) then we have that the former condition on x_0 is the most restrictive. Therefore, we have the following result.

Lemma B.2 (Lower bound II). *Under the assumptions in Lemma B.1, let us define*

$$f_R(x) := e^{-V_R(x)}, \qquad x\in\mathbb{R}$$
$$V_R(x) := V(x) + \chi_{\bar I_R}(x), \quad x\in\mathbb{R},$$

for any $R>0$, where $\chi_{\bar I_R}$ is the characteristic function associated to $\bar I_R$ (cf. (1.23)). Then, we have

$$(G*f_R)(x_0+\delta) \geq G(2\delta)\, f_R(x_0-\delta)\int_0^{\frac{\gamma}{\gamma+1}x_0-\frac{\delta}{\gamma+1}} \exp\left(\frac{\gamma+1}{2}z^2\right)dz, \tag{B.6}$$

for any $\delta>0$, each $\frac{\gamma+2}{\gamma}\delta < x_0 < R+\delta$, and every $R>\frac{2\delta}{\gamma}$.

Appendix C. Euclidean estimates on the displacement of the Brenier map between perturbations of log-concave measures.

In this section we present a proof of the the uniform bound of the ℓ_2 norm on the displacement of the Brenier map between perturbations of log-concave measures.

Lemma C.1. *Consider two densities $f,g\in L^1_+(\mathbb{R}^d)\cap\mathcal{P}(\mathbb{R}^d)$, assume that,*

$$\{z\in\mathbb{R}^d:\, f(z)>0\} = \{z\in\mathbb{R}^d:\, g(z)>0\} = \bar B_R,$$

where B_R is the Euclidean ball, and suppose that $f=e^{-W}, g=e^{-\tilde W}$ are γ-log-concave for some $\gamma>0$ and $f,g\in C^{1,\delta}(\bar B_R)$ for some $\delta>0$. Let $T=\nabla\phi:\bar B_R\longrightarrow\bar B_R$ be the Brenier map from f to g. Then,

$$W_{\infty,2}(f,g)\leq \|\|T-I\|_2\|_{L^\infty(\bar B_R)} \leq \frac{1}{\gamma}\|\|\nabla(W-\tilde W)\|_2\|_{L^\infty(\bar B_R)},$$

As mentioned in Remark 3.1, this result is not enough for the sake of this paper, but was the starting point to prove Lemma 2.6. The technique to prove it is essentially based on the computations in [22], but we provide the proof here since the statement is not a direct consequence of it. On the other hand, this very result has its own interest, as one can see from the recent paper [29].

Proof of Lemma C.1. Since $f,g\in C^{1,\delta}(\bar B_R)$ are bounded below on B_R by a positive constant, $f=g=0$ outside B_R, and B_R is uniformly convex, then the Caffarelli's theory [12] proves that $T\in C^{2,\delta}(\bar B_R)$. We consider $T(z)-z=\nabla\psi(z)$, where $\psi(z)=\phi(z)-\frac12\|z\|_2^2$. The function ψ solves the Monge-Ampere equation, that we write in logarithmic form:

$$\log\det(D^2\psi(z)+I) = \tilde W(\nabla\psi(z)+z) - W(z), \quad z\in\mathbb{R}^d. \tag{C.1}$$

Taking partial derivatives ∂_{x_k} in (C.1) we have

$$\operatorname{tr}\left((D^2\phi)^{-1}\partial_{x_k}D^2\psi\right) = \nabla\tilde W(\nabla\psi+z)\cdot\partial_{x_k}\nabla\psi + (\nabla\tilde W(\nabla\psi+z)-\nabla W)\cdot e_k, \quad z\in\mathbb{R}^d,$$

for $1\leq k\leq d$. We then multiply times $\partial_{x_k}\psi$, sum over k, so that we obtain

$$\operatorname{tr}\left((D^2\phi)^{-1}\sum_k \partial_{x_k}D^2\psi\,\partial_{x_k}\psi\right)$$
$$= \nabla\tilde W(\nabla\psi+z)\cdot\partial_{x_k}\left(\frac12\|\nabla\psi\|_2^2\right) + (\nabla\tilde W(\nabla\psi+z)-\nabla W)\cdot\nabla\psi(z), \quad z\in\mathbb{R}^d.$$

