THE WASSERSTEIN GRADIENT FLOW OF THE FISHER INFORMATION AND THE QUANTUM DRIFT-DIFFUSION EQUATION

UGO GIANAZZA, GIUSEPPE SAVARÉ, AND GIUSEPPE TOSCANI

Abstract. We prove the global existence of nonnegative variational solutions to the "drift diffusion" evolution equation

$$\partial_t u + \operatorname{div}\left(u\mathrm{D}\Big(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\Big)\right) = 0$$

under variational boundary condition. Despite the lack of a maximum principle for fourth order equations, nonnegative solutions can be obtained as a limit of a variational approximation scheme by exploiting the particular structure of this equation, which is the gradient flow of the (perturbed) *Fisher Information* functional

$$\mathscr{F}^f(u) := \frac{1}{2}\int \big|\mathrm{D}\log u\big|^2\, u\,\mathrm{d}x + \int f u\,\mathrm{d}x$$

with respect to the Kantorovich-Rubinstein-Wasserstein distance between probability measures.

We also study long time behaviour of the solutions, proving their exponential decay to the equilibrium state $g = e^{-V}$ characterized by

$$-\Delta V + \frac{1}{2}\big|\mathrm{D}V\big|^2 = f, \qquad \int e^{-V}\,\mathrm{d}x = \int u_0\,\mathrm{d}x,$$

when the potential V is uniformly convex: in this case the functional $\mathscr{F}^f$ coincides with the *Relative Fisher Information*

$$\mathscr{F}^f(u) = \frac{1}{2}\mathscr{I}(u|g) = \int \big|\mathrm{D}\log(u/g)\big|^2\, u\,\mathrm{d}x.$$

Contents

Key words and phrases. Fisher information, Entropy, quantum drift diffusion, Wasserstein distance, log-concave measures, nonnegative solutions, fourth order evolution equations.

This work was partially supported by grants of M.U.R.S.T. and by the I.M.A.T.I.-C.N.R., Pavia, Italy.

1. Introduction and main result

The aim of this paper is to study global existence in time for *nonnegative* solutions u of the fourth order "Quantum drift diffusion" equation

$$\partial_t u + \operatorname{div}\left(u\,\mathrm{D}\Big(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\Big)\right) = 0 \quad \text{in } \Omega_T := \Omega\times(0,T), \tag{EE$_1$}$$

subject to the initial Cauchy condition

$$u(x,0) = u_0(x) \geq 0 \quad \text{in } \Omega, \quad \text{with } u_0 \in L^1(\Omega), \quad \int_\Omega |x|^2 u_0(x)\,\mathrm{d}x < +\infty, \tag{1.1}$$

and to the variational boundary conditions

$$\partial_{\boldsymbol{n}} u = 0 \quad \text{on } (\partial\Omega)_T := \partial\Omega\times(0,T), \tag{1.2}$$

$$u\partial_{\boldsymbol{n}}\left(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\right) = 0 \quad \text{on } (\partial\Omega)_T. \tag{1.3}$$

Here $\partial_t := \frac{\partial}{\partial t}$, D denotes the gradient with respect to (w.r.t.) spatial variables, Ω is an open *convex* (possibly unbounded) domain of $\mathbb{R}^d$ with exterior unit normal $\boldsymbol{n}$, $T \in (0,+\infty]$ is the final time, f is a given perturbation term, which satisfies

$$f \in C^1(\Omega), \quad C_f := \sup_{x\in\Omega}\left(\frac{|f(x)|}{1+|x|^2} + \frac{|\mathrm{D}f(x)|}{1+|x|}\right) < +\infty, \quad \liminf_{|x|\to\infty}\frac{f(x)}{|x|^2} \geq 0. \tag{1.4}$$

1.1. **An important case.**

A particularly important case concerns functions f of the type

$$f = 2\frac{\Delta\sqrt{g}}{\sqrt{g}} = \frac{1}{2}\big|\mathrm{D}V\big|^2 - \Delta V, \tag{1.5a}$$

where g is a *strictly positive* function which is induced by a λ-*convex* potential $V : \mathbb{R}^d \to \mathbb{R}$ through the formula

$$g(x) := Z^{-1}e^{-V(x)}, \qquad V \in C^2(\mathbb{R}^d), \quad \mathrm{D}^2V(x) \geq \lambda\,\mathsf{Id} \quad \forall\, x\in\mathbb{R}^d, \quad \inf_{x\in\Omega} V(x) > -\infty, \tag{1.5b}$$

where the normalization constant $Z > 0$ (and the initial datum u_0) have been chosen so that

$$\int_\Omega g(x)\,\mathrm{d}x = \int_\Omega u_0(x)\,\mathrm{d}x = Z^{-1}\int_\Omega e^{-V(x)}\,\mathrm{d}x = 1. \tag{1.5c}$$

We also assume that the potential V satisfies

$$\partial_{\boldsymbol{n}} V = 0 \quad \text{on } \partial\Omega, \tag{1.5d}$$

so that g represents the density of an *invariant probability measure* $\gamma := g\mathcal{L}^d$ associated to the equation (EE$_1$); besides (1.5b), we will also often ask for the further technical assumption

$$\sup_{x\in\Omega}\big\|\mathrm{D}^2V(x)\big\| < +\infty. \tag{1.5e}$$

When V is convex (i.e. $\lambda \geq 0$ in (1.5b)) γ is a *log-concave* measure. Let us observe that the equation is invariant with respect to addition of constants to f or to V, and to multiplication of g by nonnegative constants; in particular the normalization condition (1.5c) is not restrictive, at least when γ is a finite measure. Particularly interesting cases are

$$\begin{gathered}\partial_t u + 2\operatorname{div}\left(u\,\mathrm{D}\frac{\Delta\sqrt{u}}{\sqrt{u}}\right) = 0 \qquad \text{for} \quad f = V \equiv 0,\\ \text{with} \quad g \equiv Z^{-1}, \quad Z := \mathcal{L}^d(\Omega), \quad \gamma = Z^{-1}\mathcal{L}^d, \text{ the (normalized) Lebesgue measure,}\end{gathered} \tag{1.6}$$

which has been introduced by [DLSS91] and studied in a bounded domain Ω by [BLS94] and subsequently by [JP00], and its (suitably) rescaled version in the whole space $\mathbb{R}^d$ [CT02]

$$\partial_t u + \operatorname{div}\left(u\left(2\mathrm{D}\frac{\Delta\sqrt{u}}{\sqrt{u}} - \lambda^2 x\right)\right) = 0, \tag{1.7}$$

which corresponds to

$$V(x)=\frac{\lambda|x|^2}{2},\quad Z:=(2\pi\lambda^{-1})^{d/2},\quad g(x)=\left(\tfrac{\lambda}{2\pi}\right)^{d/2}e^{-\frac{\lambda}{2}|x|^2},\quad f(x)=\frac{\lambda^2}{2}|x|^2-d\,\lambda \tag{1.8}$$

for some $\lambda>0$; in this case γ is the centered Gaussian measure with variance λ^{-1}.

1.2. **Other forms of the equation.** Before discussing the precise meaning of the equation and the related notion of solutions, let us recall that equation (EE_1) has a rich structure and, at least for regular positive solutions, other equivalent forms.

Neglecting the boundary conditions, it is not difficult to recognize that (EE_1) is (formally) equivalent to

$$\partial_t u+\sum_{i,j=1}^d\partial^2_{ij}\Big(u\,\partial^2_{ij}\log u\Big)-\sum_{i=1}^d\partial_i\Big(u\,\partial_i f\Big)=0\quad\text{in }\Omega_T, \tag{EE_2}$$

and to

$$\partial_t u+\Delta^2 v-\sum_{i,j=1}^d\partial^2_{ij}\left(\frac{\partial_i u\,\partial_j u}{u}\right)-\sum_{i=1}^d\partial_i\Big(u\,\partial_i f\Big)=0\quad\text{in }\Omega_T, \tag{EE_3}$$

where we set $\partial_i:=\frac{\partial}{\partial x_i}$, $\partial^2_{ij}=\frac{\partial^2}{\partial x_i\partial x_j}$. We will present in Theorem 5.9 the calculations which show the equivalence of (EE_1), (EE_2), and (EE_3), under suitable regularity assumptions on u and its square root $r:=\sqrt{u}$.

In the case $f\equiv 0$, formulation (EE_3) was exploited by [BLS94] in order to show local existence in time, uniqueness, and regularity of a mild solution for strictly positive and sufficiently regular initial data u_0 and periodic boundary conditions. [BLS94] considers the singular part of the equation as a perturbation term of the linear fourth order equation governed by the biharmonic operator. Techniques related to analytic semigroups are then involved and a connection between the positivity-preservation and the global existence in time is provided, although neither of them is proved except for an indication of the positivity exhibited by numerical experiments.

1.3. **Changing the reference measure.** In the case f is associated to a potential V as in (1.5a,b,c,d), these equations can be rewritten in more expressive forms. First of all, let us remark that, being g a stationary solution of (EE_1), the relative density of u with respect to g (or, better, the density of the measure $\mu:=u\cdot\mathcal{L}^d$ with respect to γ) and its square root play an important role in many formulae: we will systematically employ the short notation

$$v:=u/g,\quad s=\sqrt{v}=\sqrt{u/g},\quad u=vg=s^2g,\quad \mu=u\cdot\mathcal{L}^d=v\cdot\gamma, \tag{1.9}$$

and we will introduce the partial differential operators $\tilde\partial_i,\mathrm{div}_\gamma,\tilde\partial^2_{ij},\Delta_\gamma$, which (for sufficiently regular functions) are defined by

$$\begin{gathered}\tilde\partial_i v:=g^{-1}\partial_i(gv)=\partial_i v-v\,\partial_i V,\quad \mathrm{div}_\gamma\boldsymbol{v}:=\sum_i\tilde\partial_i v_i=g^{-1}\,\mathrm{div}(g\boldsymbol{v})=\mathrm{div}\,\boldsymbol{v}-\boldsymbol{v}\cdot DV,\\ \tilde\partial^2_{ij}v:=g^{-1}\,\partial^2_{ij}(gv)=\tilde\partial_i\tilde\partial_j v=\tilde\partial_j\tilde\partial_i v,\quad \Delta_\gamma v:=\mathrm{div}_\gamma(\mathrm{D}v)=\Delta v-\mathrm{D}v\cdot DV.\end{gathered} \tag{1.10}$$

They satisfy the "integration by parts formulae" against test functions $\zeta\in C^\infty_c(\Omega)$ with respect to the measure γ

$$\begin{gathered}\int_\Omega v\,\partial_i\zeta\,\mathrm{d}\gamma=-\int_\Omega\tilde\partial_i v\,\zeta\,\mathrm{d}\gamma,\quad \int_\Omega\boldsymbol{v}\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma=-\int_\Omega\mathrm{div}_\gamma\boldsymbol{v}\,\zeta\,\mathrm{d}\gamma,\\ \int_\Omega v\,\partial^2_{ij}\zeta\,\mathrm{d}\gamma=\int_\Omega\tilde\partial^2_{ij}v\,\zeta\,\mathrm{d}\gamma,\quad \int_\Omega\mathrm{D}v\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma=-\int_\Omega\Delta_\gamma v\,\zeta\,\mathrm{d}\gamma.\end{gathered} \tag{1.11}$$

Of course, in the case (1.6) when $g\equiv Z^{-1}$, all these operators coincide with the usual ones and the sub/superscripts $_\gamma$, $\tilde{}$ can be removed in each formula; in the case of (1.8) for $\lambda=1$, Δ_γ is the *Ornstein-Uhlenbeck* operator $\Delta-x\cdot\mathrm{D}$.

One of the main advantages of this point of view relies on a simplification in the nonlinear term inside (EE_1), which can be rewritten as

$$2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f \overset{(1.9)}{=} 2\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}. \tag{1.12}$$

In fact, (1.5a) yields up to the factor 2

$$\frac{\Delta\sqrt{u}}{\sqrt{u}} - \frac{\Delta\sqrt{g}}{\sqrt{g}} = \frac{1}{s\sqrt{g}}\Big(\Delta(s\sqrt{g}) - s\Delta\sqrt{g}\Big) = \frac{1}{s}\Big(\Delta s - \mathrm{D}V\cdot\mathrm{D}s\Big) = \frac{\Delta_\gamma s}{s} = \frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}, \tag{1.13}$$
$$\text{where}\quad s = \sqrt{v} = \sqrt{u/g}.$$

Taking into account (1.13) and (1.5d), we obtain the formulation of (EE_1) in terms of the density $v = u/g$:

$$\begin{cases} \partial_t v + 2\,\mathrm{div}_\gamma\Big(v\mathrm{D}\Big(\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}\Big)\Big) = 0 \quad \text{in } \Omega_T, \\ \partial_{\boldsymbol{n}} v = v\partial_{\boldsymbol{n}}\Big(\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}\Big) = 0 \quad \text{on } (\partial\Omega)_T. \end{cases} \tag{$\mathrm{EE}_{1,\gamma}$}$$

Neglecting the boundary conditions, one finds that ($\mathrm{EE}_{1,\gamma}$) is (formally) equivalent to

$$\partial_t v + \sum_{i,j=1}^d \tilde\partial^2_{ij}\Big(v\,\partial^2_{ij}\log v\Big) - \sum_{i=1}^d \tilde\partial_i\Big(v\,\partial^2_{ij}V\,\partial_j\log v\Big) = 0 \quad \text{in } \Omega_T, \tag{$\mathrm{EE}_{2,\gamma}$}$$

and to

$$\partial_t v + \Delta^2_\gamma v - \sum_{i,j=1}^d \tilde\partial^2_{ij}\left(\frac{\partial_i v\,\partial_j v}{v}\right) = 0 \quad \text{in } \Omega_T. \tag{$\mathrm{EE}_{3,\gamma}$}$$

Remark 1.1. All these equations can also be studied for a more general potential V satisfying only (1.5a,b,c) but not (1.5d): if, e.g., V has a super quadratic growth or it does not satisfies the homogeneous Neumann boundary condition on $\partial\Omega$ or ΔV does not belong to $C^1(\Omega)$, then the link with formulations (EE_1), (EE_2), (EE_3) is no more available but still ($\mathrm{EE}_{1,\gamma}$), ($\mathrm{EE}_{2,\gamma}$), ($\mathrm{EE}_{3,\gamma}$) make sense.

1.4. **Lyapunov functionals.** One of the main contributions of [BLS94] consists in the identification of a certain number of Lyapunov functionals, which are decreasing along regular positive solutions of (EE_1) in the case (1.6) of $f \equiv 0,\ g \equiv 1$.

Considering here the general case, *the formulation* ($\mathrm{EE}_{1,\gamma}$) through the invariant measure γ suggests the correct way to adapt these functionals in this wider setting. In particular, two of them seem to play a crucial role. The first one is the γ-Relative Entropy of $\mu_t := v_t(\cdot)\gamma = u_t(\cdot)\mathcal{L}^d$

$$\mathscr{H}(\mu_t|\,\gamma) := \int_\Omega u_t(x)\big(\log u_t(x) + V(x)\big)\,\mathrm{d}x = \int_\Omega v_t(x)\log v_t(x)\,\mathrm{d}\gamma(x), \tag{1.14}$$

which formally satisfies

$$\mathscr{H}(\mu_0|\,\gamma) - \mathscr{H}(\mu_T|\,\gamma) = \iint_{\Omega_T}\Big(\big|\mathrm{D}^2\log v_t\big|^2 + \mathrm{D}^2V\,\mathrm{D}\log v_t\cdot\mathrm{D}\log v_t\Big)\,v_t\,\mathrm{d}\gamma(x)\,\mathrm{d}t; \tag{1.15}$$

here D^2h denotes the (symmetric) matrix of the second order partial derivatives $\partial^2_{ij}h$ of the function h and

$$\big|\mathrm{D}^2h\big|^2 = \sum_{i,j=1}^d\big(\partial^2_{ij}h\big)^2, \quad \mathrm{D}^2h\,\mathrm{D}h_1\cdot\mathrm{D}h_2 = \sum_{i,j=1}^d \partial^2_{ij}h\,\partial_i h_1\partial_j h_2. \tag{1.16}$$

The second one is the γ-*Relative Fisher information*

$$\mathscr{I}_2(\mu_t|\,\gamma) := \int_\Omega |\mathrm{D}\log v_t|^2\, v_t\,\mathrm{d}\gamma(x) = 4\int_\Omega \big|\mathrm{D}\sqrt{v_t}\big|^2\,\mathrm{d}\gamma(x) \tag{1.17a}$$

$$= \int_\Omega \frac{|\mathrm{D}u_t + u_t\mathrm{D}V|^2}{u_t}\,\mathrm{d}x \tag{1.17b}$$

$$\overset{(1.5\text{a,b,c,d,e})}{=} \int_\Omega \big|\mathrm{D}\log u_t\big|^2 u_t\,\mathrm{d}x + 2\int_\Omega f u_t\,\mathrm{d}x = 4\int_\Omega \big|\mathrm{D}\sqrt{u_t}\big|^2\,\mathrm{d}x + 2\int_\Omega f u_t\,\mathrm{d}x, \tag{1.17c}$$

which satisfies

$$\frac{1}{2}\mathscr{I}_2(\mu_T|\,\gamma) + \int_0^T\int_\Omega \Big|2\mathrm{D}\Big(\frac{\Delta_\gamma\sqrt{v_t}}{\sqrt{v_t}}\Big)\Big|^2\, v_t\,\mathrm{d}\gamma(x)\,\mathrm{d}t = \frac{1}{2}\mathscr{I}_2(\mu_0|\,\gamma). \tag{1.18}$$

Remark 1.2. (1.17c) holds only if f and V are linked by (1.5a,b,c,d,e); in this case (see also the next Remark 2.5) the γ-Relative Fisher Information is in fact a linear perturbation of the *Lebesgue*-Relative Fisher Information, i.e.

$$\mathscr{I}_2(\mu_t|\,\gamma) = \mathscr{I}_2(\mu_t|\,\mathcal{L}^d) + 2\langle f,\mu_t\rangle, \quad \langle f,\mu_t\rangle := \int_\Omega f(x)\,\mathrm{d}\mu_t(x); \tag{1.19}$$

This identity, related to the change of reference measure, has also been used by [CL04] to investigate the relationships between Logarithmic Sobolev and Poincaré inequalities.

If (1.19) does not hold, one can still consider the functional

$$\mathscr{F}^f(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \langle f,\mu\rangle = 2\int_\Omega \big|\mathrm{D}\sqrt{u}\big|^2\,\mathrm{d}x + \int_\Omega f u\,\mathrm{d}x, \quad \mu = u\cdot\mathcal{L}^d, \tag{1.20}$$

which is a Lyapunov functional for (EE_1), since it satisfies

$$\mathscr{F}^f(\mu_t) + \int_0^T\int_\Omega \Big|\mathrm{D}\Big(2\frac{\Delta_\gamma\sqrt{u_t}}{\sqrt{u_t}} - f\Big)\Big|^2\, u_t\,\mathrm{d}\gamma\,\mathrm{d}t = \mathscr{F}^f(\mu_0). \tag{1.21}$$

Besides the possibility to get crucial *a priori* estimates, these functionals have also been extremely useful to study asymptotic behavior of the solutions (as in [CCT05] and [DGJ06]), since convergence to the constant steady state can be derived by studying the precise rate of decay of Lyapunov functionals. It is interesting to remark that, for strictly positive solutions, the proof of many results took advantage of the rewriting of the equation in terms of $\log u$ as in (EE_2). The new form of the equation is indeed suitable to recover the all the Lyapunov functionals (as in the Entropy dissipation formula (1.14)) which are decreasing along regular positive solutions in the case (1.6) $f\equiv 0,\ g\equiv 1$.

In [JP00] the point of view of the other formulation (EE_2) is preferred and a global existence in time for a suitable notion of weak solution is proved in the one dimensional case with Dirichlet-Neumann boundary conditions. Since H^1-Sobolev imbeddings play a crucial role, it is far from obvious how to extend the ideas and the result of [JP00] to higher dimensions or unbounded domains.

The 1–dimensional case with periodic or inhomogeneous boundary conditions have also been recently considered in [DGJ06] and [GJT06] respectively; the large–time behavior of the solution to the initial–boundary value problem has been studied in [JT03] and [DGJ06].

1.5. **The link with the Wasserstein distance and the Fisher Information.** Following a remark by Y. Brenier, in our approach (EE_1) and ($\mathrm{EE}_{1,\gamma}$) provide the most interesting form since they are strictly related to the Lyapunov estimate (1.18) and to a suitable variational formulation involving the so called "Wasserstein distance" between probability measures and the Fisher Information functional. We shall also see that it suggests a variational formulation which allows to recover the other two equations ($\mathrm{EE}_{2,3}$), ($\mathrm{EE}_{2,3,\gamma}$).

We could summarize this point of view by saying that

$$
\begin{gathered}
(\mathrm{EE}_1)\text{ and }(\mathrm{EE}_{1,\gamma})\text{ are the \textit{Gradient flow} of the functionals}\\
\mathscr{F}^f(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \langle f,\mu\rangle, \qquad \mathscr{G}(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma)\\
\text{with respect to the \textit{Wasserstein distance.}}
\end{gathered}
\tag{1.22}
$$

As we have already observed in Remarks 1.1 and 1.2, these functionals (and the corresponding gradient flows) coincide when f and V are linked by (1.5a,b,c,d,e). Let us first provide a formal justification of (1.22).