We now consider the point $z^*\in\bar B_R$ which maximizes $\frac12\|\nabla\psi\|_2^2$, which is also the maximum point for the displacement $\|T-I\|_2$. Such a point exist since the ball $\bar B_R$ is compact. Moreover, [22, Lemma 3.1] shows that such a maximum cannot be attained on the boundary ∂B_R. Hence, we can apply first and second-order optimality conditions. In particular, we have $\partial_{x_k}\left(\frac12\|\nabla\psi\|_2^2\right)(z^*)=0$ and the Hessian matrix $D^2\left(\frac12\|\nabla\psi\|_2^2\right)(z^*)$ has to be negative-definite, *i.e.*

$$\sum_k \partial_{x_k}D^2\psi(z^*)\partial_{x_k}\psi(z^*) + (D^2\psi(z^*))^2 \leq 0.$$

Using the fact that $(D^2\psi(z^*))^2$ is the square of a symmetric matrix, and hence is negative, we obtain that $\sum_k\partial_{x_k}D^2\psi(z^*)\partial_{x_k}\psi(z^*)$ is itself negative definite, and the trace of its product times $(D^2\phi)^{-1}$ is also negative. This allows to obtain

$$(\nabla\tilde W(\nabla\psi(z^*)+z^*) - \nabla W(z^*)\cdot\nabla\psi(z^*)\leq 0,$$

which implies

$$(\nabla W(\nabla\psi(z^*)+z^*) - \nabla W(z^*)\cdot\nabla\psi(z^*) \leq \|\nabla(\tilde W-W)\|_{L^\infty}\,\|\nabla\psi(z^*)\|_2,$$

and hence by γ-convexity of W we have

$$\gamma\,||\nabla\psi(z^*)||_2^2 \leq \|\nabla(\tilde{W}-W)\|_{L^\infty}\,||\nabla\psi(z^*)||_2,$$

which ends the proof. □

Similarly to Lemma 2.6 for the ℓ_1 norm of the displacement of the Brenier map, a more general result holds for strictly positive densities $f, g \in C^{1,\delta}_{\text{loc}}(\mathbb{R}^d)$ supported in the full space $\mathbb{R}^d$.

Corollary C.2. *Consider two densities $f, g \in L^1_+(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$, assume that $f, g > 0$, and suppose that $f = e^{-W}, g = e^{-\tilde{W}}$ are γ-log-concave for some $\gamma > 0$ and $f, g \in C^{1,\delta}_{\text{loc}}(\mathbb{R}^d)$ for some $\delta > 0$. Let $T = \nabla\phi : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be the Brenier map from f to g. Then,*

$$W_{\infty,2}(f,g) \leq \|\|T-I\|_2\|_{L^\infty(\mathbb{R}^d} \leq \frac{1}{\gamma}\|\|\nabla(W-\tilde{W})\|_2\|_{L^\infty(\mathbb{R}^d)}.$$

The proof is similar to the one of Lemma 2.6 arguing by a truncation argument and applying the local version in Lemma C.1. Specifically, we truncate W and $\tilde{W}$ and accordingly f and g to an increasing sequence B_R of Euclidean balls preserving the Lipschitz and convexity bounds. We obtain a sequence of optimal transport maps T_R transporting the associated truncations f_R onto g_R and satisfying

$$\|\,||T_R - I||_2\|_{L^\infty(\bar{B}_R)} \leq \frac{1}{\gamma}\|\nabla(W-\tilde{W})\|_{L^\infty(\mathbb{R}^d)},$$

for all $R > 0$. Finally, we pass to the limit in the above estimate as $R \to \infty$.