First of all observe that, at least for strictly positive regular functions, (the opposite of) the interior expression in (EE_1)

$$
\begin{gathered}
-\mathfrak{A}u := -\Big(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\Big) \quad \text{is the Euler-Lagrange first variation}\\
\text{of the functional} \quad u \mapsto \mathscr{F}^f(u\,\mathcal{L}^d) := \frac{1}{2}\mathscr{I}_2(u\mathcal{L}^d|\,\mathcal{L}^d) + \langle f, u\mathcal{L}^d\rangle \quad \text{defined in (1.17).}
\end{gathered}
\tag{1.23}
$$

$\mathscr{F}^f(\cdot\,\mathcal{L}^d)$ can indeed be written as an integral functional

$$
\mathscr{F}^f(u\,\mathcal{L}^d) = \int_\Omega L^f(x,u,\mathrm{D}u)\,\mathrm{d}x, \quad \text{with} \quad L^f(x,z,\boldsymbol{p}) := \frac{1}{2}\frac{|\boldsymbol{p}|^2}{z} + f(x)z \quad \text{for } z\in(0,+\infty),\ \boldsymbol{p}\in\mathbb{R}^d,
$$

and we easily find

$$
\begin{aligned}
\frac{\delta\mathscr{F}^f(u\,\mathcal{L}^d)}{\delta u} &= L^f_z(x,u,\mathrm{D}u) - \sum_i \partial_i L^f_{p_i}(x,u,\mathrm{D}u) = f(x) - \frac{|\mathrm{D}u|^2}{2u^2} - \sum_i \partial_i\Big(\frac{\partial_i u}{u}\Big)\\
&= f(x) - \frac{|\mathrm{D}u|^2}{2u^2} + \frac{|\mathrm{D}u|^2}{u^2} - \frac{\Delta u}{u} = f - 2\Big(\frac{\Delta u}{2u} - \frac{|\mathrm{D}u|^2}{4u^2}\Big) = f - 2\frac{\Delta\sqrt{u}}{\sqrt{u}}.
\end{aligned}
\tag{1.24}
$$

Therefore (EE_1) can be rewritten as a system of three equations

$$
\partial_t u + \operatorname{div}(u\boldsymbol{v}) = 0 \qquad \text{in } \Omega_T, \quad u\boldsymbol{v}\cdot\boldsymbol{n} = 0 \qquad \text{on } (\partial\Omega)_T,
\tag{1.25a}
$$

$$
\boldsymbol{v} = \mathrm{D}\psi \qquad \text{in } \Omega_T,
\tag{1.25b}
$$

$$
\psi = -\frac{\delta\mathscr{F}^f(u\,\mathcal{L}^d)}{\delta u} \qquad \text{in } \Omega_T, \quad \sum_i L^f_{p_i}(x,u,\mathrm{D}u)\boldsymbol{n}_i = 0 \qquad \text{on } (\partial\Omega)_T,
\tag{1.25c}
$$

which exhibits the typical structure of a Wasserstein Gradient flow (see [AGS05, Chap. XI]).

It is interesting that also $(\mathrm{EE}_{1,\gamma})$ can be recovered in a completely similar way. In this case, as we have already mentioned before, it is more convenient to write the integrals in terms of the reference measure γ and of the density $v = u/g$ of μ with respect to γ. Therefore we consider the functional

$$
v \mapsto \mathscr{G}(v\,\gamma) = \frac{1}{2}\mathscr{I}_2(v\gamma|\,\gamma) = \int_\Omega L(v,\mathrm{D}v)\,\mathrm{d}\gamma(x),
$$

relative to the Lagrangian

$$
L(z,\boldsymbol{p}) := \frac{1}{2}\frac{|\boldsymbol{p}|^2}{z} \quad \text{for } z\in(0,+\infty),\ \boldsymbol{p}\in\mathbb{R}^d;
$$

again, an easy calculation of the first variation of $\mathscr{G}(\cdot\,\gamma)$ (where integration by parts are expressed in terms of γ and therefore $\tilde\partial_i$ differentiation (1.10) should be used) yields

$$
\begin{aligned}
\frac{\tilde\delta\mathscr{G}(v\,\gamma)}{\delta v} &= L_z(v,\mathrm{D}v) - \sum_i \tilde\partial_i L_{p_i}(v,\mathrm{D}v) = -\frac{|\mathrm{D}v|^2}{2v^2} - \sum_i \tilde\partial_i\Big(\frac{\partial_i v}{v}\Big)\\
&= -\frac{|\mathrm{D}v|^2}{2v^2} + \frac{|\mathrm{D}v|^2}{v^2} - \frac{\Delta_\gamma v}{v} = -2\Big(\frac{\Delta_\gamma v}{2v} - \frac{|\mathrm{D}v|^2}{4v^2}\Big) = -2\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}.
\end{aligned}
\tag{1.26}
$$

Therefore ($\mathrm{EE}_{1,\gamma}$) can be rewritten as a system of three equations

$$\partial_t v + \mathrm{div}_\gamma(v\,\boldsymbol{v}) = 0 \qquad \text{in } \Omega_T, \quad v\,\boldsymbol{v}\cdot\boldsymbol{n} = 0 \qquad \text{on } (\partial\Omega)_T, \tag{1.27a}$$

$$\boldsymbol{v} = \mathrm{D}\psi \qquad \text{in } \Omega_T, \tag{1.27b}$$

$$\psi = -\frac{\tilde\delta\mathscr{G}(v\,\gamma)}{\delta v} \qquad \text{in } \Omega_T, \quad \sum_i L_{p_i}(v,\mathrm{D}v)\boldsymbol{n}_i = 0 \qquad \text{on } (\partial\Omega)_T. \tag{1.27c}$$

The divergence structure of the continuity equation (1.27a) and the boundary condition show

$$\frac{d}{dt}\int_\Omega v_t(x)\,\mathrm{d}\gamma(x) = 0, \quad \text{i.e.} \quad \int_\Omega v_t(x)\,\mathrm{d}\gamma(x) = \int_\Omega v_0(x)\,\mathrm{d}\gamma(x). \tag{1.28}$$

1.6. **Densities and measures.** Since the equation is homogeneous of degree one, it is not restrictive to assume (as in (1.5c))

$$\int_\Omega u_0(x)\,\mathrm{d}x = \int_\Omega v_0(x)\,\mathrm{d}\gamma(x) = 1, \tag{1.29}$$

and therefore to identify u_t and v_t with the Borel probability measure

$$\mu_t := v_t\cdot\gamma = u_t\cdot\mathscr{L}^d. \tag{1.30}$$

Thus the continuity equations (1.25a)–(1.27a) simply state that $\boldsymbol{v}_t$ is a "velocity vector" of μ_t, (1.25b)–(1.27b) show that $\boldsymbol{v}$ is a gradient vector field related to a potential ψ and thus identify $\boldsymbol{v}_t$ as the unique "Wasserstein" velocity vector, while (1.25c)–(1.27c) provide the nonlinear characterization of $\boldsymbol{v}$ in terms of the densities u_t, v_t of μ_t with respect to $\mathscr{L}^d, \gamma$ respectively. We can therefore think $\mathscr{F}^f(\cdot), \mathscr{G}(\cdot)$ as functionals defined on measures μ, whose different realizations $u\mapsto\mathscr{F}^f(u\,\mathscr{L}^d), v\mapsto\mathscr{G}(v\,\gamma)$ depend on the densities of μ w.r.t. the chosen reference measure. In particular, when (1.5a,b,c,d,e) hold,

$$\mathscr{G}(\mu) \overset{(1.5\mathrm{a,b,c,d,e})}{=} \mathscr{F}^f(\mu), \quad \mathscr{F}^f(u\,\mathscr{L}^d) = \mathscr{G}(v\,\gamma), \quad \mu = u\mathscr{L}^d = v\gamma, \tag{1.31}$$

and (1.25a,b,c) and (1.27a,b,c) can be reformulated in terms of μ_t as

$$\partial_t\mu + \mathrm{div}(\mu\boldsymbol{v}) = 0 \qquad \text{in } \Omega_T, \quad \boldsymbol{v}\cdot\boldsymbol{n} = 0 \qquad \text{on } (\partial\Omega)_T, \tag{1.32a}$$

$$\boldsymbol{v} = \mathrm{D}\psi \qquad \text{in } \Omega_T, \tag{1.32b}$$

$$\psi = \begin{cases} -\frac{\delta\mathscr{F}^f(u\,\mathscr{L}^d)}{\delta u} & \text{as in (1.25c)}, \ u := \frac{d\mu}{d\mathscr{L}^d} \\ -\frac{\tilde\delta\mathscr{G}(v\,\gamma)}{\delta v} & \text{as in (1.27c)}, \ v := \frac{d\mu}{d\gamma} \end{cases} \qquad \text{in } \Omega_T, \quad L_{\boldsymbol{p}}\cdot\boldsymbol{n} = 0 \qquad \text{on } (\partial\Omega)_T, \tag{1.32c}$$

where (1.32a) should be intended in the usual weak formulation

$$-\iint_{\Omega_T}\Big(\partial_t\zeta_t(x) + \mathrm{D}\zeta_t(x)\cdot\boldsymbol{v}_t(x)\Big)\,d\mu_t(x)\,\mathrm{d}t = 0 \quad \forall\,\zeta\in C^\infty_c(\mathbb{R}^d\times(0,T)). \tag{1.33}$$

When Ω is unbounded, we impose a condition on μ_t at ∞ by assuming the finiteness both of its quadratic moment and of the "kinetic energy",

$$\mathsf{m}_2^2(\mu_t) := \int_\Omega |x|^2\,\mathrm{d}\mu_t(x), \quad E^2(\mu_t,\boldsymbol{v}_t) := \int_\Omega |\boldsymbol{v}_t(x)|^2\,\mathrm{d}\mu_t(x); \tag{1.34}$$

allowing test functions ζ with a quadratic growth in (1.33), we observe that

$$\begin{aligned}\frac{d}{dt}\mathsf{m}_2^2(\mu_t) &= \frac{d}{dt}\int_\Omega |x|^2\,\mathrm{d}\mu_t(x) \overset{(1.33)}{=} 2\int_\Omega x\cdot\boldsymbol{v}_t(x)\,\mathrm{d}\mu_t(x)\\ &\le 2\Big(\int_\Omega |x|^2\,\mathrm{d}\mu_t(x)\Big)^{1/2}\Big(\int_\Omega |\boldsymbol{v}_t|^2\,\mathrm{d}\mu_t(x)\Big)^{1/2} \le 2\mathsf{m}_2(\mu_t)E(\mu_t,\boldsymbol{v}_t),\end{aligned}$$

and therefore

$$\mathsf{m}_2(\mu_T) \le \mathsf{m}_2(\mu_0) + \int_0^T E(\mu_t,\boldsymbol{v}_t)\,\mathrm{d}t. \tag{1.35}$$

Since we are looking for solutions with finite kinetic energy and bounded quadratic moment (a condition which is in fact restrictive only if Ω is unbounded), the natural ambient spaces for u_t, v_t are

$$
\begin{aligned}
\mathcal{S}_{\mathcal{L}^d}(\Omega) &:= \Big\{ u \in L^1(\Omega) : u \geq 0, \quad \int_\Omega u(x)\,\mathrm{d}x = 1, \quad \int_\Omega |x|^2 u(x)\,\mathrm{d}x < +\infty \Big\}, \\
\mathcal{S}_\gamma(\Omega) &:= \Big\{ v \in L^1_\gamma(\Omega) : v \geq 0, \quad \int_\Omega v(x)\,\mathrm{d}\gamma(x) = 1, \quad \int_\Omega |x|^2 v(x)\,\mathrm{d}\gamma(x) < +\infty \Big\},
\end{aligned} \tag{1.36}
$$

which we will systematically identify with the subset

$$
\mathcal{P}_2^r(\Omega) := \Big\{ \mu = u \cdot \mathcal{L}^d = v \cdot \gamma : u \in \mathcal{S}_{\mathcal{L}^d}(\Omega),\ v \in \mathcal{S}_\gamma(\Omega) \Big\} \subset \mathcal{P}_2(\Omega) \tag{1.37}
$$

of the collection $\mathcal{P}_2(\Omega)$ of all the Borel probability measures on Ω with finite quadratic moment.

The role of the third relation (1.32c) was clarified by [JKO98, Ott01]: considering the case of Fokker-Planck and porous medium equations, those papers proposed a formal interpretation (the so called "Otto calculus") of equations like (1.32a,b,c) as the gradient flow of the appropriate functional $\mathscr{F}^f$, $\mathscr{G}$ with respect to the Wasserstein distance in the space $\mathcal{P}_2(\Omega)$.

In order to understand this point of view (see also [Vil03, AGS05] for further insights on this aspect), let us suppose that $\mu_t = v_t\gamma$ is a smooth solution of (1.32a,b,c) (e.g. with respect to $\mathscr{G}$) with a strictly positive density v_t; we can evaluate the derivative of $\mathscr{G}$ along the trajectory

$$
\begin{aligned}
\frac{d}{dt}\mathscr{G}(\mu_t) &= \frac{d}{dt}\mathscr{G}(v_t\,\gamma) = \frac{d}{dt}\int_\Omega L(v_t, \mathrm{D}v_t)\,\mathrm{d}\gamma = \int_\Omega \Big(L_z(v, \mathrm{D}v)\partial_t v + \sum_i L_{p_i}(v, \mathrm{D}v)\partial_i\partial_t v \Big)\,\mathrm{d}\gamma \\
&\overset{(1.32c)}{=} -\int_\Omega \psi\partial_t v\,\mathrm{d}\gamma \overset{(1.33)}{=} -\int_\Omega \mathrm{D}\psi \cdot \boldsymbol{v}\,\mathrm{d}\mu_t \geq -\Big(\int_\Omega |\mathrm{D}\psi|^2\,\mathrm{d}\mu_t\Big)^{1/2}\Big(\int_\Omega |\boldsymbol{v}|^2\,\mathrm{d}\mu_t\Big)^{1/2}.
\end{aligned}
$$

Since we interpreted $\boldsymbol{v}$ as a sort of "velocity vector" of the curve $\{\mu_t\}_{t\in(0,T)}$ in $\mathcal{P}_2(\Omega)$, whose squared norm is provided by the kinetic energy (1.34), it follows that the maximum rate of decay for $\mathscr{F}$ can be obtained when

$$
\boldsymbol{v} = \mathrm{D}\psi = 2\mathrm{D}\Big(\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}\Big) = \mathrm{D}\frac{\tilde{\delta}\mathscr{G}(v\,\gamma)}{\delta v}, \quad \text{i.e.} \quad (1.32\text{b}), (1.32\text{c}) \text{ hold.}
$$

(1.34) induces a notion of "energy" (and "length") of a curve in $\mathcal{P}_2(\Omega)$:

$$
\mathscr{E}_T(\mu) := \inf\Big\{ \int_0^T \Big(\int_\Omega |\boldsymbol{v}_t(x)|^2\,\mathrm{d}\mu_t(x)\Big)\,\mathrm{d}t : \quad \boldsymbol{v} \text{ satisfying (1.33)} \Big\} \tag{1.38}
$$

and therefore a "geodesic" distance between two element of $\mathcal{P}_2(\Omega)$:

$$
\begin{aligned}
W^2(\mu_1, \mu_2) := \inf\Big\{ T^{-1}\mathscr{E}_T(\mu) : \quad &\mu : [0, T] \to \mathcal{P}_2(\Omega) \text{ is a continuous} \\
&\text{curve in } \mathcal{P}_2(\Omega) \text{ connecting } \mu_1 \text{ to } \mu_2 \Big\}.
\end{aligned} \tag{1.39}
$$

J.D. Benamou and Y. Brenier [BB00, AGS05] showed that W coincide with the so called *Kantorovich-Rubinstein-Wasserstein distance* in $\mathcal{P}_2(\Omega)$ (see also the next section 2.4)

$$
\begin{aligned}
W^2(\mu_1, \mu_2) := \inf\Big\{ &\int_{\Omega\times\Omega} |x_1 - x_2|^2\,\mathrm{d}\boldsymbol{\mu}(x_1, x_2) : \boldsymbol{\mu} \in \mathcal{P}_2(\Omega\times\Omega), \\
&\int_{\Omega\times\Omega} \zeta(x_i)\,\mathrm{d}\boldsymbol{\mu}(x_1, x_2) = \int_\Omega \zeta(x_i)\,\mathrm{d}\mu_i(x_i) \quad \forall\,\zeta \in C^0_0(\Omega),\ i = 1, 2 \Big\}.
\end{aligned} \tag{1.40}
$$

Thus the system (1.32a,b,c) can be considered as the gradient flow of the Fisher information $\frac{1}{2}\mathscr{I}_2(\mu|\,\gamma)$ in the infinite dimensional Riemannian structure associated to W. We refer to [Vil03, AGS05] for other examples, applications and developments of these ideas; here we only recall that

(see Lemma 2.8)

$$\mu_n \rightharpoonup \mu \quad \text{narrowly in } \mathcal{P}(\Omega) \quad \Leftrightarrow \quad \lim_{n\to\infty} \int_\Omega \zeta \,\mathrm{d}\mu_n = \int_\Omega \zeta \,\mathrm{d}\mu \quad \forall\, \zeta \in C^0_b(\Omega), \tag{1.41}$$

$$\mu_n \to \mu \quad \text{in } \mathcal{P}_2(\Omega) \quad \Leftrightarrow \quad \lim_{n\to\infty} W(\mu_n,\mu) = 0 \quad \Leftrightarrow \quad \begin{cases} \mu_n \rightharpoonup \mu \quad \text{narrowly in } \mathcal{P}(\Omega), \\ \lim\limits_{n\to\infty} \mathsf{m}_2(\mu_n) = \mathsf{m}_2(\mu). \end{cases} \tag{1.42}$$

1.7. **The variational approximation scheme.** F. Otto suggested that the previous formal discussion can be rigorously justified by introducing a variational approximation scheme which, in our case, can be used to prove the existence of a solution of (EE_1) and ($\mathrm{EE}_{1,\gamma}$). This algorithm is in fact a particular example of a procedure which can be performed in any metric spaces and also in more general frameworks, as proposed by E. De Giorgi [DGMT80, De 93, AGS05].

We fix a time step $\tau > 0$ and we consider the partition of the time interval $(0,+\infty)$ given by

$$\mathcal{P}_\tau := \Big\{ 0 < t^1_\tau < \cdots < t^n_\tau < \cdots \Big\}, \quad t^n_\tau := n\tau, \quad n = 0,1,\cdots \tag{1.43}$$

Starting from a given approximation M^0_τ of $\mu_0 = v_0 \cdot \gamma = u_0 \mathcal{L}^d$, we recursively solve the minimum problem in the unknowns $\{M^n_\tau\}_{n=1}^\infty$

$$\begin{cases} \Phi(\tau, M^{n-1}_\tau; M^n_\tau) = \min\limits_{\mu\in\mathcal{P}_2(\Omega)} \Phi(\tau, M^{n-1}_\tau; \mu) \\ \text{where} \quad \Phi(\tau, M; \mu) := \dfrac{W^2(M,\mu)}{2\tau} + \phi(\mu) \quad \forall\, M,\mu \in \mathcal{P}_2(\Omega); \end{cases} \tag{1.44}$$

here ϕ is one of the functionals

$$\mathscr{F}^f(\mu) = \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \int_\Omega f \,\mathrm{d}\mu, \qquad \mathscr{G}(\mu) = \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma), \tag{1.45}$$

and $\mathscr{I}_2(\cdot|\,\mathcal{L}^d), \mathscr{I}_2(\cdot|\,\gamma)$ are the Relative Fisher information functionals defined as in (1.17a,b,c)

$$\mathscr{I}_2(\mu|\,\mathcal{L}^d) := \begin{cases} 4\int_\Omega |\mathrm{D}r|^2 \,\mathrm{d}x & \text{if } \mu = r^2 \cdot \mathcal{L}^d \text{ and } r \in W^{1,2}(\Omega), \\ +\infty & \text{otherwise;} \end{cases} \tag{1.46}$$

$$\mathscr{I}_2(\mu|\,\gamma) := \begin{cases} 4\int_\Omega |\mathrm{D}s|^2 \,\mathrm{d}\gamma & \text{if } \mu = s^2 \cdot \gamma \text{ and } s \in W^{1,2}_\gamma(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{1.47}$$

$W^{k,p}_\gamma(\Omega)$ is the "weighted Sobolev space" of functions with derivatives in $L^p_\gamma(\Omega)$ up to the order k (see next section 2.2). By standard lower semicontinuity and compactness argument, it is not difficult (see the next Theorem 2.13) to check that the minimum problem (1.47) admits a solution and therefore a minimizing sequence $\{M^n_\tau\}_{n=1}^\infty$ always exists; a convexity argument shows that it is also uniquely determined by the algorithm. We denote by $\overline{M}_\tau : [0,+\infty) \mapsto \mathcal{P}_2(\Omega)$ the piecewise constant interpolant of the values M^n_τ on the grid $\mathcal{P}_\tau$, i.e.

$$\overline{M}_\tau(t) \equiv M^n_\tau \quad \text{if } t \in (t^{n-1}_\tau, t^n_\tau]. \tag{1.48}$$

Definition 1.3 (**Minimizing Movements (De Giorgi** [De 93]**)).** Let $\phi : \mathcal{P}_2(\Omega) \to (-\infty,+\infty]$ be a given functional (as in (1.45)). $GMM(\mu_0,\phi)$ (the so called "*Generalized Minimizing Movements*") is the set of all the (pointwise) accumulation points of M_τ as $\tau \downarrow 0$ in $\mathcal{P}(\Omega)$ (endowed with the narrow topology, i.e. the weak* topology induced by the duality with continuous and bounded real functions, see § 2.1) provided

$$M^0_\tau \rightharpoonup \mu_0 \quad \text{in } \mathcal{P}(\Omega), \qquad \limsup_{\tau\downarrow 0} \mathsf{m}_2(M^0_\tau) < +\infty, \qquad \limsup_{\tau\downarrow 0} \phi(M^0_\tau) < +\infty, \tag{1.49}$$

i.e.

$$\mu \in GMM(\mu_0,\phi) \quad \Leftrightarrow \quad \begin{cases} \exists\, \tau_n \downarrow 0 : \overline{M}_{\tau_n,t} \rightharpoonup \mu_t \quad \text{in } \mathcal{P}(\Omega) \quad \forall\, t \in [0,+\infty), \\ \sup\limits_n \mathsf{m}_2(M^0_{\tau_n}) < +\infty, \quad \sup\limits_n \phi(M^0_{\tau_n}) < +\infty. \end{cases} \tag{1.50}$$

We shall show that if $\mu_t = u_t\mathcal{L}^d = v_t\gamma$ is an element of $GMM(\mu_0,\phi)$, then u, v are a "variational" solution of (EE_1) and ($\mathrm{EE}_{1,\gamma}$) respectively: we are now making precise the related definition.

1.8. **The notions of weak solution.** (EE_3) and ($\mathrm{EE}_{3,\gamma}$) provide the easiest form of the equation which admits a weak formulation. Recalling the definition (1.36) of $\mathcal{S}_{\mathcal{L}^d}(\Omega), \mathcal{S}_\gamma(\Omega)$ and the fact that

$$\frac{\partial_i v\,\partial_j v}{v} = 4\partial_i\sqrt{v}\partial_j\sqrt{v}, \qquad \int_\Omega \Big|\frac{\partial_i v\,\partial_j v}{v}\Big|\,\mathrm{d}\gamma \le \mathscr{I}_2(v\,\gamma|\,\gamma),$$

we introduce the following definition:

Definition 1.4 (**Weak solutions** of (EE_3),($\mathrm{EE}_{3,\gamma}$))**.** Let $u : [0,+\infty) \mapsto \mathcal{S}_{\mathcal{L}^d}(\Omega)$ (resp. $v : [0,+\infty) \mapsto \mathcal{S}_\gamma(\Omega)$) be narrowly continuous such that

$$r := \sqrt{u} \in L^2_{\mathrm{loc}}(0,+\infty; W^{1,2}(\Omega)), \qquad s := \sqrt{v} \in L^2_{\mathrm{loc}}(0,+\infty; W^{1,2}_\gamma(\Omega)). \tag{1.51}$$

We say that u is a distributional solution of

$$\partial_t u + \Delta^2 u - \sum_{i,j=1}^d \partial^2_{ij}\Big(\frac{\partial_i u\,\partial_j u}{u}\Big) - \sum_{i=1}^d \partial_i(u\partial_i f) = 0 \quad \text{in } \Omega_\infty = \Omega\times(0,+\infty), \tag{1.52}$$

with boundary conditions (1.2) and (1.3) if

$$\iint_{\Omega_\infty} \Big(- u\partial_t\zeta - \sum_{i=1}^d \partial_i u\,\partial_i\Delta\zeta - 4\sum_{i,j=1}^d \partial^2_{ij}\zeta\,\partial_i r\,\partial_j r + \sum_{i=1}^d u\partial_i f\partial_i\zeta\Big)\,\mathrm{d}x\,\mathrm{d}t = 0 \tag{1.53}$$
$$\forall \zeta \in C^\infty_{\mathrm{c}}(\mathbb{R}^d\times(0,+\infty)) \quad \text{with} \quad \partial_{\boldsymbol{n}}\zeta = 0 \quad \text{on } \partial\Omega\times(0,+\infty).$$

We say that v is a distributional solution of

$$\partial_t v + \Delta^2_\gamma v - \sum_{i,j=1}^d \tilde\partial^2_{ij}\Big(\frac{\partial_i v\,\partial_j v}{v}\Big) = 0 \quad \text{in } \Omega_\infty, \quad \partial_{\boldsymbol{n}} v = \partial_{\boldsymbol{n}}\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}} = 0 \quad \text{on } (\partial\Omega)_\infty, \tag{1.54}$$

if

$$\iint_{\Omega_\infty} \Big(- v\partial_t\zeta - \sum_{i=1}^d \partial_i v\,\partial_i\Delta_\gamma\zeta - 4\sum_{i,j=1}^d \partial^2_{ij}\zeta\,\partial_i s\,\partial_j s\Big)\,\mathrm{d}\gamma(x)\,\mathrm{d}t = 0 \tag{1.55}$$
$$\forall \zeta \in C^\infty_{\mathrm{c}}(\mathbb{R}^d\times(0,+\infty)) \quad \text{with} \quad \partial_{\boldsymbol{n}}\zeta = 0 \quad \text{on } \partial\Omega\times(0,+\infty).$$

The main difficulties in the rigorous formulation of the Lyapunov identity (1.18) and in the definition of a reasonable notion of variational solution of (EE_1) and ($\mathrm{EE}_{1,\gamma}$) are represented by the terms (recall (1.9))

$$\boldsymbol{q} = u\mathrm{D}\Big(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\Big), \quad \tilde{\boldsymbol{q}} = 2v\mathrm{D}\Big(\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}\Big). \tag{1.56}$$

Since the particular structure of the Fisher information shows that $r := \sqrt{u},\ s = \sqrt{v} = \sqrt{u/g}$ should play a crucial role, we can use the identities

$$\boldsymbol{q} := r^2\mathrm{D}\Big(2\frac{\Delta r}{r} - f\Big) = 2\mathrm{D}\big(r\Delta r\big) - 4\Delta r\mathrm{D}r - r^2 Df, \tag{1.57a}$$

$$\tilde{\boldsymbol{q}} := 2s^2\mathrm{D}\Big(\frac{\Delta_\gamma s}{s}\Big) = 2\mathrm{D}\big(s\Delta_\gamma s\big) - 4\Delta_\gamma s\mathrm{D}s, \tag{1.57b}$$

yielding, together with the boundary conditions $\boldsymbol{q}\cdot\boldsymbol{n} = 0$, $\tilde{\boldsymbol{q}}\cdot\boldsymbol{n} = 0$ on $(\partial\Omega)_\infty$,

$$\int_\Omega \zeta\ \mathrm{div}\,\boldsymbol{q}\,\mathrm{d}x = -\int_\Omega \boldsymbol{q}\cdot\mathrm{D}\zeta\,\mathrm{d}x = \int_\Omega \Big(2r\Delta r\Delta\zeta + 4\Delta r\mathrm{D}r\cdot\mathrm{D}\zeta + u\mathrm{D}f\cdot\mathrm{D}\zeta\Big)\,\mathrm{d}x, \tag{1.58a}$$

$$\int_\Omega \zeta\ \mathrm{div}_\gamma\tilde{\boldsymbol{q}}\,\mathrm{d}\gamma = -\int_\Omega \tilde{\boldsymbol{q}}\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma = \int_\Omega \Big(2s\Delta_\gamma s\Delta_\gamma\zeta + 4\Delta_\gamma s\mathrm{D}s\cdot\mathrm{D}\zeta\Big)\,\mathrm{d}\gamma, \tag{1.58b}$$

for every test function $\zeta \in C^\infty_c(\mathbb{R}^d)$ with $\partial_{\boldsymbol{n}}\zeta = 0$ on $(\partial\Omega)_\infty$.

Definition 1.5 (Weak solutions of (EE$_1$) **and** (EE$_{1,\gamma}$)**).** Let $u : [0,+\infty) \mapsto \mathcal{S}_{\mathcal{L}^d}(\Omega)$ (resp. $v : [0,+\infty) \mapsto \mathcal{S}_\gamma(\Omega)$) be narrowly continuous such that

$$r = \sqrt{u},\ \ s = \sqrt{v} \in L^2_{\rm loc}(0,+\infty; W^{2,2}_{\rm loc}(\overline{\Omega})), \quad \partial_{\boldsymbol n} r = 0,\ \partial_{\boldsymbol n} s = 0 \quad \text{on } (\partial\Omega)_\infty. \tag{1.59}$$

We say that u is a weak solution of (EE$_1$)

$$\partial_t u + \operatorname{div}\Big(u\mathrm{D}\Big(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\Big)\Big) = 0 \quad \text{in } \Omega_\infty, \tag{1.60}$$

$$\partial_{\boldsymbol n}\sqrt{u} = 0,\ u\partial_{\boldsymbol n}\Big(2\frac{\Delta\sqrt{u}}{\sqrt{u}} - f\Big) = \boldsymbol q\cdot\boldsymbol n = 0 \quad \text{on } (\partial\Omega)_\infty, \tag{1.61}$$

if

$$\begin{gathered}\iint_{\Omega_\infty}\Big(-u\partial_t\zeta + 2r\Delta r\,\Delta\zeta + 4\Delta r\,\mathrm{D}r\cdot\mathrm{D}\zeta + u\mathrm{D}f\cdot\mathrm{D}\zeta\Big)\,\mathrm{d}x\,\mathrm{d}t = 0\\ \forall\zeta\in C^\infty_{\rm c}(\mathbb{R}^d\times(0,+\infty)),\quad \partial_{\boldsymbol n}\zeta = 0 \quad \text{on } (\partial\Omega)_\infty.\end{gathered} \tag{1.62}$$

We say that v is a weak solution of (EE$_{1,\gamma}$)

$$\partial_t v + 2\operatorname{div}_\gamma\Big(v\mathrm{D}\Big(\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}}\Big)\Big) = 0 \quad \text{in } \Omega_\infty, \tag{1.63}$$

$$\partial_{\boldsymbol n}\sqrt{v} = 0,\ v\partial_{\boldsymbol n}\frac{\Delta_\gamma\sqrt{v}}{\sqrt{v}} = \tilde{\boldsymbol q}\cdot\boldsymbol n = 0 \quad \text{on } (\partial\Omega)_\infty, \tag{1.64}$$

if

$$\begin{gathered}\iint_{\Omega_\infty}\Big(-v\partial_t\zeta + 2s\Delta_\gamma s\,\Delta_\gamma\zeta + 4\Delta_\gamma s\,\mathrm{D}s\cdot\mathrm{D}\zeta\Big)\,\mathrm{d}\gamma(x)\,\mathrm{d}t = 0\\ \forall\zeta\in C^\infty_{\rm c}(\mathbb{R}^d\times(0,+\infty)),\quad \partial_{\boldsymbol n}\zeta = 0 \quad \text{on } (\partial\Omega)_\infty.\end{gathered} \tag{1.65}$$

An expression of the type $s \in L^2_{\rm loc}(0,+\infty; W^{k,p}_{\rm loc}(\overline{\Omega}))$ as in (1.59) simply means that for every *bounded* open subset $\Omega' \subset \Omega$ and every $0 < a < b < T$ the restriction of s to $\Omega'\times(a,b)$ belongs to $L^2(a,b; W^{k,p}(\Omega'))$.

Remark 1.6 (Boundary conditions). Observe that the first boundary condition (1.61), (1.64) is simply imposed in the sense of traces in $W^{2,2}(\Omega')$, Ω' being an arbitrary bounded open subset of Ω. A weak integral formulation of the second boundary condition of (1.61)–(1.64) is contained in (1.62)–(1.65) since the test functions ζ are not required to vanish on $\partial\Omega$.

Remark 1.7 (Other formulations). Starting from (1.65) it is not difficult to show that u (resp. v) is also a weak solution of (EE$_3$) (resp. EE$_{3,\gamma}$) according to Definition 1.4 (see the calculations in Corollary 5.9).
By using the identities for $s = \sqrt{v} = \sqrt{u/g}$

$$v\,\mathrm{D}^2\log v = 2\Big(s\mathrm{D}^2 s - \mathrm{D}s\otimes\mathrm{D}s\Big), \quad v\partial^2_{ij}\log v = \partial^2_{ij}v - \frac{\partial_i v\partial_j v}{v} = 2\Big(s\partial^2_{ij}s - \partial_i s\partial_j s\Big) \tag{1.66}$$

(and the corresponding ones for $\log u$ and $r = \sqrt{u}$) it is also possible to obtain the weak formulation of (EE$_{2,\gamma}$) (and of (EE$_2$)) which reads as

$$\iint_{\Omega_\infty}\Big(-v\partial_t\zeta + 4\sum_{i,j}\big(s\partial^2_{ij}s - \partial_i s\,\partial_j s\big)\,\partial^2_{ij}\zeta + \partial^2_{ij}V\,\partial_j v\,\partial_i\zeta\Big)\,\mathrm{d}\gamma\,\mathrm{d}t = 0 \tag{1.67}$$

for every $\zeta \in C^\infty_{\rm c}(\Omega\times(0,+\infty))$.

Remark 1.8 (Classical solutions). It should be clear from the previous discussion that any classical solution u, v with $u(x,t), v(x,t) > 0\ \forall\,(x,t)\in\Omega_\infty$ satisfies the previous variational formulations.

1.9. **Main results: equations ($\mathbf{EE}_{1,2,3}$).** We are presenting here our main results concerning equations ($\mathrm{EE}_{1,2,3}$) as gradient flows of the "perturbed" Fisher information functional $\mathscr{F}^f$. The particular structure of the Fisher information w.r.t. the Lebesgue measure is quite useful to overcome some technical difficulties, arising when a general log-concave measure γ is involved. Nevertheless, when (1.5a,b,c,d,e) hold, the invariant measure γ and its related functionals play a crucial role in the study of the asymptotic behaviour of the solution; in particular, when γ is induced by a potential V which is λ-convex for some $\lambda > 0$, we can prove the exponential decay of the solution to the equilibrium.

We will distinguish two cases: the first one, Theorem 1, deals with initial data with finite (relative) Fisher information and provides the strongest properties for the solution. Theorems 2 and 3 concern the behaviour of the quadratic moments, the Relative Entropy, and the Relative Fisher Information of the solutions.

Theorem 4 allows to relax the assumptions on the initial data, which are only supposed to have finite entropy: we shall show that in this case the solution exhibits a nice regularization effect, which is sufficient to recover almost all the result of the previous theorems.

Theorem 5 focuses on the particular case $f \equiv 0$ in a bounded set Ω: even if the related potential V is constant (thus convex but not λ-convex for any strictly positive λ), still we can prove an exponential convergence of the solution to the (constant) steady state. In the one-dimensional case with periodic boundary conditions this result has been obtained by [DGJ06].

In the sequel, given a measure ν on the set A and two ν-measurable functions $\boldsymbol{a} : A \to \mathbb{R}^k$, $b : A \to \mathbb{R}$, we adopt the usual convention

$$\begin{aligned} &\boldsymbol{a}/b \quad \text{is well defined only if} \quad \boldsymbol{a}(x) = 0 \quad \text{for } \nu\text{-a.e. } x \in A \text{ with } b(x) = 0, \\ &\quad \text{with} \quad \boldsymbol{a}(x)/b(x) = 0 \quad \text{if} \quad \boldsymbol{a}(x) = 0,\ b(x) = 0. \end{aligned} \tag{1.68}$$

In particular, observe that if $\mu = v\gamma = s^2\gamma$, the following expressions are equivalent

$$\int_\Omega \frac{|\boldsymbol{a}(x)|^2}{v(x)}\,\mathrm{d}\gamma(x) = \int_\Omega \left|\frac{\boldsymbol{a}(x)}{s(x)}\right|^2 \mathrm{d}\gamma(x) = \int_\Omega \left|\frac{\boldsymbol{a}(x)}{v(x)}\right|^2 \mathrm{d}\mu(x). \tag{1.69}$$

We introduce the second order functionals (cf. (1.15) and (2.8))

$$\mathscr{K}_{-1}(\mu|\,\mathcal{L}^d) := 4\int_\Omega \left|\frac{r\mathrm{D}^2 r - \mathrm{D}r\otimes\mathrm{D}r}{r}\right|^2 \mathrm{d}x \quad \text{if } \mu = r^2\mathcal{L}^d \in \mathcal{P}_2^r(\Omega),\ r \in W^{2,2}_{\mathrm{loc}}(\Omega), \tag{1.70}$$

$$\mathscr{K}_{-1}(\mu|\,\gamma) := 4\int_\Omega \left|\frac{s\mathrm{D}^2 s - \mathrm{D}s\otimes\mathrm{D}s}{s}\right|^2 \mathrm{d}\gamma \quad \text{if } \mu = s^2\gamma \in \mathcal{P}_2^r(\Omega),\ s \in W^{2,2}_{\mathrm{loc}}(\Omega). \tag{1.71}$$

Theorem 1 **(G.M.M. are solutions of** (EE_1)**).** *Let us suppose that Ω is an open and* convex *subset of $\mathbb{R}^d$, $f : \Omega \to \mathbb{R}$ satisfies* (1.4)*, ϕ is the functional $\mathscr{F}^f(\cdot)$ of* (1.20)*, and $\mu_0 = u_0\mathcal{L}^d, M^0_\tau \in \mathcal{P}_2^r(\Omega)$ are the (continuous and discrete) initial data satisfying* (1.49)*, i.e.*

$$\begin{gathered} M^0_\tau \rightharpoonup \mu_0 \quad \text{in } \mathcal{P}(\Omega), \quad \limsup_{\tau\downarrow 0} \mathsf{m}_2(M^0_\tau) = \mathsf{m}_{2,0} < +\infty, \\ \mathscr{I}_2(\mu_0|\,\mathcal{L}^d) \le \mathscr{I}_0 := \limsup_{\tau\downarrow 0} \mathscr{I}_2(M^0_\tau|\,\mathcal{L}^d) < +\infty. \end{gathered} \tag{1.72}$$

i) ***Existence and regularity of Generalized Minimizing Movements.*** *For every step $\tau > 0$, the variational scheme* (1.44) *admits a unique solution $\{M^n_\tau\}_{n\in\mathbb{N}}$ with $M^n_\tau = U^n_\tau\mathcal{L}^d = (R^n_\tau)^2\mathcal{L}^d$ and each infinitesimal sequence τ_k of time steps admits a subsequence (still denoted by τ_k), such that*

$$\overline{M}_{\tau_{k,t}} \rightharpoonup \mu_t \quad \text{in } \mathcal{P}(\Omega), \tag{1.73}$$

$$\overline{U}_{\tau_{k,t}} \to u_t \quad \text{strongly in } L^p(\Omega) \quad \forall\, t > 0,\ 1 \le p < \tfrac{1}{2}2^\star, \tag{1.74}$$

where $2^\star$ denotes the usual Sobolev exponent $\quad 2^\star := \begin{cases} \frac{2d}{d-2} & \text{if } d > 2, \\ +\infty & \text{if } d \le 2, \end{cases}$

$$\lim_{k\uparrow+\infty} \mathscr{H}(\overline{M}_{\tau_k,t}|\,\mathcal{L}^d) = \mathscr{H}(\mu_t|\,\mathcal{L}^d) \quad \forall\, t \ge 0, \tag{1.75}$$

$$\lim_{k\uparrow+\infty} \mathscr{I}_2(\overline{M}_{\tau_k,t}|\,\mathcal{L}^d) = \mathscr{I}_2(\mu_t|\,\mathcal{L}^d) \quad \text{for a.e. } t > 0, \tag{1.76}$$

$$\overline{R}_{\tau_k} \to r \quad \text{strongly in } L^2(0,T;W^{1,2}(\Omega)) \quad \text{and weakly in } L^2(0,T;W^{2,2}(\Omega)) \quad \forall\, T > 0; \tag{1.77}$$

$\mu_t = u_t\mathcal{L}^d = r_t^2\mathcal{L}^d \in \mathcal{P}_2^r(\Omega)$ for every $t > 0$ and r satisfies the regularity properties (slightly stronger than (1.59)*)*

$$\begin{aligned} &r = \sqrt{u} \in L^2(0,T;W^{2,2}(\Omega)), \quad \partial_{\boldsymbol{n}} r = 0 \quad \text{on } (\partial\Omega)_\infty, \\ &r\Delta r \in L^1((0,T);W^{1,1}(\Omega)), \quad \sqrt{r} \in L^4(0,T;W^{1,4}(\Omega)), \end{aligned} \tag{1.78}$$

for every $T > 0$. All the limit curves μ obtained in this way are the elements of $GMM(\mu_0;\mathscr{F}^f)$, which in particular is non empty.

ii) ***Generalized Minimizing Movement are variational solutions.*** *If $\mu = u\mathcal{L}^d$ belongs to $GMM(\mu_0,\mathscr{F}^f)$ then u is a variational solution of* (EE$_1$) *according to Definition 1.5 and it satisfies*

$$\partial_t u + \operatorname{div} \boldsymbol{q} = 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (0,+\infty)), \quad \text{for } \boldsymbol{q} := u\boldsymbol{v} = 2\mathrm{D}(r\Delta r) - 4\Delta r\,\mathrm{D}r - u\mathrm{D}f \tag{1.79}$$

with

$$\int_0^T \int_\Omega \frac{|\boldsymbol{q}_t|^2}{u_t}\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega |\boldsymbol{v}_t|^2\,\mathrm{d}\mu_t(x)\,\mathrm{d}t < +\infty \qquad \forall\, T > 0. \tag{1.80}$$

iii) ***Entropy inequalities.*** *The map $t \mapsto \mathscr{H}(\mu_t|\,\mathcal{L}^d)$ is absolutely continuous and for a.e. $t > 0$*

$$\Big(\mathscr{I}_2(\mu_t|\,\mathcal{L}^d)\int_\Omega |\boldsymbol{v}_t|^2\,\mathrm{d}\mu_t\Big)^{1/2} \ge -\frac{d}{dt}\mathscr{H}(\mu_t|\,\mathcal{L}^d) \ge \mathscr{K}_{-1}(\mu_t|\,\mathcal{L}^d) + \int_\Omega \mathrm{D}f\cdot\mathrm{D}u_t\,\mathrm{d}x. \tag{1.81}$$

Moreover

$$\mathscr{K}_{-1}(\mu_t|\,\mathcal{L}^d) \ge \frac{12}{2+d}\int_\Omega \big|\mathrm{D}^2 r_t\big|^2\,\mathrm{d}x + \frac{64}{3(2+d)}\int_\Omega |\mathrm{D}\sqrt{r_t}|^4\,\mathrm{d}x. \tag{1.82}$$

When Ω is a *bounded* set, there is no difference between narrow convergence in $\mathcal{P}(\Omega)$ and convergence in $\mathcal{P}_2(\Omega)$ (recall (1.42)). In the *unbounded* case $\mathcal{P}_2(\Omega)$ is endowed with the finer topology induced by the Wasserstein distance and it could be interesting to know if (1.73) can be improved to obtain this stronger convergence. We consider here the case when Ω is a ***convex cone***, i.e. $\alpha\Omega = \Omega$, $\forall\,\alpha > 0$, a condition trivially satisfied when $\Omega = \mathbb{R}^d$. In this case it is possible to find a differential equation satisfied by $t \mapsto \mathsf{m}_2^2(\mu_t)$, which is useful to prove the convergence of the quadratic moments of the approximating family $\overline{M}_\tau$.