References

1. L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Birkhäuser, Basel, 2008.
2. A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter, *On Convex Sobolev Inequalities and the Rate of Convergence to Equilibrium for Fokker-Planck Type Equations*, Commun. Partial. Differ. Equ. **8** (26), no. 1-2, 43–100.
3. D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, Lectures on Probability Theory: Ecole d'Eté de Probabilités de Saint-Flour XXII-1992 (P. Bernard, ed.), Lecture Notes in Mathematics, vol. 1581, Springer, Berlin, Heidelberg, 1994, pp. 1–114.
4. D. Bakry, I. Gentil, and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, A Series of Comprehensive Studies in Mathematics, vol. 348, Springer, Cham, 2014.
5. G. Barles, S. Mirrahimi, and B. Perthame, *Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result*, Methods and Applications of Analysis **16** (2009), no. 3, 321–340.
6. N. H. Barton, A. M. Etheridge, and A. Véber, *The infinitesimal model: Definition, derivation, and implications*, Theoret. Population Biol. **118** (2017), 50–73.
7. H. Berestycki, J. Coville, and H.-H. Vo, *Persistence criteria for populations with non-local dispersion*, J. Math. Biol. **72** (2016), 1693–1745.
8. H. J. Brascamp and E. 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** (1991), 366–389.
9. Y Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417.
10. L. A. Caffarelli, *Boundary regularity of maps with convex potentials*, Comm. Pure Appl. Math. **45** (1992), no. 9, 1141–1151.
11. ______, *The regularity of mappings with a convex potential*, J. Amer. Math. Soc. **5** (1992), no. 1, 99–104.
12. ______, *Boundary regularity of maps with convex potentials II*, Comm. Pure Appl. Math. **45** (1996), no. 9, 1141–1151.
13. ______, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys. **214** (2000), no. 3, 547–563.
14. ______, *Erratum: "Monotonicity properties of optimal transportation and the FKG and related inequalities" [Comm. Math. Phys. 214 (2000), no. 3, 547-563]*, Comm. Math. Phys. **225** (2002), no. 2, 449–450.
15. V. Calvez, J. Garnier, and F. Patout, *Asymptotic analysis of a quantitative genetics model with nonlinear integral operator*, J. Éc. Polytech. Math. **6** (2019), 537–579.
16. V. Calvez, L. Lepoutre, and D. Poyato, *Ergodicity of the Fisher infinitesimal model with quadratic selection*, Nonlinear Anal. Theory Methods Appl. **238** (2024), 113392.
17. M. Colombo and M. Fathi, *Bounds on optimal transport maps onto log-concave measures*, J. Differ. Equ. **271** (2021), 1007–1022.
18. M. Colombo, A. Figalli, and Y. Jhaveri, *Lipschitz changes of variables between perturbations of log-concave measures*, Ann. Sc. Norm. Super. Pisa Cl. Sci. **17** (2017), 1491–1519.
19. J. Coville, F. Li, and X. Wang, *On eigenvalue problems arising from nonlocal diffusion models*, Discrete Contin. Dyn. Syst. Ser. A **37** (2017), no. 2, 879–903.
20. L. Dekens, *Evolutionary dynamics of complex traits in sexual populations in a heterogeneous environment: how normal?*, J. Math. Biol. **84** (2022), no. 3, 15.
21. O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame, *The dynamics of adaptation: An illuminating example and a Hamilton–Jacobi approach*, Theor. Popul. Biol. **67** (2005), no. 4, 257–271.

22. V. Ferrari and F. Santambrogio, *Lipschitz estimates on the JKO scheme for the Fokker–Planck equation on bounded convex domains*, Appl. Math. Lett. **112** (2021), 106806.
23. R. A. Fisher, *The correlation between relatives on the supposition of mendelian inheritance*, Trans. Roy. Soc. Edinburgh **52** (1918), 399–433.
24. ———, *On the mathematical foundations of theoretical statistics*, Philos. Trans. Royal Soc. A **222** (1922), no. 594-604, 309–368.
25. A. Frouvelle and C. Taing, *On the fisher infinitesimal model without variability*, 2023.
26. J. Garnier, O. Cotto, E. Bouin, T. Bourgeron, T. Lepoutre, O. Ronce, and V. Calvez, *Adaptation of a quantitative trait to a changing environment: New analytical insights on the asexual and infinitesimal sexual models*, Theor. Popul. Biol. **152** (2023), 1–22.
27. J Guerand, M Hillairet, and S Mirrahimi, *A moment-based approach for the analysis of the infinitesimal model in the regime of small variance*, 2023.
28. Y. Jhaveri, *On the (in)stability of the identity map in optimal transportation*, Calc. Var. Partial Differ. Equ. **58** (2019), 96.
29. K. A. Khudiakova, J. Maas, and F. Pedrotti, *L^∞-optimal transport of anisotropic log-concave measures and exponential convergence in Fisher's infinitesimal model*, 2024, arXiv:2402.04151.
30. K. Kuwada, *Duality on gradient estimates and Wasserstein controls*, J. Funct. Anal. **258** (2010), no. 11, 3758–3774.
31. R. Mahadevan, *A note on a non-linear Krein-Rutman theorem*, Nonlinear Anal. Theory Methods Appl. **67** (2007), no. 11, 3084–3090.
32. S. Mirrahimi and G. Raoul, *Dynamics of sexual populations structured by a space variable and a phenotypical trait*, Theoret. Population Biol. **84** (2013), 87–103.
33. R. D. Nussbaum, *Hilbert's Projective Metric and Iterated Nonlinear Maps*, American Mathematical Society, Basel, 1099.
34. ———, *Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations*, Differ. Integral Equ. **7** (1994), no. 5-6, 1649–1707.
35. Y. Ollivier, *Ricci curvature of metric spaces*, C. R. Acad. Sci. Paris, Ser. I **345** (2007), no. 11, 643–646.
36. ———, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), no. 3, 810–864.
37. F. Patout, *The Cauchy problem for the infinitesimal model in the regime of small variance*, Anal. PDE **16** (2023), no. 6, 1289–1350.
38. B. Perthame and G. Barles, *Dirac concentrations in Lotka-Volterra parabolic PDEs*, Indiana Univ. Math. J. **57** (2008), no. 7, 3275–3301.
39. G. Raoul, *Macroscopic limit from a structured population model to the Kirkpatrick-Barton model*, 2017, arXiv:1706.04094.
40. ———, *Exponential convergence to a steady-state for a population genetics model with sexual reproduction and selection*, 2021, arXiv:2104.06089.
41. A. Saumard and J. A. Wellner, *Log-concavity and strong log-concavity: A review*, Statist. Surv. **8** (2014), no. 45, 45–114.
42. S. Stigler, *Fisher in 1921*, Stat. Sci. **20** (2005), no. 1, 32–49.
43. C. Villani, *Topics in optimal transportation*, American Mathematical Society, Providence, RI, 2003.
44. M.-K. Von Renesse and K.-T. Sturm, *Transport inequalities, gradient estimates, entropy and Ricci curvature*, Commun. Pure Appl. Math. **58** (2005), no. 7, 923–940.
45. G. F. Webb, *An operator-theoretic formulation of asynchronous exponential growth*, Trans. Amer. Math. Soc. **303** (1987), no. 2, 751–763.

CNRS, Univ Brest, UMR 6205, Laboratoire de Mathématiques de Bretagne Atlantique, France
Email address: vincent.calvez@math.cnrs.fr

Departamento de Matemática Aplicada and Research Unit "Modeling Nature" (MNat), Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain
Email address: davidpoyato@ugr.es

Universite Claude Bernard Lyon 1, CNRS, Ecole Centrale de Lyon, INSA Lyon, Université Jean Monnet, Institut Camille Jordan UMR5208, 43 bd du 11 novembre 1918, 69622 Villeurbanne Cedex, France.
Email address: santambrogio@math.univ-lyon1.fr