Theorem 2 (**Quadratic moments**). *Under the same assumptions of the previous Theorem 1, let us further assume that Ω is a (convex)* cone. *Then the map $t \mapsto \mathsf{m}_2^2(\mu_t)$ (which is always absolutely continuous) satisfies*

$$\frac{d}{dt}\frac{1}{2}\int_\Omega |x|^2\,\mathrm{d}\mu_t(x) = \mathscr{I}_2(\mu_t|\,\mathcal{L}^d) - \int_\Omega \mathrm{D}f\cdot x\,\mathrm{d}\mu_t(x) \quad \text{for a.e. } t > 0. \tag{1.83}$$

Moreover, if

$$M_\tau^0 \to \mu_0 \quad \text{in } \mathcal{P}_2(\Omega), \tag{1.84}$$

then each narrowly convergent subsequence $\overline{M}_{\tau_k}$ according to Theorem 1 also satisfies

$$\overline{M}_{\tau_k,t} \to \mu_t \quad \text{in } \mathcal{P}_2(\Omega) \quad \forall\, t \ge 0. \tag{1.85}$$

The next Theorem deals with the Energy inequality and asymptotic behaviour of the solutions given by Theorem 1.

Theorem 3 (**Energy inequalities and asymptotic behaviour**). *Under the same assumptions of Theorem 1 let us further suppose that at least one of the following conditions is satisfied:*

H1) *The function $f:\Omega\to\mathbb{R}^d$ has a sub-quadratic growth, i.e.* $\limsup_{|x|\to\infty}\dfrac{|f(x)|}{|x|^2}=0$

(a condition trivially satisfied when Ω is bounded).

H2) Ω *is a* cone *and the initial data satisfy* (1.84).

Then the map $t\mapsto \mathscr{F}^f(\mu_t)$ is essentially decreasing (i.e. it coincides with a decreasing function up to a negligible set) and satisfies

$$\mathscr{F}^f(\mu_{t_1})+\int_{t_0}^{t_1}\int_\Omega\Big|\frac{\boldsymbol{q}_t(x)}{u_t(x)}\Big|^2\,\mathrm{d}\mu_t(x)\,\mathrm{d}t\le\mathscr{F}^f(\mu_{t_0})\le\frac12\mathscr{J}_0+\langle f,\mu_0\rangle \quad \text{for a.e. } t_0>0 \quad \text{and every } t_1\ge t_0. \tag{1.86}$$

Moreover, if f is of the type (1.5a,b,c,d,e) *for a λ-convex potential V inducing the invariant measure $\gamma=e^{-V}\mathcal{L}^d$, then $\mathscr{F}^f(\cdot)=\frac12\mathscr{I}_2(\cdot|\,\gamma)$ and*

$$\Big(-\frac{d}{dt}\big(\frac12\mathscr{I}_2(\mu_t|\,\gamma)\big)^2\Big)^{1/2}\ge-\frac{d}{dt}\mathscr{H}(\mu_t|\,\gamma)\ge\mathscr{K}_{-1}(\mu_t|\,\gamma)+\lambda\mathscr{I}_2(\mu_t|\,\gamma)\qquad \text{for a.e. } t>0. \tag{1.87}$$

In particular, when $\lambda\ge0$,

$$\mathscr{H}(\mu_t|\,\gamma)\le\mathscr{H}(\mu_0|\,\gamma)e^{-2\lambda^2t}, \tag{1.88}$$

and, assuming $\mathscr{J}_0=\mathscr{I}_2(\mu_0|\,\mathcal{L}^d)$,

$$\mathscr{I}_2(\mu_t|\,\gamma)\le\mathscr{I}_2(\mu_0|\,\gamma)e^{-2\lambda^2t}. \tag{1.89}$$

Remark 1.9 (The role of the Logarithmic Sobolev inequality). (1.88) and (1.89) are strictly related to the Logarithmic-Sobolev inequality [Gro76]

$$\mathscr{H}(\mu|\,\gamma)\le\frac{1}{2\lambda}\mathscr{I}_2(\mu|\,\gamma)\quad\forall\,\mu\in\mathcal{P}_2^r(\Omega), \tag{1.90}$$

which holds for every Log-concave measure γ whose inducing potential V is λ-convex, $\lambda>0$. thanks to Bakry-Emery criterion [BÉ85] (see also [Tos97, OV00]).

Remark 1.10 (L^1-estimates). Recall that by the Csiszár-Kullback-Pinsker inequality

$$\|u_t-g\|^2_{L^1(\Omega)}\le2\mathscr{H}(\mu_t|\,\gamma), \tag{1.91}$$

(1.88) provides an exponential convergence of u_t to the density g of the invariant measure γ in the usual L^1-norm.

Theorem 4 (**Regularizing effect under finite Entropy**). *Let us suppose that $\Omega\subset\mathbb{R}^d$ is either a* bounded *convex set or a convex* cone, *$f:\Omega\to\mathbb{R}$ satisfies* (1.4), *ϕ is the functional $\mathscr{F}^f(\cdot)$ of* (1.20).

> ***All the same statements of the previous Theorems 1, 2, and 3 still hold,***
> *except for* (1.80), *where the time integral should be restricted to any interval (ε,T),*
> $0<\varepsilon<T<+\infty$,

even if condition (1.72) *on the (continuous and discrete) initial data $\mu_0=u_0\mathcal{L}^d, M^0_\tau\in\mathcal{P}_2^r(\Omega)$ is replaced by the finite Entropy condition*

$$\begin{gathered} M^0_\tau\rightharpoonup\mu_0\quad\text{in }\mathcal{P}(\Omega),\quad \limsup_{\tau\downarrow0}\mathsf{m}_2(M^0_\tau)=\mathsf{m}_{2,0}<+\infty,\\ \lim_{\tau\downarrow0}\mathscr{H}(M^0_\tau|\,\mathcal{L}^d)=\mathscr{H}(\mu_0|\,\mathcal{L}^d)<+\infty.\end{gathered} \tag{1.92}$$

In particular for every $t>0$ the Fisher information $\mathscr{I}_2(\mu_t|\,\mathcal{L}^d)$ is finite and there exists a constant C only dependent on C_f and the dimension d such that

$$\mathscr{H}(\mu_t|\,\mathcal{L}^d)+\big(\pi+\tfrac12\big)\mathsf{m}_2^2(\mu_t)+\frac14\int_0^t\mathscr{K}_{-1}(\mu_s|\,\mathcal{L}^d)\,\mathrm{d}s\le\big(\mathscr{H}(\mu_0|\,\mathcal{L}^d)+(\pi+\tfrac12)\mathsf{m}_{2,0}^2+Ct\big)e^{Ct}, \tag{1.93}$$

$$\limsup_{t\downarrow0}\sqrt{t}\mathscr{I}_2(\mu_t|\,\gamma)\le C\Big(\mathscr{H}(\mu_0|\,\mathcal{L}^d)+(\pi+\tfrac12)\mathsf{m}_{2,0}^2\Big)^{1/2}. \tag{1.94}$$

In the particular case of a constant potential V and a bounded open set Ω, corresponding to the reference (normalized) Lebesgue measure $\gamma := Z^{-1}\mathscr{L}^d$, $Z := \mathscr{L}^d(\Omega)$, we can still obtain interesting information on the asymptotic decay of the solution. Let us first introduce the best positive constants $\alpha_\Omega, \beta_\Omega > 0$ in the "vectorial" Poincaré inequality

$$\int_\Omega |\mathrm{D}\boldsymbol{\xi}|^2\,\mathrm{d}x \ge \alpha_\Omega \int_\Omega |\boldsymbol{\xi}|^2\,\mathrm{d}x \quad \forall\,\boldsymbol{\xi} \in W^{1,2}(\Omega;\mathbb{R}^d), \quad \boldsymbol{\xi}\cdot\boldsymbol{n} = 0 \text{ on } \partial\Omega, \tag{1.95}$$

and in the Logarithmic-Sobolev inequality

$$\mathscr{I}_2(\mu|\,\gamma) \ge \beta_\Omega \mathscr{H}(\mu|\,\gamma) \quad \forall\,\mu \in \mathcal{P}^r(\Omega), \tag{1.96}$$

which, up to the transformation $\mu = r^2\mathscr{L}^d$, is equivalent to

$$4\int_\Omega |\mathrm{D}r|^2\,\mathrm{d}x \ge \beta_\Omega \int_\Omega r^2 \log\Big(Z\frac{r^2}{\int_\Omega r^2\,\mathrm{d}x}\Big)\,\mathrm{d}x \quad \forall\,r \in W^{1,2}(\Omega), \quad r \ne 0. \tag{1.97}$$

Theorem 5 (**Asymptotic decay in the case of Lebesgue measure**). *Let us suppose that Ω is a* bounded *convex open subset of $\mathbb{R}^d$, let $Z := \mathscr{L}^d(\Omega)$ (so that $\gamma = Z^{-1}\mathscr{L}^d \in \mathcal{P}^r(\Omega)$), and let $\mu_t = u_t\mathscr{L}^d \in GMM(\mu_0, \mathscr{F}^0)$ be a variational solution of* (EE_1) *with $f = 0$. Then*

$$\mathscr{H}(\mu_t|\,Z^{-1}\mathscr{L}^d) \le e^{-\beta t}\mathscr{H}(\mu_0|\,Z^{-1}\mathscr{L}^d), \quad \mathscr{I}_2(\mu_t|\,Z^{-1}\mathscr{L}^d) \le e^{-\alpha t}\mathscr{I}_2(\mu_0|\,Z^{-1}\mathscr{L}^d) \quad \forall\,t > 0, \tag{1.98}$$

where

$$\alpha := 2\Big(\frac{3\alpha_\Omega}{d+2}\Big)^2, \qquad \beta := \frac{3\,\alpha_\Omega\,\beta_\Omega}{d+2}. \tag{1.99}$$

In any case, even if Ω is not bounded, we have

$$\mathscr{I}_2(\mu_t|\,\mathscr{L}^d) \le \frac{d+2}{\sqrt{t}}. \tag{1.100}$$

1.10. **Main results: equations $(\mathbf{EE}_{1,2,3\,\gamma})$ for a general measure γ.** Let us now consider the case of the relative Fisher information with respect to a general measure $\gamma = e^{-V}\mathscr{L}^d$

$$\phi(\mu) := \mathscr{G}(\mu) = \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma), \tag{1.101}$$

allowing potentials V with arbitrary growth at infinity.

Theorem 6 (**G.M.M. are solutions of** $(\mathrm{EE}_{1,\gamma})$). *Let us suppose that Ω is an open* convex *subset of $\mathbb{R}^d$, $V : \mathbb{R}^d \to \mathbb{R}$ is a C^2 λ-convex function for some $\lambda \in \mathbb{R}$ bounded from below, $\phi = \mathscr{G}$ as in* (1.101)*, and $\mu_0 = v_0\gamma, M^0_\tau \in \mathcal{P}^r_2(\Omega)$ are the (continuous and discrete) initial data satisfying* (1.49) *and*

$$\mathscr{I}_2(\mu_0|\,\gamma) = \int_\Omega \frac{|\mathrm{D}v_0|^2}{v_0}\,\mathrm{d}\gamma \le \mathscr{J}_0 := \limsup_{\tau\downarrow 0} \mathscr{I}_2(M^0_\tau|\,\gamma) < +\infty, \quad \limsup_{\tau\downarrow 0} \mathsf{m}_2(M^0_\tau) < +\infty. \tag{1.102}$$

i) ***Existence and regularity of Generalized Minimizing Movements*** *For every step $\tau > 0$ the variational scheme* (1.44) *admits a unique solution $\{M^n_\tau\}_{n\in\mathbb{N}}$ and each infinitesimal sequence τ_k of time steps admits a subsequence (still denoted by τ_k), a non increasing map $\mathscr{J} : (0,+\infty) \to \mathbb{R}$ and an absolutely continuous function $\mathscr{H} : (0,+\infty) \to \mathbb{R}$ such that*

$$\overline{M}_{\tau_k,t} \to \mu_t \quad \textit{narrowly in } \mathcal{P}(\Omega) \quad \forall\,t \in [0,+\infty), \tag{1.103}$$

$$\mathscr{J}_t \le \mathscr{J}_0, \quad \mathscr{I}_2(\mu_t|\,\gamma) \le \mathscr{J}_t := \lim_{k\uparrow+\infty} \mathscr{I}_2(\overline{M}_{\tau_k,t}|\,\gamma) < +\infty \quad \forall\,t > 0, \tag{1.104}$$

$$\mathscr{H}(\mu_t|\,\gamma) \le \mathscr{H}_t = \lim_{k\uparrow+\infty} \mathscr{H}(\overline{M}_{\tau_k,t}|\,\gamma) < +\infty \quad \forall\,t \ge 0. \tag{1.105}$$

$\mu_t = v_t\gamma = s_t^2\gamma \in \mathcal{P}^r_2(\Omega)$ for every $t > 0$ and s satisfies the regularity properties (slightly stronger than (1.59)*)*

$$\begin{gathered} s = \sqrt{v} \in L^2(0,T;W^{2,2}_{\mathrm{loc}}(\overline{\Omega})) \cap L^\infty(0,+\infty;W^{1,2}_\gamma(\Omega)), \quad \partial_{\boldsymbol{n}} s = 0 \quad \textit{on } (\partial\Omega)_\infty, \\ s\Delta_\gamma s \in L^1((0,T);W^{1,1}_{\mathrm{loc}}(\overline{\Omega})) \quad \sqrt{s} \in L^4(0,T;W^{1,4}_{\mathrm{loc}}(\overline{\Omega})), \end{gathered} \tag{1.106}$$

for every $T > 0$, with $T = +\infty$ if $\lambda \ge 0$. All the limit curves μ obtained in this way are the elements of $GMM(\mu_0;\mathscr{G})$, which in particular is not empty.

ii) ***Generalized Minimizing Movement are variational solutions*** *If* $\mu = v\gamma \in GMM(\mu_0, \mathscr{G})$ *then* v *is a variational solution of* $(\mathrm{EE}_{1,\gamma})$ *according to Definition 1.5 and it satisfies*

$$\partial_t v + \mathrm{div}_\gamma \tilde{\boldsymbol{q}} = 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (0,+\infty)), \quad \text{for } \tilde{\boldsymbol{q}} := 2\mathrm{D}(s\Delta_\gamma s) - 4\Delta_\gamma s\,\mathrm{D}s. \tag{1.107}$$

iii) ***Lyapunov inequality*** *The non increasing map* $\mathscr{J}_t$ *satisfies*

$$\dot{\mathscr{J}}_t := \frac{d}{dt}\mathscr{J}_t \le -2\int_\Omega \left|\frac{\tilde{\boldsymbol{q}}_t(x)}{v_t(x)}\right|^2 \mathrm{d}\mu_t(x) \qquad \text{for a.e. } t \in (0,+\infty), \tag{1.108}$$

or equivalently

$$\frac{1}{2}\mathscr{J}_{t_1} + \int_{t_0}^{t_1}\int_\Omega \left|\frac{\tilde{\boldsymbol{q}}_t(x)}{v_t(x)}\right|^2 \mathrm{d}\mu_t(x)\,\mathrm{d}t \le \frac{1}{2}\mathscr{J}_{t_0} \qquad \text{for every } 0 < t_0 < t_1. \tag{1.109}$$

iv) ***Entropy and Fisher dissipation*** *The functions* $\mathscr{H}_t$ *and* $\mathscr{J}_t$ *are related by the Entropy-Entropy dissipation inequalities*

$$-\frac{d}{dt}\mathscr{H}_t \ge \mathscr{K}_{-1}(\mu_t|\,\gamma) + \lambda\,\mathscr{J}_t, \qquad \text{for a.e. } t > 0, \tag{1.110}$$

$$\Big(-\frac{1}{4}\frac{d}{dt}\mathscr{J}_t^2\Big)^{1/2} \ge \mathscr{K}_{-1}(\mu_t|\,\gamma) + \lambda\,\mathscr{J}_t, \qquad \text{for a.e. } t > 0. \tag{1.111}$$

v) ***Asymptotic decay when*** $\lambda > 0$ *When* $\lambda > 0$ *and we normalize* γ *so that* $\gamma(\Omega) = 1$*, then* $\mathscr{H}_t$ *is a nonincreasing function satisfying the "Log-Sobolev-like" inequality*

$$\mathscr{H}_t \le \frac{1}{2\lambda}\mathscr{J}_t \quad \forall\, t \ge 0, \tag{1.112}$$

and we have

$$\mathscr{H}(\mu_t|\,\gamma) \le \mathscr{H}_t \le \mathscr{H}_0 \mathrm{e}^{-2\lambda^2 t}, \tag{1.113}$$

$$\mathscr{I}_2(\mu_t|\,\gamma) \le \mathscr{J}_t \le \mathscr{J}_0 \mathrm{e}^{-2\lambda^2 t}. \tag{1.114}$$

When V is uniformly convex we can relax the assumption on the initial datum (finite relative Entropy instead of relative Information) by showing that the solution exhibits a regularizing effect.

Theorem 7 (Regularizing effect). *Let us suppose that* Ω *is an open* convex *subset of* $\mathbb{R}^d$*,*

$$V \textbf{ satisfies the } \lambda\textbf{-convexity condition for some } \lambda > 0, \tag{1.115}$$

$\gamma(\Omega) = 1$*, and* $\mu_0 = v_0\gamma, M^0_\tau \in \mathcal{P}_2^r(\Omega)$ *are the (continuous and discrete) initial data satisfying the convergence and the finite Entropy condition*

$$M^0_\tau \rightharpoonup \mu_0 \quad \text{narrowly in} \quad \mathcal{P}(\Omega), \quad \mathscr{H}(\mu_0|\,\gamma) \le \limsup_{\tau\downarrow 0}\mathscr{H}(M^0_\tau|\,\gamma) = \mathscr{H}_0 < +\infty. \tag{1.116}$$

Then

all the same conclusions of Theorem 6 still hold*;*

in particular for every $t > 0$ *the Fisher information* $\mathscr{I}_2(\mu_t|\,\gamma)$ *is finite and satisfies*

$$\mathscr{I}_2(\mu_t|\,\gamma) \le \mathscr{J}_t \le \frac{1}{\lambda t}\mathscr{H}_0 \qquad \forall\, t > 0. \tag{1.117}$$

If moreover V ***has a super-quadratic growth,*** *i.e.*

$$\lim_{|x|\to\infty}\frac{V(x)}{|x|^2} = +\infty, \tag{1.118}$$

then

$$\lim_{k\to\infty} W(\overline{M}_{\tau_k,t}, \mu_t) = 0, \qquad \lim_{k\to\infty}\mathscr{H}(\overline{M}_{\tau_k,t}|\,\gamma) = \mathscr{H}(\mu_t|\,\gamma) = \mathscr{H}_t \quad \forall\, t > 0. \tag{1.119}$$

Remark 1.11. We have already observed that when V has an at most quadratic growth and satisfies (1.5a,b,c,d,e), then it is possible to work either with the perturbed functional $\mathscr{F}^f$ involving the Lebesgue measure or with the relative Fisher information w.r.t. γ $\mathscr{G}(\cdot) = \frac{1}{2}\mathscr{I}_2(\cdot|\,\gamma)$. In this case the first approach yields more refined convergence results and there is no need to introduce the auxiliary limit functions $\mathscr{H}_t$, $\mathscr{J}_t$.

By using γ as a reference measure and the relative density v as natural unknown, more general potentials V can be considered; however, in the most relevant case of an unbounded domain Ω, global convergence results for the approximating variational scheme are available only when V has a super-quadratic growth. In the intermediate case of a potential V which neither satisfies conditions (1.5a,b,c,d,e) nor exhibits a super-quadratic growth, only local convergence results are available at the present time.

Remark 1.12. When $M^0_\tau \equiv \mu_0$ for every $\tau > 0$, we have

$$\mathscr{H}_0 = \mathscr{H}(\mu_0|\,\gamma), \quad \mathscr{J}_0 = \mathscr{I}_2(\mu_0|\,\gamma).$$

In particular, (1.113), (1.114) read as

$$\mathscr{H}(\mu_t|\,\gamma) \le \mathscr{H}(\mu_0|\,\gamma)e^{-2\lambda^2 t}, \qquad \mathscr{I}_2(\mu_t|\,\gamma) \le \mathscr{I}_2(\mu_0|\,\gamma)e^{-2\lambda^2 t}. \tag{1.120}$$

1.11. **Plan of the paper and final remarks.** Apart from the next section, where we will collect some preliminary material (on the relative Entropy and Fisher information, the Wasserstein distance, the gradient flows and the Minimizing Movements in Wasserstein spaces, and the Fokker-Planck equation), the rest of the paper is devoted to the proof of the main theorems stated in this introduction.

We have tried to divide the proof in main steps, some of which are of independent interest, by following a general strategy which in principle could be adapted to many other examples. Let us quote here the main ideas, mainly referring to the case of the functional $\mathscr{F}^f$; the treatment of $\mathscr{G}(\cdot) = \frac{1}{2}\mathscr{I}_2(\cdot|\,\gamma)$ is completely similar.

Section 2.5 is our starting point: it provides the abstract compactness and convergence result for the variational scheme (1.44) and it reduces the whole evolution problem to the study of a single stationary problem (1.44) and to the characterization of a sort of "Wasserstein subdifferential" $\partial_\ell\mathscr{F}^f$, attached to the Fisher information functional; this general approach has been developed in [AGS05]. Following this strategy it is possible to show that any generalized Minimized Movement $\mu = u\mathcal{L}^d$ associated with the functional $\phi = \mathscr{F}^f$ satisfies a system like (1.25a,b,c)

$$\partial_t u + \operatorname{div}(u\boldsymbol{v}) = 0 \quad \text{in } \Omega_\infty, \qquad \boldsymbol{v}_t = -\partial_\ell\mathscr{F}^f(\mu_t) \quad \text{for a.e. } t > 0. \tag{1.121}$$

Recalling (1.25c), one can formally expect that

$$\partial_\ell\mathscr{F}^f(\mu) = \mathrm{D}\Big(\frac{\delta\mathscr{F}^f(u\mathcal{L}^d)}{\delta u}\Big) \qquad \text{if } \mu = u\mathcal{L}^d. \tag{1.122}$$

Section 4 contains the computation of this subdifferential, which justifies (1.122): here we follow the ideas of [JKO98], by taking a sort of "first variation" of the minimizing functional along the flow generated by a smooth vector field with compact support (cf. paragraph 4.1). It turns out that the formula for $\partial_\ell\mathscr{F}^f(\mu_t)$ is strictly related to formulation (1.53) of (EE_3), since

$$\boldsymbol{v} = -\partial_\ell\mathscr{F}^f(\mu_t) \quad \Rightarrow \quad \int_\Omega \boldsymbol{v}\cdot\mathrm{D}\zeta\, u\,\mathrm{d}x = \int_\Omega \Big(4\mathrm{D}^2\zeta\mathrm{D}\sqrt{u}\cdot\mathrm{D}\sqrt{u} + \mathrm{D}\Delta\zeta\cdot\mathrm{D}u - u\mathrm{D}f\cdot\mathrm{D}\zeta\Big)\,\mathrm{d}x, \tag{1.123}$$

for every smooth test function $\zeta \in C^\infty_c(\mathbb{R}^d)$ with $\partial_{\boldsymbol{n}}\zeta = 0$ on $\partial\Omega$.

The main technical difficulty is to prove the closure of the first variation formula (1.123) with respect to weak convergence of u; we also need higher regularity for $\sqrt{u}$ in order to write the formulation proposed in Definition 1.5. In order to get an a priori bound on the second order derivatives of u, we will exploit the Lyapunov identity for the Entropy (1.15). Its differential counterpart (we consider here the simplest case $f \equiv 0$) reads formally as

$$\int_\Omega \boldsymbol{v}\cdot\mathrm{D}u\,\mathrm{d}x = \int_\Omega \big|\mathrm{D}^2\log u\big|^2 u\,\mathrm{d}x \quad \boldsymbol{v} \text{ given by (1.123)}. \tag{1.124}$$

In order to prove (1.124) we need a correct definition of the right hand side and we should justify difficult integration by parts under low regularity assumptions on u; moreover, we also need to extract more information from the second order derivatives of the logarithm of u. These improvements are performed in two independent steps, in Section 3 and in Section 5.

Section 5 provides new *a priori* estimates, by taking the "first variation" of $\mathscr{F}^f$ along the flow generated by the Fokker-Planck equation (this part also justifies the introductory remarks of § 2.6; notice that the Fokker-Planck equation is exactly the Wasserstein gradient flow of the Entropy, which is the Lyapunov functional involved in (1.15)). This procedure, which is also strictly related to the contribution of [BÉ85], [OV00], and [DPL04a, DPL04b], provides the stationary counterpart of (1.15), in particular a uniform control, of

$$\int_\Omega \left|\mathrm{D}^2 \log u\right|^2 u \,\mathrm{d}x, \qquad \text{or, more generally,} \qquad \int_\Omega |\mathrm{D}^2 \log v|^2\, v \,\mathrm{d}\gamma. \tag{1.125}$$

The Fokker-Planck flow (at least in a bounded domain) has also the advantage of a smooth and strictly positive regularization of the discrete solution, and therefore allows for many calculations, which would be delicate in a weaker setting.

Concerning this wide subject, a recent study [MV00] reveals in fact the various links between the Fokker-Planck equation and differential inequalities. This part of the paper takes advantage of a well–known technique, first used in connection with the logarithmic entropy and Fisher information functional in two pioneering papers [Bla65, McK66], where the smoothing of a probability density function by means of the heat kernel is used to recover refined inequalities. In [Bla65] N.M. Blachman presented a simple proof of Shannon's convolution inequality previously obtained by A. Stam [Sta59], showing convolution inequalities for Fisher information, and subsequently for the logarithmic entropy using their link in terms of the heat kernel. The key idea is that convolution inequalities for Fisher information (with respect to the logarithmic entropy) are relatively easy to prove thanks to its quadratic structure. A detailed study of the properties of Fisher information in connection with logarithmic Sobolev inequalities can be found in [Car91]. In particular, E. Carlen shows that the Blachman–Stam inequality is very refined, in that it implies the logarithmic Sobolev inequality.

Section 5.3 contains the analysis of the "weak closure" of the graph of the Wasserstein subdifferential (given by (1.123)) of the Fisher information: thanks to the previous "a priori" estimates and to the lower semicontinuity results of Section 3, one gains enough compactness to characterize the weak limit of the nonlinear differential operator (1.57a).

Section 3 collects a systematic study of the relationships between "logarithmic" second order functionals like the ones of (1.125) and provides a relevant new estimate of the second order derivatives of the square root $\sqrt{v}$ in terms of them, which in the case of Neumann boundary conditions $\partial_{\boldsymbol{n}} v = 0$ on $\partial\Omega$ reads as (cf. (3.24) and (3.25) for $\alpha = -1$)

$$\int_\Omega \Big(2\big|\mathrm{D}^2\sqrt{v}\big|^2 + \big(\Delta\sqrt{v}\big)^2\Big)\,\mathrm{d}\gamma \le \frac{1}{4}\int_\Omega \Big(2\big|\mathrm{D}^2\log v\big|^2 + \big(\Delta_\gamma \log v\big)^2\Big) v\,\mathrm{d}\gamma \tag{1.126}$$

and

$$4^4 \int_\Omega \big|\mathrm{D}\sqrt[4]{v}\big|^4\,\mathrm{d}x \le 3\int_\Omega \Big(2\big|\mathrm{D}^2\log v\big|^2 + \big(\Delta_\gamma \log v\big)^2\Big) v\,\mathrm{d}\gamma. \tag{1.127}$$

The section is of intrinsic interest and its results, which are strictly related to the contribution of [LT95], are independent of the rest of the paper: therefore we decided to put it before developing all the other more specific arguments, just after the preliminary remarks.

Observe that when V satisfies a Lipschitz condition on Ω, then the terms containing $\Delta_\gamma \log v$ (and therefore $\mathrm{D}V\cdot\mathrm{D}\log v$) in the right hand side of (1.126) and (1.127) can be controlled by the squared norm of the Hessian $\mathrm{D}^2\log v$ and by the Fisher Information. For general potentials V, those terms can be only locally bounded by $\mathrm{D}^2 \log v$: the lack of control on $\Delta_\gamma \log v$ is one of the main technical difficulties and it causes the local nature of the convergence results of Theorems 6 and 7.

Let us also notice that the possible importance of higher–order functionals obtained as higher–order derivatives in time of the logarithmic entropy of a probability density function smoothed by means of the heat kernel was questioned by H.P. McKean [McK66], who first considered Fisher's information in connection with the large–time behavior of a kinetic model.

A new general and systematic way to attack the problem of constructing entropies for higher-order nonlinear PDE's has been recently introduced by [JM06].

Sections 6 and 7 connect all the previous contribution and provide a detailed guide to the final steps of the proofs of our main theorems.

Final remarks and open problems. This paper represents only a first step towards the investigation of gradient flows of first order functionals w.r.t. the Wasserstein distance: up to now the existing theory covers the case of R.J. McCann geodesically/displacement convex functionals (see Definition 2.14), whose main examples are only confined to functionals which does not depend on the gradient of their argument. It would be interesting to find first order examples of geodesically convex functionals, or other general principles which could be useful to study their gradient flow.

We are collecting here only a few selection of open problems which could deserve further investigation.

- In the present case of the Fisher information, finer regularity properties of the solution could be deduced by the energy estimate (1.18), in particular by the L^2 estimate of the Wasserstein subdifferential of a measure. This investigation seems also related to the possibility to prove other higher order "logarithmic Sobolev" inequalities, extending what we will present in Section 3 to third order functionals. The ideas recently introduced by [JM06] could be very useful at this respect.
- The regularity issue is also related to the uniqueness question, which is known only for regular and strictly positive solutions [JP00, BLS94].
- Further study of the long time behaviour of the solution of (EE_1) could also be interesting to get solutions V of (1.5a) satisfying (1.5d), which corresponds to the equilibrium state of the equation.
- It is not clear if the asymptotic behaviour of (EE_1) is also related to the perturbation approach of [CL04] to Logarithmic Sobolev inequalities.
- The stability of the solutions to ($\mathrm{EE}_{1,\gamma}$) with respect to perturbation of γ is another interesting question: it should not be difficult to use a regularization argument to deal with potentials V with $\lim_{x\to\partial\Omega} V(x) = +\infty$. The case of equations settled in nonconvex open domains and potentials whose second derivatives are not bounded form below are also interesting.
- Concerning other kind of boundary conditions (see the recent contributions [DGJ06, GJT06] in the 1-dimensional case), it would not be too difficult to adapt our arguments to e.g. Dirichlet boundary condition on u instead of (1.2), whereas (1.3) is a crucial condition for the "Wasserstein" approach, since it guarantees the conservation of the total mass.
- It is not clear if the Lyapunov inequalities (1.86) and (1.108) could be improved to obtain identities (with moreover the identification $\mathscr{J}_t = \mathscr{I}_2(\mu_t|\,\gamma)$): this would also imply the absolute continuity of the Fisher information along the trajectories of its gradient flow, preventing jumps during the evolutions.
- Finally, it would be interesting to know if the convergence results of Theorems 6 and 7 could be reinforced, obtaining at least convergence of the quadratic moment of the approximating solutions and of their relative entropy.

2. Notation and preliminary results

In the following table we resume the notation used without further explanation throughout the text, Ω being a given open subset of $\mathbb{R}^d$:

$\mathscr{B}(\Omega)$	The collection of the Borel subsets of Ω
$C^0_b(\Omega)$	Continuous and bounded real functions defined in Ω
$\mathcal{L}^d$	The Lebesgue measure in $\mathbb{R}^d$
$\mathcal{H}^{d-1}$	The Hausdorff $(d-1)$ dimensional measure
$\gamma = e^{-V}\mathcal{L}^d$	The invariant measure, see § 2.2
$L^p_\gamma(\Omega)$ $\big(L^p_\gamma(\Omega;\mathbb{R}^k)\big)$	the Lebesgue spaces of p-summable real ($\mathbb{R}^k$-valued) functions w.r.t. the measure γ (§2.2)
$W^{k,p}_\gamma(\Omega)$	the weighted Sobolev spaces (§2.2)
$\mathrm{div}_\gamma, \Delta_\gamma, \tilde\partial_i, \tilde\partial^2_{ij}$	the partial differential operators induced by γ, see (1.10)
$\mathcal{M}_{\mathrm{loc}}(\Omega)$	(the space of all) real Radon Measures defined in Ω (§2.1)
$\mathcal{M}^+_{\mathrm{loc}}(\Omega)$	nonnegative Radon measures (§2.1)
$\mathcal{M}(\Omega)$	Real Borel measures with finite total variation (§2.1)
$\big[\mathcal{M}_{\mathrm{loc}}(\Omega)\big]^k, \big[\mathcal{M}(\Omega)\big]^k, \ldots$	The corresponding spaces of vector measures (§2.1)
$\mathcal{P}(\Omega)$ $\big(\mathcal{P}_2(\Omega)\big)$	Probability measures (with finite quadratic moment) (2.1)
$\mathcal{P}^r(\Omega), \mathcal{P}^r_2(\Omega)$	Probability measures $\ll \mathcal{L}^d$
$\mu_n \rightharpoonup \mu$ in $\mathcal{P}(\Omega)$	Narrow convergence of probability measures, (2.3) and Remark 2.1
$\dfrac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}\nu}$	the density of $\boldsymbol{\mu}$ w.r.t. ν
$\mathsf{m}^2_2(\mu)$	the quadratic moment of μ (1.34)
$\boldsymbol{r}_\#\mu$	Push forward of the measure μ through the map $\boldsymbol{r}$, (2.55)
$\boldsymbol{i}$	The identity map
$W(\mu_1,\mu_2)$	the L^2-Wasserstein distance between the measures $\mu_1,\mu_2 \in \mathcal{P}_2(\Omega)$, § 2.4
$T_o(\mu_1,\mu_2)$	the optimal transport map between $\mu_1,\mu_2 \in \mathcal{P}^r_2(\Omega)$
$\lvert\mathsf{A}\rvert$	Euclidean norm of a matrix $\mathsf{A} \in \mathbb{M}^{d\times d}$, $\lvert\mathsf{A}\rvert^2 = \sum_{i,j=1}^d \lvert\mathsf{A}_{ij}\rvert^2$
$\lVert\mathsf{A}\rVert$	Operator norm of $\mathsf{A} \in \mathbb{M}^{d\times d}$, $\lVert\mathsf{A}\rVert = \sup\{\lvert\mathsf{A}\boldsymbol{\xi}\rvert : \boldsymbol{\xi} \in \mathbb{R}^d,\ \lvert\boldsymbol{\xi}\rvert \le 1\}$

2.1. **Measures.** Recall that a positive Borel measure ν on Ω is *Radon* if it is finite on the compact subset of Ω; the space of positive Radon measure is denoted by $\mathcal{M}^+_{\mathrm{loc}}(\Omega)$. If $\nu(\Omega) < +\infty$ (resp. $\nu(\Omega) = 1$) we say that ν is a *finite* (resp. *probability*) measure in $\mathcal{M}^+(\Omega)$ (resp. $\mathcal{P}(\Omega)$). $\mathcal{P}^r(\Omega)$ denotes the subset of all the measures in $\mathcal{P}(\Omega)$ which are *absolutely continuous* with respect to the Lebesgue measure $\mathcal{L}^d$. The sets $\mathcal{P}_2(\Omega), \mathcal{P}^r_2(\Omega)$ are defined as

$$\mathcal{P}_2(\Omega) := \Big\{\mu \in \mathcal{P}(\Omega) : \mathsf{m}^2_2(\mu) = \int_\Omega |x|^2 \,\mathrm{d}\mu(x) < +\infty\Big\}, \qquad \mathcal{P}^r_2(\Omega) := \mathcal{P}_2(\Omega) \cap \mathcal{P}^r(\Omega). \tag{2.1}$$

A real (or $\mathbb{R}^k$-valued) set function μ defined on the relatively compact Borel subsets of Ω that is a (finite) measure on each compact set $K \subset\subset \Omega$ is called a real (resp. $\mathbb{R}^k$-valued) Radon measure: the corresponding space is $\mathcal{M}_{\mathrm{loc}}(\Omega)$ (resp. $\big[\mathcal{M}_{\mathrm{loc}}(\Omega)\big]^k$: vector Radon measures $\boldsymbol{\mu}$ can be identified with a k-tuple of Radon measures $\big\{\boldsymbol{\mu}^{(j)}\big\}_{j=1}^k$). On $\mathcal{M}_{\mathrm{loc}}(\Omega)$ (and $\mathcal{M}^+_{\mathrm{loc}}(\Omega)$) we consider the topology of the weak convergence in the sense of distributions, i.e.

$$\mu_n \rightharpoonup \mu \quad \text{in} \quad \mathcal{M}_{\mathrm{loc}}(\Omega) \qquad \Leftrightarrow \qquad \lim_{n\uparrow+\infty} \int_\Omega \zeta(x)\,\mathrm{d}\mu_n(x) = \int_\Omega \zeta(x)\,\mathrm{d}\mu(x) \quad \forall\, \zeta \in C^\infty_c(\Omega). \tag{2.2}$$

Analogously $\boldsymbol{\mu}_n \rightharpoonup \boldsymbol{\mu}$ in $\big[\mathcal{M}_{\mathrm{loc}}(\Omega)\big]^k$ if $\boldsymbol{\mu}^{(j)}_n \rightharpoonup \boldsymbol{\mu}^{(j)}$ for each component $\boldsymbol{\mu}^{(j)}_n, \boldsymbol{\mu}^{(j)}$, $j = 1, \ldots, k$.

According to the probabilistic terminology, if $\mu_n \in \mathcal{P}(\Omega)$ we say that μ_n *narrowly* converges to μ if

$$\lim_{n\to\infty} \int_\Omega \zeta(x)\,\mathrm{d}\mu_n(x) = \int_\Omega \zeta(x)\,\mathrm{d}\mu(x) \quad \forall\, \zeta \in C^0_b(\Omega); \tag{2.3}$$

in this case $\mu \in \mathcal{P}(\Omega)$ and we will write $\mu_n \rightharpoonup \mu$ in $\mathcal{P}(\Omega)$.

An easy application of Prokhorov's Theorem yields

$$\mu_n \in \mathcal{P}(\mathbb{R}^d), \quad \sup_n \mathsf{m}_2(\mu_n) < +\infty \qquad \Rightarrow \qquad \exists\, \mu_{n_k}, \mu: \quad \mu_{n_k} \rightharpoonup \mu \quad \text{in } \mathcal{P}(\mathbb{R}^d). \tag{2.4}$$

Finally, if $\mu_n, \mu \in \mathcal{M}^+_{\rm loc}(\Omega)$ and $\mu_n \rightharpoonup \mu$ in $\mathcal{M}_{\rm loc}(\Omega)$ then

$$\liminf_{n\uparrow+\infty} \int_\Omega f(x)\,\mathrm{d}\mu_n(x) \ge \int_\Omega f(x)\,\mathrm{d}\mu(x) \tag{2.5}$$

for every proper, lower semicontinuous function $f : \Omega \to [0, +\infty]$; in particular

$$\liminf_{n\uparrow+\infty} \mathsf{m}_2(\mu_n) \ge \mathsf{m}_2(\mu). \tag{2.6}$$

Remark 2.1 (Trivial extension and convergence comparison). Since Ω is an open (thus Borel) subset of $\mathbb{R}^d$, we can identify $\mathcal{M}^+(\Omega)$ with the subset of $\mathcal{M}^+(\mathbb{R}^d)$ of the measures concentrated in Ω, i.e.

$$\mathcal{M}^+(\Omega) \approx \{\mu \in \mathcal{M}^+(\mathbb{R}^d) : \mu(\mathbb{R}^d \setminus \Omega) = 0\big\}. \tag{2.7}$$

If μ_n is a sequence in $\mathcal{P}(\Omega) \subset \mathcal{M}^+(\Omega)$ we thus have at least the following four notions of convergence at our disposal (each of the following items induces a finer topology than the previous ones):
- the convergence in the sense of distribution in $\mathcal{D}'(\Omega)$ (2.2),
- the distributional convergence in $\mathcal{D}'(\mathbb{R}^d)$ inherited from the identification (2.7),
- the narrow convergence in $\mathcal{P}(\mathbb{R}^d)$ inherited from the identification (2.7) (i.e. against bounded test functions which admits a continuous extension to $\mathbb{R}^d$),
- the narrow convergence in $\mathcal{P}(\Omega)$ (2.3).

If the limit μ is still a probability measure in $\mathcal{P}(\Omega)$, *all these notions of convergence are in fact equivalent* to the (a priori strongest) narrow convergence.

2.2. **The invariant measure and the weighted Sobolev spaces.** As detailed in the introduction, we will at least assume that

$$\begin{gathered} V : \mathbb{R}^d \to \mathbb{R} \quad \text{is a } C^2 \text{ semi-convex function with } V_{\min} := \inf_{\mathbb{R}^d} V > -\infty, \\ g := e^{-V}, \quad \gamma := g \cdot \mathcal{L}^d, \end{gathered} \tag{2.8}$$

i.e. there exists some $\lambda \in \mathbb{R}$ such that $V - \frac{1}{2}\lambda|\cdot|^2$ is a convex function, which is equivalent to

$$\mathrm{D}^2 V(x)\boldsymbol{\zeta}\cdot\boldsymbol{\zeta} \ge \lambda\,|\boldsymbol{\zeta}|^2 \quad \forall\, x, \boldsymbol{\zeta} \in \mathbb{R}^d. \tag{2.9}$$

Observe that the restrictions of V and g to bounded sets are bounded and Lipschitz continuous, in particular γ is a Radon measure absolutely continuous w.r.t. $\mathcal{L}^d$. For a given open set Ω, we denote by $L^p_\gamma(\Omega)$ the usual Lebesgue space w.r.t. the (restriction to Ω of the) measure γ; since $0 < \inf_{\Omega'} g \le \sup_{\Omega'} g < +\infty$ for every bounded set $\Omega' \subset \Omega$, we have

$$L^p_{\rm loc}(\Omega) = L^p_{\gamma,\,\rm loc}(\Omega); \quad L^p(\Omega) = L^p_\gamma(\Omega) \quad \text{if } \Omega \text{ is bounded}. \tag{2.10}$$

If a function v belongs to $W^{k,p}_{\rm loc}(\Omega)$ and its derivatives up to the order k are in $L^p_\gamma(\Omega)$ we say that $v \in W^{k,p}_\gamma(\Omega)$. In particular

$$W^{1,p}_\gamma(\Omega) := \Big\{v \in W^{1,p}_{\rm loc}(\Omega) : \int_\Omega \Big(|v|^p + |\mathrm{D}v|^p\Big)\,\mathrm{d}\gamma < +\infty\Big\}, \tag{2.11}$$

$$W^{2,p}_\gamma(\Omega) := \Big\{v \in W^{2,p}_{\rm loc}(\Omega) : \int_\Omega \Big(|v|^p + |\mathrm{D}v|^p + |\mathrm{D}^2 v|^p\Big)\,\mathrm{d}\gamma < +\infty\Big\}. \tag{2.12}$$

2.3. **Convex functionals with superlinear growth.** If $\boldsymbol{\mu} \in \big[\mathcal{M}_{\rm loc}(\Omega)\big]^k, \nu \in \mathcal{M}^+_{\rm loc}(\Omega)$, and $\boldsymbol{\mu} \ll \nu$, we will denote by

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}\nu} \in L^1_{\nu,\rm loc}(\Omega;\mathbb{R}^k) \quad \text{the density of } \boldsymbol{\mu} \text{ w.r.t. } \nu. \tag{2.13}$$

Let $Q : \mathbb{R}^k \to [0,+\infty]$ be a lower semicontinuous convex function which grows super-linearly at infinity, i.e.

$$\lim_{|\boldsymbol{\xi}|\uparrow+\infty} \frac{Q(\boldsymbol{\xi})}{|\boldsymbol{\xi}|} = +\infty. \tag{2.14}$$

For $\boldsymbol{\mu} \in \big[\mathcal{M}_{\rm loc}(\Omega)\big]^k, \nu \in \mathcal{M}^+_{\rm loc}(\Omega)$ we define the functional

$$\mathscr{Q}(\boldsymbol{\mu}|\nu) := \begin{cases} \displaystyle\int_\Omega Q\left(\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}\nu}(x)\right)\mathrm{d}\nu(x) & \text{if } \boldsymbol{\mu} \ll \nu,\\ +\infty & \text{otherwise.}\end{cases} \tag{2.15}$$

We refer to [GS64, But89, AFP00] for the proof of the following result.

Proposition 2.2 (**Joint lower semicontinuity under distributional convergence**). *If $\boldsymbol{\mu}_n \in \big[\mathcal{M}_{\rm loc}(\Omega)\big]^k$, $\nu_n, \nu \in \mathcal{M}^+_{\rm loc}(\Omega)$ satisfy*

$$\nu_n \rightharpoonup \nu \quad \text{in } \mathcal{M}^+_{\rm loc}(\Omega) \text{ as } n \to \infty, \quad \liminf_{n\to\infty} \mathscr{Q}(\boldsymbol{\mu}_n|\nu_n) < +\infty, \tag{2.16}$$

then there exists a subsequence (still labeled $\boldsymbol{\mu}_n$) and $\boldsymbol{\mu} \in \big[\mathcal{M}_{\rm loc}(\Omega)\big]^k$ such that $\boldsymbol{\mu}_n \rightharpoonup \boldsymbol{\mu} \in \big[\mathcal{M}_{\rm loc}(\Omega)\big]^k$, $\boldsymbol{\mu} \ll \nu$, and

$$\mathscr{Q}(\boldsymbol{\mu}|\nu) \le \liminf_{n\uparrow+\infty} \mathscr{Q}(\boldsymbol{\mu}_n|\nu_n). \tag{2.17}$$

If moreover $k = 1$, $\boldsymbol{\mu}_n = \mu_n, \nu_n, \nu \in \mathcal{P}(\Omega)$ with $\nu_n \rightharpoonup \nu$ in $\mathcal{P}(\Omega)$, then $\boldsymbol{\mu} = \mu \in \mathcal{P}(\Omega)$ and $\mu_n \rightharpoonup \mu$ narrowly in $\mathcal{P}(\Omega)$.
Finally, if $\nu_n \equiv \nu$ then

$$\frac{\mathrm{d}\boldsymbol{\mu}_n}{\mathrm{d}\nu} \rightharpoonup \frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}\nu} \quad \textit{weakly in } L^1_\nu(\Omega;\mathbb{R}^k). \tag{2.18}$$

Lebesgue densities. The simplest example is provided by $Q(\boldsymbol{\xi}) := |\boldsymbol{\xi}|^p$ for $p > 1$, thus obtaining the functional

$$\mathscr{L}_p\big(\boldsymbol{\mu}|\nu\big) := \int_\Omega \left|\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}\nu}(x)\right|^p \mathrm{d}\nu(x) \quad \text{if } \boldsymbol{\mu} \ll \nu. \tag{2.19}$$

In that case $\mathscr{L}_p\big(\boldsymbol{\mu}|\nu\big) < +\infty$ if and only if $\boldsymbol{\mu} = \boldsymbol{w} \cdot \nu$ for some $\boldsymbol{w} \in L^p_\nu(\Omega;\mathbb{R}^k)$.
Relative entropy. Another interesting application of the above result relies in the definition of the Relative Entropy $\mathscr{H}(\mu|\,\gamma)$

$$\mathscr{H}(\mu|\,\gamma) := \int_\Omega \frac{\mathrm{d}\mu}{\mathrm{d}\gamma}(x)\,\log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\gamma}(x)\right)\mathrm{d}\gamma(x) \quad \text{if } \mu \ll \gamma, \tag{2.20}$$

of a probability measure $\mu \in \mathcal{P}^r_2(\Omega)$. When γ is a probability measure, then $\mathscr{H}(\mu|\,\gamma)$ can be expressed in the form (2.15) for the choice

$$Q(\xi) := \begin{cases} \xi\log\xi - \xi + 1 & \text{if } \xi > 0,\\ 1 & \text{if } \xi = 0,\\ +\infty & \text{if } \xi < 0,\end{cases} \tag{2.21}$$

thus showing that $\mathscr{H}(\mu|\,\gamma) \ge 0$. More generally, if γ is a positive finite measure with $\gamma(\Omega) > 0$, $\mu = v\gamma$, and $\tilde\gamma := \gamma/\gamma(\Omega)$, we easily get

$$\mathscr{H}(\mu|\,\gamma) = \int_\Omega v\log v\,\mathrm{d}\gamma = \int_\Omega \big(v\gamma(\Omega)\big)\log\big(v\gamma(\Omega)\big)\,\mathrm{d}\tilde\gamma - \log\gamma(\Omega) = \mathscr{H}(\mu|\,\tilde\gamma) - \log\gamma(\Omega), \tag{2.22}$$

so that

$$\mathscr{H}(\mu|\,\gamma) \ge -\log\gamma(\Omega). \tag{2.23}$$

When $\gamma \in \mathcal{M}^+_{\rm loc}(\Omega)$ is not finite, in order to show that (2.20) is well defined we introduce the auxiliary weights

$$h_\vartheta := e^{-\frac{\vartheta}{2}|x|^2 - c_\vartheta}, \quad c_\vartheta := \log\Big(\int_\Omega e^{-\frac{\vartheta}{2}|x|^2}\,\mathrm{d}\gamma\Big), \qquad \gamma_\vartheta := h_\vartheta\cdot\gamma, \qquad \vartheta>0, \tag{2.24}$$

where ϑ has been chosen so that (recall that the density g of γ is bounded)

$$\int_\Omega h_\vartheta(x)\,\mathrm{d}\gamma(x) = 1, \quad \text{i.e.}\quad \gamma_\vartheta\in\mathcal{P}(\Omega).$$

For $\mu = v\cdot\gamma$ and $\vartheta>0$ we obtain

$$\begin{aligned}\mathscr{H}(\mu|\,\gamma) &= \int_\Omega v\log v\,\mathrm{d}\gamma = \int_\Omega v\log\big(v/h_\vartheta\big)\,\mathrm{d}\gamma - \frac{\vartheta}{2}\int_\Omega |x|^2 v\,\mathrm{d}\gamma - c_\vartheta\\ &= \int_\Omega \frac{v}{h_\vartheta}\log\big(v/h_\vartheta\big)\,\mathrm{d}\gamma_\vartheta - \frac{\vartheta}{2}\mathsf{m}_2^2(\mu) - c_\vartheta = \mathscr{H}(\mu|\,\gamma_\vartheta) - \frac{\vartheta}{2}\mathsf{m}_2^2(\mu) - c_\vartheta \ge -\frac{\vartheta}{2}\mathsf{m}_2^2(\mu) - c_\vartheta.\end{aligned} \tag{2.25}$$

In particular, in the case of the Lebesgue Measure we have

$$\mathscr{H}(\mu|\,\mathcal{L}^d) \ge -\frac{\vartheta}{2}\mathsf{m}_2^2(\mu) + \frac{d}{2}\log\big(\vartheta/2\pi\big), \qquad \mathscr{H}(\mu|\,\mathcal{L}^d) + \pi\mathsf{m}_2^2 \ge 0. \tag{2.26}$$

Lemma 2.3 (**Lower semicontinuity of the relative entropy**). *Let $\gamma_n,\gamma\in\mathcal{M}^+_{\rm loc}(\mathbb{R}^d)$ be satisfying*

$$\gamma(\mathbb{R}^d)>0, \quad \gamma_n\rightharpoonup\gamma \quad \textit{in } \mathcal{M}_{\rm loc}(\mathbb{R}^d), \quad \textit{with} \quad m_\vartheta := \sup_n\int_\Omega e^{-\frac{\vartheta}{2}|x|^2}\,\mathrm{d}\gamma_n(x) < +\infty \quad \forall\,\vartheta>0, \tag{2.27}$$

and let $\mu_n = v_n\cdot\gamma_n\in\mathcal{P}_2(\mathbb{R}^d)$ be a sequence satisfying

$$\sup_n \mathsf{m}_2(\mu_n) < +\infty, \quad \sup_n \mathscr{H}(\mu_n|\,\gamma_n) < +\infty; \tag{2.28}$$

then there exist $\mu = v\cdot\gamma\in\mathcal{P}_2(\mathbb{R}^d)$ and a subsequence (still labeled by μ_n) narrowly converging to μ in $\mathcal{P}(\mathbb{R}^d)$. Moreover

$$\liminf_{n\to\infty}\mathscr{H}(\mu_n|\,\gamma_n) \ge \mathscr{H}(\mu|\,\gamma), \qquad \liminf_{n\to+\infty}\mathsf{m}_2(\mu_n)\ge\mathsf{m}_2(\mu). \tag{2.29}$$

Proof. (2.27) and the lower semicontinuity property (2.5) show that

$$e^{c_\vartheta} := \int_{\mathbb{R}^d} e^{-\frac{\vartheta}{2}|x|^2}\,\mathrm{d}\gamma(x) \le m_\vartheta < +\infty \quad \forall\,\vartheta>0. \tag{2.30}$$

For $\vartheta>0$ we introduce the probability measure γ_ϑ as in (2.24) and the corresponding sequence

$$\gamma_{\vartheta,n} := h_{\vartheta,n}\cdot\gamma_n\in\mathcal{P}(\mathbb{R}^d), \quad h_{\vartheta,n}(x) := e^{-\frac{\vartheta}{2}|x|^2 - c_{\vartheta,n}}, \quad c_{\vartheta,n} := \log\Big(\int_{\mathbb{R}^d} e^{-\frac{\vartheta}{2}|x|^2}\,\mathrm{d}\gamma_n\Big). \tag{2.31}$$

By (2.27) it is not difficult to check that

$$\lim_{n\to\infty} c_{\vartheta,n} = c_\vartheta \quad \forall\,\vartheta>0; \tag{2.32}$$

in fact, choosing $0<\vartheta_0<\vartheta$, we can express $c_{\vartheta,n}$ through the formula

$$e^{c_{\vartheta,n}} = \int_{\mathbb{R}^d} e^{-\frac{\vartheta-\vartheta_0}{2}|x|^2} e^{-\frac{\vartheta_0}{2}|x|^2}\,\mathrm{d}\gamma_n(x)$$

where $e^{-\frac{\vartheta_0}{2}|x^2|}\cdot\gamma_n$ are uniformly bounded measures (thanks to (2.27)) weakly converging to $e^{-\frac{\vartheta_0}{2}|x^2|}\cdot\gamma$ in the duality with the continuous functions of $C^0_0(\mathbb{R}^d)$ vanishing at ∞.

We therefore deduce that $\gamma_{\vartheta,n}\rightharpoonup\gamma_\vartheta$ in $\mathcal{P}(\mathbb{R}^d)$. Owing to (2.25) we obtain

$$\mathscr{H}(\mu_n|\,\gamma_{\vartheta,n}) = \mathscr{H}(\mu_n|\,\gamma_n) + \frac{\vartheta}{2}\mathsf{m}_2^2(\mu_n) + c_{\vartheta,n}, \tag{2.33}$$

so that by (2.28)

$$\sup_n \mathscr{H}(\mu_n|\,\gamma_{\vartheta,n}) < +\infty. \tag{2.34}$$

Applying Proposition 2.2 we can find a subsequence (still denoted by μ_n) such that $\mu_n \rightharpoonup \mu$ in $\mathcal{P}(\mathbb{R}^d)$; moreover, by (2.5) and (2.28), μ belongs to $\mathcal{P}_2(\mathbb{R}^d)$ and its quadratic moment satisfies (2.29). (2.17) of Proposition 2.2, (2.32), and (2.33) yield

$$\begin{aligned}\liminf_{n\to\infty} \mathscr{H}(\mu_n|\,\gamma_n) &\geq \liminf_{n\to\infty} \mathscr{H}(\mu_n|\,\gamma_{\vartheta,n}) - \frac{\vartheta}{2} m_\vartheta - c_\vartheta \\ &\overset{(2.17)}{\geq} \mathscr{H}(\mu|\,\gamma_\vartheta) - \frac{\vartheta}{2} m_\vartheta - c_\vartheta \overset{(2.25)}{=} \mathscr{H}(\mu|\,\gamma) - \frac{\vartheta}{2}\big(m_\vartheta - \mathsf{m}_2^2(\mu)\big).\end{aligned}$$

Since ϑ can be chosen arbitrarily small in the previous inequality, we get (2.29). $\square$

Logarithmic gradient and Relative Fisher information. Let $\mu = v \cdot \gamma \in \mathcal{P}^r(\Omega)$ be a given probability measure and let us suppose that for some $p > 1$

$$\exists\, \boldsymbol{\eta} \in L^p_\mu(\Omega;\mathbb{R}^d): \quad -\int_\Omega \mathrm{div}_\gamma \boldsymbol{\zeta}(x)\,\mathrm{d}\mu(x) = \int_\Omega \boldsymbol{\zeta}(x)\cdot\boldsymbol{\eta}(x)\,\mathrm{d}\mu(x) \quad \forall\, \boldsymbol{\zeta} \in C^\infty_c(\Omega;\mathbb{R}^d). \tag{2.35a}$$

Recalling that $\mathrm{div}_\gamma \boldsymbol{\zeta} = \mathrm{div}\,\boldsymbol{\zeta} - \boldsymbol{\zeta}\cdot \mathrm{D}V$, (2.35a) is equivalent to

$$-\int_\Omega \mathrm{div}\,\boldsymbol{\zeta}(x)\,\mathrm{d}\mu(x) = \int_\Omega \boldsymbol{\zeta}(x)\cdot\big(\boldsymbol{\eta}(x) - \mathrm{D}V(x)\big)\,\mathrm{d}\mu(x) \quad \forall\, \boldsymbol{\zeta} \in C^\infty_c(\Omega;\mathbb{R}^d); \tag{2.35b}$$

since V is locally Lipschitz, (2.35b) implies that the distribution associated to μ belongs to $BV_{\mathrm{loc}}(\Omega)$; in particular $\mu \ll \gamma$, $\mu = v\cdot\gamma$ with $v \in L^{d/(d-1)}_{\mathrm{loc}}(\Omega)$. (2.35a,b) yield

$$\mathrm{D}v\cdot\gamma = \boldsymbol{\eta}\cdot\mu = \boldsymbol{\eta}\,v\cdot\gamma, \tag{2.36}$$

so that $v \in W^{1,1}_\gamma(\Omega)$ since

$$\int_\Omega |\mathrm{D}v(x)|\,\mathrm{d}\gamma(x) = \int_\Omega |\boldsymbol{\eta}(x)| v(x)\,\mathrm{d}\gamma(x) = \int_\Omega |\boldsymbol{\eta}(x)|\,\mathrm{d}\mu(x) \leq \|\boldsymbol{\eta}\|_{L^p_\mu(\Omega;\mathbb{R}^d)}.$$

(2.35a,b) thus represent a weak definition of the *logarithmic gradient* (see e.g. [Bog98, Def. 5.2.7])

$$\boldsymbol{\eta} = \frac{\mathrm{D}v}{v} = \text{“}\mathrm{D}\log v\,\text{”}, \tag{2.37}$$

$\boldsymbol{\eta}$ being determined up to μ-negligible sets. The γ-Relative Fisher Information of μ is defined as

$$\mathscr{I}_p(\mu|\,\gamma) := \int_\Omega |\boldsymbol{\eta}(x)|^p\,\mathrm{d}\mu(x) = \int_\Omega \Big|\frac{\mathrm{D}v}{v}\Big|^p\,\mathrm{d}\mu = \mathscr{L}_p\big(\mathrm{D}v\cdot\gamma|\mu\big), \tag{2.38}$$

where $\mathscr{L}_p$ is defined as in (2.19). As usual, we put $\mathscr{I}_p(\mu|\,\gamma) = +\infty$ if $v \notin W^{1,1}_\gamma(\Omega)$. Being μ a probability measure, we easily get

$$\Big(\mathscr{I}_{p_1}(\mu|\gamma)\Big)^{1/p_1} \leq \Big(\mathscr{I}_{p_2}(\mu|\gamma)\Big)^{1/p_2} \quad \text{if } 1 \leq p_1 < p_2 < +\infty. \tag{2.39}$$

Here we recall a list of useful and well known properties of the Relative Fisher Information, whose proofs are briefly sketched.

Lemma 2.4 (**Main properties of the Fisher information**). *If $\mu = v\cdot\gamma \in \mathcal{P}^r(\Omega)$ then*

$$\begin{aligned}&\mathscr{I}_p(\mu|\,\gamma) < +\infty \quad \Leftrightarrow \quad s := \sqrt[p]{v} \in W^{1,p}_\gamma(\Omega),\\ &p^p \int_\Omega |\mathrm{D}s(x)|^p\,\mathrm{d}\gamma(x) = \mathscr{I}_p(\mu|\,\gamma), \quad p\mathrm{D}s = s\,\boldsymbol{\eta},\end{aligned} \tag{2.40}$$

where $\boldsymbol{\eta}$ is the logarithmic gradient of v defined by (2.35a,b). *In particular $v \in L^{d/(d-p)}_{\mathrm{loc}}(\Omega)$ if $p < d$, $v \in L^r_{\mathrm{loc}}(\Omega)$ for every $r \in [1,+\infty)$ if $p = d$, and $v \in C^0(\Omega)$ if $p > d$.*
If $\mu_n = v_n\cdot\gamma = s_n^p\cdot\gamma \in \mathcal{P}^r_2(\Omega)$ is a sequence satisfying

$$\sup_n \Big(\mathsf{m}_2(\mu_n) + \mathscr{I}_p(\mu_n|\,\gamma)\Big) < +\infty, \tag{2.41}$$

and Ω is a Lipschitz open set, then there exists $\mu = v \cdot \gamma \in \mathcal{P}_2^r(\Omega)$ and a subsequence (still labeled by μ_n) such that

$$\mu_n \rightharpoonup \mu \quad \text{in } \mathcal{P}(\Omega), \quad v_n \to v \quad \text{strongly in } L^1_\gamma(\Omega), \quad s_n \to s \quad \text{strongly in } L^p_\gamma(\Omega) \tag{2.42}$$

$$v_n \to v \quad \text{strongly in } \begin{cases} L^r_{\mathrm{loc}}(\overline{\Omega}), & 1 \le r < \frac{d}{d-p} \quad \text{if } p < d, \\ L^r_{\mathrm{loc}}(\overline{\Omega}), & 1 \le r < +\infty \quad \text{if } p = d, \\ L^\infty_{\mathrm{loc}}(\overline{\Omega}) & \quad \text{if } p > d, \end{cases} \tag{2.43}$$

$$\liminf_{n\uparrow+\infty} \mathscr{I}_p(\mu_n|\,\gamma) \ge \mathscr{I}_p(\mu|\,\gamma), \tag{2.44}$$

$$\mathrm{D}s_n \rightharpoonup \mathrm{D}s \quad \text{weakly in } L^p_\gamma(\Omega;\mathbb{R}^d), \quad \mathrm{D}v_n \rightharpoonup \mathrm{D}v \quad \text{weakly in } L^1_\gamma(\Omega;\mathbb{R}^d). \tag{2.45}$$

Finally, if $\limsup_{n\uparrow+\infty} \mathscr{I}_p(\mu_n|\,\gamma) \le \mathscr{I}_p(\mu|\,\gamma)$, *then*

$$\mathrm{D}s_n \to \mathrm{D}s \quad \text{strongly in } L^p_\gamma(\Omega;\mathbb{R}^d), \qquad \mathrm{D}v_n \to \mathrm{D}v \quad \text{strongly in } L^1_\gamma(\Omega;\mathbb{R}^d). \tag{2.46}$$

Proof. The implication "⇒" in (2.40) can be easily proved by introducing the functions

$$s_\varepsilon(x) := (v(x)+\varepsilon)^{1/p}, \quad \varepsilon > 0,$$

which satisfy

$$s_\varepsilon \in W^{1,1}_{\mathrm{loc}}(\Omega), \quad \mathrm{D}s_\varepsilon = \frac{1}{p}(v+\varepsilon)^{1/p-1}\mathrm{D}v = \frac{1}{p}(v+\varepsilon)^{1/p-1} v\boldsymbol{\eta}$$

with

$$p^p \int_\Omega |\mathrm{D}s_\varepsilon|^p \,\mathrm{d}\gamma \le \int_\Omega \Big(\frac{v}{v+\varepsilon}\Big)^{p-1} v\,|\boldsymbol{\eta}|^p \,\mathrm{d}\gamma \le \int_\Omega |\boldsymbol{\eta}|^p \,\mathrm{d}\mu = \mathscr{I}_p(\mu|\,\gamma).$$

As $\varepsilon \downarrow 0$ we have $s_\varepsilon(x) \uparrow s(x) := \big(v(x)\big)^{1/p}$ and its gradient is uniformly bounded in $L^p_\gamma(\Omega;\mathbb{R}^d)$; hence it follows that

$$s \in W^{1,p}_\gamma(\Omega), \quad \mathrm{D}s = \frac{\mathrm{D}v}{p\, v^{1-1/p}} = \frac{1}{p} s\boldsymbol{\eta} \in L^p_\gamma(\Omega;\mathbb{R}^d). \tag{2.47}$$

The converse implication of (2.40) is an easy consequence of a truncation argument and the Chain rule for Sobolev functions. The improved summability properties for v follow from the application of Sobolev Imbedding Theorem to $s = v^{1/p}$. Combining Rellich and Prokhorov Theorems we get (2.42) and (2.43). (2.44) is a consequence of Proposition 2.2.

The first weak convergence of (2.45) is a simple consequence of the uniform bound of $\mathrm{D}s_n$ in $L^p_\gamma(\Omega;\mathbb{R}^d)$. The second one follows from the fact that $\mathrm{D}v_n$ is the product of a sequence strongly converging in $L^{p'}_\gamma(\Omega)$ (i.e. s_n^{p-1}) and $p\mathrm{D}s_n$ which is weakly converging in $L^p_\gamma(\Omega;\mathbb{R}^d)$.

Finally, the "lim sup" assumption of (2.46) yields the strong convergence of $\mathrm{D}s_n$ in $L^p_\gamma(\Omega;\mathbb{R}^d)$, and therefore the strong convergence of $\mathrm{D}v_n$ in $L^1_\gamma(\Omega;\mathbb{R}^d)$. □

Remark 2.5 (Relative Fisher information in terms of the Lebesgue Measure). It will be useful to write the Relative Fisher information functional in terms of the Lebesgue measure: suppose that as in (1.5d,e)

$$\sup_{x\in\Omega} \big|\mathrm{D}^2 V(x)\big| < +\infty, \quad \mathrm{D}V \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega, \tag{2.48}$$

so that, for a suitable constant A,

$$f = \frac{1}{2}|\mathrm{D}V|^2 - \Delta V \quad \text{satisfies} \quad |f(x)| \le A(1+|x|^2) \quad \forall\, x \in \Omega. \tag{2.49}$$

Then, as we already observed in (1.19),

$$\frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) = \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \int_\Omega f(x)\,\mathrm{d}\mu(x) \quad \forall\, \mu \in \mathcal{P}_2(\Omega); \tag{2.50}$$

in particular when γ is the centered Gaussian measure of variance λ^{-1} in $\Omega = \mathbb{R}^d$ as in (1.8) we have

$$\frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) = \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \frac{\lambda^2}{2}\mathsf{m}_2^2(\mu) - \lambda d \quad \forall\, \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{2.51}$$

For, setting as usual, $\mu = s^2\gamma = r^2\mathcal{L}^d$, we have

$$s = r\, e^{\frac{1}{2}V}, \quad \mathrm{D}s = \big(\mathrm{D}r + \frac{1}{2} r\, \mathrm{D}V\big) e^{\frac{1}{2}V}, \tag{2.52}$$

so that

$$\begin{aligned}
\frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) &= 2\int_\Omega \big|\mathrm{D}s\big|^2\,\mathrm{d}\gamma = 2\int_\Omega \big|\mathrm{D}r + \frac{1}{2} r\mathrm{D}V\big|^2\,\mathrm{d}x\\
&= 2\int_\Omega \big|\mathrm{D}r\big|^2\,\mathrm{d}x + \frac{1}{2}\int_\Omega r^2\big|\mathrm{D}V\big|^2\,\mathrm{d}x + 2\int_\Omega r\mathrm{D}r\cdot\mathrm{D}V\,\mathrm{d}x\\
&\overset{(2.48)}{=} \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \int_\Omega \big(\frac{1}{2}|\mathrm{D}V|^2 - \Delta V\big) r^2\,\mathrm{d}x.
\end{aligned}$$

The last integration by parts can be justified by a standard localization argument recalling that the quadratic moment of μ is finite and the differential of V has a linear growth by (2.48). For, introducing a smooth function $\theta : \mathbb{R}^d \to [0,+\infty)$ satisfying $\theta(x) \equiv 1$ if $|x| \le 1$ and $\theta(x) \equiv 0$ if $|x| \ge 2$, and setting $\theta_k(x) = \theta(x/k)$, we get

$$\begin{aligned}
2\int_\Omega r\mathrm{D}r\cdot\mathrm{D}V\,\mathrm{d}x &= \lim_{k\uparrow\infty}\int_\Omega \mathrm{D}u\cdot\mathrm{D}V\,\theta_k(x)\,\mathrm{d}x = -\lim_{k\uparrow\infty}\int_\Omega \Big(u\,\Delta V\,\theta_k + u\mathrm{D}V\cdot\mathrm{D}\theta_k\Big)\,\mathrm{d}x\\
&= -\lim_{k\uparrow\infty}\int_\Omega u\,\Delta V\,\theta_k\,\mathrm{d}x = -\int_\Omega u\,\Delta V\,\mathrm{d}x
\end{aligned}$$

since

$$\int_\Omega \big|u\mathrm{D}V\cdot\mathrm{D}\theta_k\big|\,\mathrm{d}x \le \frac{1}{k}\sup_{\mathbb{R}^d}|\mathrm{D}\theta|\int_\Omega |\mathrm{D}V|\,\mathrm{d}\mu \to 0 \quad \text{as } k\uparrow\infty.$$

Weighted Fisher information. Let $\mu = v\cdot\gamma = s^2\cdot\gamma \in \mathcal{P}_2^r(\Omega)$ with $\mathscr{I}_2(\mu|\,\gamma) < +\infty$. If $\boldsymbol{\eta}$ is the logarithmic gradient defined as in (2.35a,b), we set

$$\mathscr{I}_2^V(\mu|\,\gamma) := \int_\Omega \mathrm{D}^2V(x)\boldsymbol{\eta}(x)\cdot\boldsymbol{\eta}(x)\,\mathrm{d}\mu(x) = 4\int_\Omega \mathrm{D}^2V(x)\mathrm{D}s(x)\cdot\mathrm{D}s(x)\,\mathrm{d}\gamma(x). \tag{2.53}$$

It is easy to check that for $\mu_n, \mu \in \mathcal{P}_2^r(\Omega)$ with $\mu_n \rightharpoonup \mu$ in $\mathcal{P}(\Omega)$, $\quad \sup_{n\in\mathbb{N}}\mathscr{I}_2(\mu_n|\,\gamma) < +\infty$, and $\lambda' \ge \lambda^-$,

$$\liminf_{n\to\infty}\Big(\mathscr{I}_2^V(\mu_n|\,\gamma) + \lambda'\mathscr{I}_2(\mu_n|\,\gamma)\Big) \ge \mathscr{I}_2^V(\mu|\,\gamma) + \lambda'\mathscr{I}_2(\mu|\,\gamma). \tag{2.54}$$

2.4. **Kantorovich-Rubinstein-Wasserstein distance.** Let Ω be an open convex subset of $\mathbb{R}^d$, $\mu_1, \mu_2 \in \mathcal{P}(\Omega)$ be two Borel probability measures; $\boldsymbol{\mu} \in \mathcal{P}(\Omega\times\Omega)$ has marginals μ_1, μ_2, and we say that $\boldsymbol{\mu}$ is a *transference plan* in $\Gamma(\mu_1,\mu_2)$, if

$$\int_{\Omega\times\Omega}\zeta(x_i)\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2) = \int_\Omega \zeta(x_i)\,\mathrm{d}\mu_i(x_i) \qquad \forall\,\zeta\in C_b^0(\Omega), \quad i=1,2.$$

$\Gamma(\mu_1,\mu_2)$ is not empty, since it contains the product measure $\mu_1\otimes\mu_2$; the admissible transference plans in $\Gamma(\mu_1,\mu_2)$ which are concentrated on graphs of Borel maps $\boldsymbol{r} : \Omega\to\Omega$ play a crucial role: in that case it is not difficult to check that integration w.r.t. $\boldsymbol{\mu}$ can be reduced to integration w.r.t. μ_1 by

$$\int_{\Omega\times\Omega}\zeta(x_1,x_2)\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2) = \int_\Omega \zeta(x_1,\boldsymbol{r}(x_1))\,\mathrm{d}\mu_1(x_1),$$

so that the second marginal condition becomes

$$\int_\Omega \zeta(x_2)\,\mathrm{d}\mu_2(x_2) = \int_\Omega \zeta(\boldsymbol{r}(x_1))\,\mathrm{d}\mu_1(x_1) \quad \forall\,\zeta\in C_b^0(\Omega). \tag{2.55}$$

If (2.55) holds, we will say that $\mu_2 = \boldsymbol{r}_\#\mu_1$ and $\boldsymbol{r}$ is a *transport map* between μ_1 and μ_2; admissible transport maps are denoted by $T(\mu_1,\mu_2)$ and the induced transference plans can be represented by $\boldsymbol{\mu} = (\boldsymbol{i}\times\boldsymbol{r})_\#\mu_1$.

The L^2-*Kantorovich-Rubinstein-Wasserstein distance* (in short, *Wasserstein distance*) between two probability measures $\mu_1, \mu_2 \in \mathcal{P}_2(\Omega)$ is defined by

$$W^2(\mu_1,\mu_2) := \inf\Big\{ \int_{\Omega\times\Omega} |x_2 - x_1|^2 \,\mathrm{d}\boldsymbol{\mu}(x_1,x_2) : \boldsymbol{\mu} \in \Gamma(\mu_1,\mu_2)\Big\}. \tag{2.56}$$

It is not difficult to show that W is a true distance and that the inf in (2.56) is attained; in fact, a much stronger and deeper result holds true: we collect in the next results various contributions by Brenier [Bre91], Knott & Smith [SK87], Gangbo & McCann [GM96]; see also [RR98, Vil03, AGS05].

Theorem 2.6. *If $\mu_i := v_i \cdot \gamma \in \mathcal{P}_2^r(\Omega)$, $i = 1,2$, then there exists an optimal transport map $\boldsymbol{r} := T_o(\mu_1,\mu_2) \in T(\mu_1,\mu_2)$ (uniquely determined μ_1-almost everywhere) which satisfies*

$$\boldsymbol{r}_\#\mu_1 = \mu_2, \qquad \int_\Omega |\boldsymbol{r}(x_1) - x_1|^2 \,\mathrm{d}\mu(x_1) = W^2(\mu_1,\mu_2). \tag{2.57}$$

Moreover the graph of $\boldsymbol{r}$ is contained in a cyclically monotone set, and there exists a Borel set $\Sigma \subset \Omega$ with $\mu_1(\Omega\setminus\Sigma) = 0$ such that

i. $\boldsymbol{r}$ is strictly monotone (in particular injective) in Σ i.e.

$$\langle \boldsymbol{r}(x) - \boldsymbol{r}(y), x - y\rangle > 0 \quad \text{for every } x, y \in \Sigma; \tag{2.58}$$

ii. $\boldsymbol{r}$ is differentiable in Σ, and

$$\mathrm{D}\boldsymbol{r}(x) \quad \text{is symmetric and strictly positive definite for every } x \in \Sigma, \tag{2.59}$$

iii. the change of variable formula holds

$$v_2(y) = 0 \quad \text{in } \Omega \setminus \boldsymbol{r}(\Sigma), \quad v_2(\boldsymbol{r}(x))g(\boldsymbol{r}(x)) = \frac{v_1(x)\,g(x)}{\det \mathrm{D}\boldsymbol{r}(x)} \quad \forall\, x \in \Sigma. \tag{2.60}$$

For $\mu_1, \mu_2 \in \mathcal{P}_2^r(\Omega)$ we denote by $T_o(\mu_1,\mu_2)$ the (unique) *optimal transport* $\boldsymbol{r}$ satisfying (2.57). Observe that

$$\mathsf{m}_2(\mu_2) \le W(\mu_1,\mu_2) + \mathsf{m}_2(\mu_1). \tag{2.61}$$

The proof of the next Lemma can be found in [Vil03].

Lemma 2.7 (**Strict convexity**)**.** *Let us fix $\nu \in \mathcal{P}_2^r(\Omega)$; the map $\mu \in \mathcal{P}_2(\Omega) \mapsto W^2(\mu,\nu)$ is strictly convex, i.e. for each $\mu_0, \mu_1 \in \mathcal{P}_2(\Omega)$ with $\mu_0 \neq \mu_1$ we have*

$$W^2((1-t)\mu_0 + t\mu_1, \nu) < (1-t)W^2(\mu_0,\nu) + tW^2(\mu_1,\nu) \quad \forall\, t \in (0,1). \tag{2.62}$$

Lemma 2.8 (**Lower semicontinuity and convergence in $\mathcal{P}_2(\Omega)$**)**.** *Let $\mu_n, \mu, \nu \in \mathcal{P}_2(\Omega)$; if $\mu_n \rightharpoonup \mu$ narrowly in $\mathcal{P}(\Omega)$ then*

$$\liminf_{n\uparrow+\infty} W(\mu_n,\nu) \ge W(\mu,\nu). \tag{2.63}$$

Moreover,

$$\lim_{n\uparrow+\infty} W(\mu_n,\mu) = 0 \qquad \Longleftrightarrow \qquad \mu_n \rightharpoonup \mu \quad \text{narrowly in } \mathcal{P}(\Omega), \quad \mathsf{m}_2(\mu_n) \to \mathsf{m}_2(\mu). \tag{2.64}$$

In this case

$$\lim_{n\to\infty} \int_\Omega \zeta(x)\,\mathrm{d}\mu_n(x) = \int_\Omega \zeta(x)\,\mathrm{d}\mu(x), \tag{2.65}$$

for every continuous function $\zeta : \Omega \to \mathbb{R}$ with quadratic growth.

For the *proof*, see [Vil03, chapter 7] or [AGS05, Chap. 7].

Remark 2.9. When Ω is convex, the canonical identification of $\mathcal{P}_2(\Omega)$ with the subset $\big\{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \mu(\mathbb{R}^d \setminus \Omega) = 0\big\}$ already discussed in Remark 2.1 is in fact an isometry with respect to the Wasserstein Distance, since the Wasserstein distance in $\mathcal{P}_2(\Omega)$ coincides with the Wasserstein distance inherited from $\mathcal{P}_2(\mathbb{R}^d)$ through this inclusion.

2.5. **The discrete scheme and its abstract convergence properties.** We recall here a general result of convergence for the variational approximation scheme we presented in § 1.7 of the Introduction. We are supposing that the functional

$$\begin{aligned}\phi : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty,+\infty] \quad &\text{is l.s.c. w.r.t. } \textit{narrow convergence} \text{ on}\\ \text{bounded subsets of } \mathcal{P}_2(\mathbb{R}^d), \text{ with} \quad &\emptyset \neq D(\phi) := \big\{\mu \in \mathcal{P}_2(\mathbb{R}^d) : \phi(\mu) < +\infty\big\} \subset \mathcal{P}_2^r(\Omega),\end{aligned} \tag{2.66a}$$

and its satisfies the lower bound

$$\exists\, \tau_o > 0 : \qquad \inf_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \phi(\mu) + \frac{1}{2\tau_o}\mathsf{m}_2^2(\mu) > -\infty. \tag{2.66b}$$

Taking into account Remark 2.1, the compactness property (2.4), and (2.66b), (2.66a) is equivalent to

$$\begin{aligned}&\text{for every sequence } (\mu_n) \subset D(\phi) \text{ with} \qquad \sup_n \Big(\phi(\mu_n) + \frac{1}{\tau_o}\mathsf{m}_2(\mu_n)\Big) < +\infty,\\ &\exists\,(\mu_{n_k}),\ \exists\,\mu \in D(\phi) \subset \mathcal{P}_2^r(\Omega) : \quad \mu_{n_k} \rightharpoonup \mu \quad \text{narrowly in } \mathcal{P}(\Omega), \quad \liminf_{n\to\infty} \phi(\mu_n) \geq \phi(\mu).\end{aligned} \tag{2.67}$$

Before stating the general approximation result, we need to introduce two relevant definitions [AGS05, 10.1.1, 11.1.5] in order to interpret the limit equation satisfied by any limit point $\mu \in GMM(\mu_0;\phi)$.

Definition 2.10 (**Strong and regular limiting subdifferentials; regular functionals**)**.** Let $\phi : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty,+\infty]$ be a given functional and let $\mu \in D(\phi)$; we say that $\boldsymbol{\xi} \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$ belongs to the *strong subdifferential* $\partial_s\phi(\mu)$ of ϕ at μ if

$$\phi(\boldsymbol{t}_{\#}\mu) - \phi(\mu) \geq \int_\Omega \boldsymbol{\xi}(x)\cdot(\boldsymbol{t}(x)-x)\,\mathrm{d}\mu(x) + o\big(\|\boldsymbol{t}-\boldsymbol{i}\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\big). \tag{2.68}$$

Let us further assume that ϕ satisfies (2.66a,b); ϕ ***has a regular limiting subdifferential*** if for every couple of sequences $\mu_k \in \mathcal{P}_2^r(\Omega)$, $\boldsymbol{\xi}_k \in L^2_{\mu_k}(\Omega;\mathbb{R}^d)$ such that

$$\begin{cases} \boldsymbol{\xi}_k \in \partial_s\phi(\mu_k), \quad \mu_k \rightharpoonup \mu \quad \text{narrowly in } \mathcal{P}(\Omega),\\ \sup_k \Big(\phi(\mu_k) + \dfrac{1}{\tau_o}\mathsf{m}_2(\mu_k) + \displaystyle\int_\Omega |\boldsymbol{\xi}_k(x)|^2\,\mathrm{d}\mu_k(x)\Big) < +\infty,\end{cases} \tag{2.69}$$

there exists a unique $\boldsymbol{\xi} \in L^2_\mu(\Omega;\mathbb{R}^d)$ such that

$$\boldsymbol{\nu}_k = \boldsymbol{\xi}_k\mu_k \rightharpoonup \boldsymbol{\nu} = \boldsymbol{\xi}\mu \quad \text{in } \big[\mathcal{M}_{\mathrm{loc}}(\Omega)\big]^d. \tag{2.70}$$

We call $\boldsymbol{\xi}$ *the limiting subdifferential* $\partial_\ell\phi(\mu)$ of ϕ; $\mu \mapsto \partial_\ell\phi(\mu)$ is thus single valued in its proper domain $D(\partial_\ell\phi)$. If moreover every sequence (μ_k) as in (2.69) also satisfies

$$\lim_{k\to\infty} \phi(\mu_k) = \phi(\mu), \tag{2.71}$$

then we say that ϕ ***is regular.***

The role of the strong subdifferential is provided by the following simple properties:

S1: Derivative of ϕ along a smooth vector field: Suppose that for $t \in (-\varepsilon,\varepsilon)$ $\boldsymbol{X}_t : \Omega \to \Omega$ is smooth family of maps such that for a suitable constant $C > 0$

$$|\boldsymbol{X}_t(x) - x| \leq Ct(1+|x|), \qquad \frac{d}{dt}\boldsymbol{X}_t(x)_{|t=0} = \boldsymbol{\zeta}(x) \qquad \forall\, x \in \Omega,\ t \in (-\varepsilon,\varepsilon). \tag{2.72}$$

If $\boldsymbol{\xi} \in \partial_s\phi(\mu)$ and the composed function $t \mapsto \phi\big((\boldsymbol{X}_t)_{\#}\mu\big)$ is left and right differentiable at $t = 0$, then

$$\frac{d}{dt}\phi\big((\boldsymbol{X}_t)_{\#}\mu\big)_{|t=0-} \leq \int_\Omega \langle\boldsymbol{\xi}(x),\boldsymbol{\zeta}(x)\rangle\,\mathrm{d}\mu(x) \leq \frac{d}{dt}\phi\big((\boldsymbol{X}_t)_{\#}\mu\big)_{|t=0+}. \tag{2.73}$$

In particular, if $t \mapsto \phi\big((\boldsymbol{X}_t)_{\#}\mu\big)$ is *differentiable*, then

$$\frac{d}{dt}\phi\big((\boldsymbol{X}_t)_{\#}\mu\big)_{|t=0} = \int_\Omega \langle\boldsymbol{\xi}(x),\boldsymbol{\zeta}(x)\rangle\,\mathrm{d}\mu(x). \tag{2.74}$$

S2: Euler equation for the variational scheme: Let $\{M_\tau^n\}_{n=1}^\infty$ be a solution of the variational scheme (1.44), starting from $\mu_0 \in \mathcal{P}_2(\Omega)$. Since $M_\tau^n \in D(\phi) \subset \mathcal{P}_2^r(\Omega)$, by Theorem 2.6 we find a unique optimal transport map $\boldsymbol{r}_\tau^n = T_o(M_\tau^n, M_\tau^{n-1})$, $n \in \mathbb{N}$, thus satisfying

$$(\boldsymbol{r}_\tau^n)_\# M_\tau^n = M_\tau^{n-1}, \quad W^2(M_\tau^n, M_\tau^{n-1}) = \int_\Omega |\boldsymbol{r}_\tau^n(x) - x|^2 \,\mathrm{d}M_\tau^n(x). \tag{2.75}$$

Then it is not difficult to show [AGS05, Lemma 10.1.2] that

$$-\boldsymbol{v}_\tau^n := \frac{\boldsymbol{r}_\tau^n - \boldsymbol{i}}{\tau} \in \partial_s \phi(M_\tau^n) \subset L^2_{M_\tau^n}(\Omega; \mathbb{R}^d). \tag{2.76}$$

S3: Smooth perturbation: let us suppose that f satisfies (1.4) and let us consider the functional

$$\mu \mapsto \langle f, \mu\rangle := \int_\Omega f(x) \,\mathrm{d}\mu(x) \quad \forall\, \mu \in \mathcal{P}_2(\Omega), \tag{2.77}$$

and the induced perturbed functional

$$\mu \mapsto \phi^f(\mu) := \phi(\mu) + \langle f, \mu\rangle \quad \forall\, \mu \in \mathcal{P}_2(\Omega), \tag{2.78}$$

having the same domain of ϕ. Thanks to (1.4), if ϕ satisfies (2.66a,b), then also $\phi^f := \phi + \langle f, \cdot\rangle$ satisfies (2.66a,b). The next perturbation result reduces the computation of the subdifferential of ϕ^f to that of ϕ:

Lemma 2.11 (**Linear perturbation of subdifferentials**)**.** *If f satisfies* (1.4) *then $\langle f, \cdot\rangle$ is Wasserstein differentiable at each measure $\mu \in \mathcal{P}_2(\Omega)$ with*

$$\boldsymbol{\xi} \in \partial_s \langle f, \mu\rangle \qquad \Longleftrightarrow \qquad \boldsymbol{\xi} = \mathrm{D}f \quad \mu\text{-a.e. in } \Omega, \tag{2.79}$$

which in particular implies

$$\langle f, \boldsymbol{t}_\#\mu\rangle - \langle f, \mu\rangle - \int_\Omega \mathrm{D}f \cdot (\boldsymbol{t} - \boldsymbol{i}) \,\mathrm{d}\mu(x) = o\big(\|\boldsymbol{t} - \boldsymbol{i}\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\big), \quad \boldsymbol{t}(\Omega) \subset \Omega. \tag{2.80}$$

If $\mu_k \rightharpoonup \mu$ narrowly in $\mathcal{P}(\Omega)$ with $\sup_k \mathsf{m}_2(\mu_k) < +\infty$, then $(\mathrm{D}f)\mu_k \rightharpoonup (\mathrm{D}f)\mu$ weakly in $\big[\mathcal{M}_{\mathrm{loc}}(\Omega)\big]^d$. We thus have the perturbation rule

$$\boldsymbol{\xi} \in \partial_s \phi^f(\mu) \quad \Longleftrightarrow \quad \boldsymbol{\xi} - \mathrm{D}f \in \partial_s \phi(\mu); \tag{2.81a}$$

finally, if ϕ has a regular limiting subdifferential, then ϕ^f satisfies the same property, with

$$\partial_\ell \phi^f(\mu) = \partial_\ell \phi(\mu) + \mathrm{D}f. \tag{2.81b}$$

Proof. Let us introduce the convex set

$$L^2_\mu(\Omega; \Omega) := \big\{\boldsymbol{t} \in L^2_\mu(\Omega; \mathbb{R}^d) : \boldsymbol{t}(x) \in \Omega \text{ for } \mu\text{-a.e. } x \in \Omega\big\}; \tag{2.82}$$

since

$$\langle f, \boldsymbol{t}_\#\mu\rangle = \int_\Omega f(\boldsymbol{t}(x)) \,\mathrm{d}\mu(x),$$

the mean value Theorem yields

$$\begin{aligned}\langle f, \boldsymbol{t}_\#\mu\rangle - \langle f, \mu\rangle - \int_\Omega \mathrm{D}f \cdot (\boldsymbol{t} - \boldsymbol{i}) \,\mathrm{d}\mu &= \int_0^1 \int_\Omega \Big(\mathrm{D}f\big((1-\theta)\boldsymbol{i} + \theta\boldsymbol{t}\big) - \mathrm{D}f\Big) \cdot (\boldsymbol{t} - \boldsymbol{i}) \,\mathrm{d}\mu \,\mathrm{d}\theta \\ &\le \big\|\boldsymbol{t} - \boldsymbol{i}\big\|_{L^2_\mu(\Omega;\mathbb{R}^d)} \int_0^1 \big\|\mathcal{R}_\theta[\boldsymbol{t}]\big\|_{L^2_\mu(\Omega;\mathbb{R}^d)} \,\mathrm{d}\theta, \qquad \forall\, \boldsymbol{t} \in L^2_\mu(\Omega; \Omega),\end{aligned} \tag{2.83}$$

where $\mathcal{R}_\theta[\cdot]$ is the transformation

$$\boldsymbol{t} \in L^2_\mu(\Omega; \Omega) \mapsto \mathcal{R}_\theta[\boldsymbol{t}](x) := R_\theta(x, \boldsymbol{t}(x)) \quad \forall\, x \in \Omega, \tag{2.84}$$

induced by the continuous maps

$$R_\theta(x, y) := \mathrm{D}f\big((1-\theta)x + \theta y\big) - \mathrm{D}f(x), \quad \theta \in [0, 1],\ x, y \in \Omega. \tag{2.85}$$

Since R_θ satisfies

$$|R_\theta(x, y)| \le C(1 + |x| + |y|), \quad R_\theta(x, x) = 0 \quad \forall\, \theta \in [0, 1],\ x, y \in \Omega, \tag{2.86}$$

$\mathcal{R}_\theta$ is a continuous (superposition) operator in $L^2_\mu(\Omega;\Omega)$ (see e.g. [AP93]), with

$$\big\|\mathcal{R}_\theta[\boldsymbol{t}]\big\|_{L^2_\mu(\Omega;\mathbb{R}^d)} \le C\big(1+\|\boldsymbol{t}\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\big) \qquad \forall\, \boldsymbol{t}\in L^2_\mu(\Omega;\Omega), \quad \mathcal{R}_\theta[\boldsymbol{i}]=0. \tag{2.87}$$

Therefore as $\boldsymbol{t}\to\boldsymbol{i}$ in $L^2_\mu(\Omega;\Omega)$ we have

$$\mathcal{R}_\theta[\boldsymbol{t}]\to 0 \quad \text{in } L^2_\mu(\Omega;\mathbb{R}^d), \qquad \int_0^1 \big\|\mathcal{R}_\theta[\boldsymbol{t}]\big\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\,\mathrm{d}\theta\to 0, \tag{2.88}$$

thanks to Lebesgue Dominated Convergence Theorem; (2.83) and (2.88) yield (2.80). The converse implication in (2.79) is immediate: if $\boldsymbol{\xi}\in\partial_s\langle f,\mu\rangle$ then choosing $\boldsymbol{t}=\boldsymbol{i}+\varepsilon\boldsymbol{\eta}$ with $\boldsymbol{\eta}\in C^\infty_{\rm c}(\Omega;\mathbb{R}^d)$ in (2.68) and (2.80) and letting $\varepsilon\downarrow 0$, we obtain

$$\int_\Omega \langle \boldsymbol{\xi}-\mathrm{D}f,\boldsymbol{\eta}\rangle\,\mathrm{d}\mu\le 0 \quad \forall\,\boldsymbol{\eta}\in C^\infty_{\rm c}(\Omega;\mathbb{R}^d), \tag{2.89}$$

which yields $\boldsymbol{\xi}(x)=\mathrm{D}f(x)$ for μ-a.e. $x\in\Omega$.

Since $\mathrm{D}f$ has a linear growth and it is continuous, the second part of the Lemma is a standard application of narrow convergence (see e.g. [AGS05, Chap. V]). □

Keeping the same notation of § 1.7, (2.75), and (2.76), we also set $\boldsymbol{\nu}^n_\tau := \boldsymbol{v}^n_\tau M^n_\tau \in \big[\mathcal{M}(\Omega)\big]^d$ and, as usual, we denote by $\overline{M}_\tau,\overline{\boldsymbol{\nu}}_\tau,\overline{\boldsymbol{v}}_\tau$ the corresponding piecewise constant functions defined on Ω_∞:

$$\overline{M}_{\tau,t}:=M^n_\tau,\quad \overline{\boldsymbol{\nu}}_\tau:=\boldsymbol{\nu}^n_\tau,\quad \overline{\boldsymbol{v}}_\tau:=\boldsymbol{v}^n_\tau,\quad \bar t_\tau:=n\tau \qquad \text{if}\quad (n-1)\tau<t\le n\tau. \tag{2.90}$$

We will also use a different interpolation family, firstly introduced by E. De Giorgi.

Definition 2.12 (**De Giorgi's variational interpolants).** For every $t=(n-1)\tau+\sigma\in\big((n-1)\tau,n\tau\big]$, $0<\sigma\le\tau$, we denote by

$$\tilde M_{\tau,t} \text{ a minimizer of} \qquad \mu\mapsto\Phi(\sigma,M^{n-1}_\tau;\mu)=\frac{1}{2\sigma}W^2(M^{n-1}_\tau,\mu)+\phi(\mu); \tag{2.91}$$

observe that for $\sigma=\tau,\ t=t_n=n\tau$, we can always choose $\tilde M_\tau(t_n)=\overline{M}_\tau(t_n)=M^n_\tau$. We also set

$$\tilde{\boldsymbol{r}}_{\tau,t}=T_o(\tilde M_{\tau,t},M^{n-1}_\tau),\quad \tilde{\boldsymbol{v}}_{\tau,t}:=\frac{\boldsymbol{i}-\tilde{\boldsymbol{r}}_{\tau,t}}{\sigma}\in L^2_{\tilde M_{\tau,t}}(\Omega;\mathbb{R}^d),\quad \tilde{\boldsymbol{\nu}}_\tau:=\tilde{\boldsymbol{v}}_\tau\,\tilde M_\tau, \tag{2.92}$$

observing that

$$-\tilde{\boldsymbol{v}}_\tau\in\partial_s\phi(\tilde M_{\tau,t})\quad\forall\,t>0. \tag{2.93}$$

The following result provides the starting point for all the further developments [AGS05, Prop. 2.2.3, Prop. 3.2.2, Cor. 3.3.4, §3,4, Theorem 11.1.6, Cor. 11.1.8].

Theorem 2.13 (**G.M.M., gradient flows, and "Wasserstein" subdifferential**). *Let ϕ be a functional satisfying* (2.66a,b).

a) ***Existence of approximate solutions and energy estimate.*** *For each $M\in\mathcal{P}_2(\Omega)$, $\tau\in(0,\tau_o)$, the functional $\mu\mapsto\Phi(\tau,M;\mu)$ admits a minimum point, which is unique if ϕ is also* convex. *Therefore for every initial choice of $M^0_\tau\in\mathcal{P}_2(\Omega)$ the variational scheme* (1.44) *admits a (unique, if ϕ is convex) solution $\{M^n_\tau\}_{n=1}^\infty$, whose piecewise constant interpolation $\overline{M}_\tau$ and variational interpolation $\tilde M_\tau$ satisfy*

$$-\overline{\boldsymbol{v}}_{\tau,t}\in\partial_s\phi(\overline{M}_{\tau,t}),\quad -\tilde{\boldsymbol{v}}_{\tau,t}\in\partial_s\phi(\tilde M_{\tau,t})\quad\forall\,t>0, \tag{2.94}$$

and, for every $0\le s<t<+\infty$

$$\phi(\overline{M}_{\tau,t})+\frac12\int_{\bar s_\tau}^{\bar t_\tau}\int_\Omega|\overline{\boldsymbol{v}}_{\tau,s}|^2\,\mathrm{d}\overline{M}_{\tau,s}\,\mathrm{d}s+\frac12\int_{\bar s_\tau}^{\bar t_\tau}\int_\Omega|\tilde{\boldsymbol{v}}_{\tau,s}|^2\,\mathrm{d}\tilde M_{\tau,s}\,\mathrm{d}s\le\phi(\overline{M}_{\tau,s}). \tag{2.95}$$

b) ***Compactness of discrete solutions*** *If*

$$M^0_\tau \rightharpoonup \mu_0, \quad \limsup_{\tau\downarrow 0} \mathsf{m}_2(M^0_\tau) =: \mathsf{m}_{2,0} < +\infty, \quad \phi(\mu_0) \le \limsup_{\tau\downarrow 0} \phi(M^0_\tau) =: \varphi_0 < +\infty, \tag{2.96}$$

then there exists a time step $\tau_* > 0$ and for every $T > 0$ there exists a constant $C > 0$ only depending on $\mathsf{m}_{2,0}, \varphi_0, T, \tau_o$ such that

$$\mathsf{m}_2(\overline{M}_{\tau,t}) \le C, \quad \phi(\overline{M}_{\tau,t}) \le C, \quad \int_0^T \int_\Omega \left|\overline{\boldsymbol{v}}_{\tau,s}\right|^2 \mathrm{d}\overline{M}_{\tau,s}\,\mathrm{d}s \le C \qquad \forall\, t\in[0,T],\ 0<\tau<\tau_*. \tag{2.97}$$

Each infinitesimal sequence τ_k of time steps admits a subsequence (still denoted by τ_k) and a non increasing map $\varphi : t \in [0,+\infty) \to \varphi_t \in [0,+\infty)$ such that

$$\overline{M}_{\tau_k,t},\ \tilde{M}_{\tau_k,t} \to \mu_t \quad \textit{narrowly in } \mathcal{P}(\Omega) \quad \forall\, t\ge 0, \tag{2.98}$$

$$\overline{\boldsymbol{\nu}}_{\tau_k} \rightharpoonup \boldsymbol{\nu} = \boldsymbol{v}\mu, \qquad \tilde{\boldsymbol{\nu}}_{\tau_k} \rightharpoonup \tilde{\boldsymbol{\nu}} = \tilde{\boldsymbol{v}}\mu \quad \textit{weakly in } \left[\mathcal{M}_{\mathrm{loc}}(\Omega_\infty)\right]^d, \tag{2.99}$$

$$\varphi_0 \ge \varphi_s \ge \lim_{k\uparrow+\infty} \phi(\tilde{M}_{\tau_k,t}) \ge \lim_{k\uparrow+\infty} \phi(\overline{M}_{\tau_k,t}) = \varphi_t \ge \phi(\mu_t) \quad \forall\, 0\le s<t, \tag{2.100}$$

$$\lim_{k\uparrow+\infty} \phi(\overline{M}_{\tau_k,t}) = \varphi_t = \lim_{k\uparrow+\infty} \phi(\tilde{M}_{\tau_k,t}) \quad \textit{if } \varphi \textit{ is left continuous at } t. \tag{2.101}$$

The map $t\mapsto\mu_t$ is continuous in $[0,+\infty)$ with values in $\mathcal{P}^r_2(\Omega)$, for every $T>0$

$$\begin{aligned} \int_0^T \int_\Omega |\boldsymbol{v}_t|^2 \ \mathrm{d}\mu_t\,\mathrm{d}t &\le \liminf_{k\uparrow+\infty} \int_0^T \int_\Omega \left|\overline{\boldsymbol{v}}_{\tau_k,t}\right|^2 \ \mathrm{d}\overline{M}_{\tau_k,t}\,\mathrm{d}t < +\infty, \\ \int_0^T \int_\Omega |\tilde{\boldsymbol{v}}_t|^2 \ \mathrm{d}\mu_t\,\mathrm{d}t &\le \liminf_{k\uparrow+\infty} \int_0^T \int_\Omega \left|\tilde{\boldsymbol{v}}_{\tau_k,t}\right|^2 \ \mathrm{d}\tilde{M}_{\tau_k,t}\,\mathrm{d}t < +\infty, \end{aligned} \tag{2.102}$$

and the trivial extension of $\mu, \boldsymbol{v}$ outside Ω satisfy the continuity equation

$$\partial_t\mu + \operatorname{div}\big(\boldsymbol{v}\,\mu\big) = 0 \quad \textit{in } \mathscr{D}'(\mathbb{R}^d\times(0,+\infty)). \tag{2.103}$$

All the limit curves μ obtained in this way are the elements of $GMM(\mu_0;\phi)$.

c) ***Gradient flow equation for regular limiting subdifferentials*** *If ϕ has a* regular limiting subdifferential *as Definition 2.10, then for $\mathscr{L}^1$-a.e. $t\in(0,T)$*

$$\boldsymbol{\nu}_t = \tilde{\boldsymbol{\nu}}_t \quad \textit{is the unique (weak) limit point of the families} \quad \overline{\boldsymbol{\nu}}_{\tau_k,t},\ \tilde{\boldsymbol{\nu}}_{\tau_k,t}, \tag{2.104}$$

$$\int_\Omega |\boldsymbol{v}_t|^2\,\mathrm{d}\mu \le \begin{cases} \liminf\limits_{k\uparrow+\infty} \displaystyle\int_\Omega |\overline{\boldsymbol{v}}_{\tau_k,t}|^2\,\mathrm{d}\overline{M}_{\tau_k,t} \\ \liminf\limits_{k\uparrow+\infty} \displaystyle\int_\Omega |\tilde{\boldsymbol{v}}_{\tau_k,t}|^2\,\mathrm{d}\tilde{M}_{\tau_k,t}, \end{cases} \tag{2.105}$$

$\mu_t, \boldsymbol{v}_t$ satisfy the "gradient flow equation"

$$-\boldsymbol{v}_t = \partial_\ell\phi(\mu_t) \quad \textit{for a.e. } t\in(0,T), \tag{2.106}$$

and the Energy-Lyapunov inequality

$$\varphi_t + \int_s^t \int_\Omega |\boldsymbol{v}_r(x)|^2\ \mu_r(x)\,\mathrm{d}r \le \varphi_s \quad 0\le s<t\le T. \tag{2.107}$$

d) ***Convergence of the energy for regular functionals*** *Finally, if ϕ is also* **regular***, then*

$$\phi(\mu_t) = \varphi_t = \lim_{k\to\infty} \phi(\overline{M}_{\tau_k,t}) = \lim_{k\to\infty} \phi(\tilde{M}_{\tau_k,t}) \quad \mathscr{L}^1\textit{-a.e. in } (0,T). \tag{2.108}$$

2.6. **Wasserstein distance, Entropy and Fisher information.** The Wasserstein distance provides an interesting link between (relative) Entropy and (relative) Fisher information. As usual, we are assuming (2.8) and (2.9).

Let us first recall the notion of *displacement interpolation* and *displacement convexity*, which have been introduced by R.J. McCann [McC97] and play a crucial role in the Wasserstein framework.

Definition 2.14 (**Displacement interpolation and displacement convexity**). For each couple $\mu_0, \mu_1 \in \mathcal{P}_2^r(\Omega)$, $\boldsymbol{r} = T_o(\mu_0, \mu_1)$ and $t \in (0,1)$, we set

$$\boldsymbol{r}_t := (1-t)\boldsymbol{i} + t\boldsymbol{r}, \qquad \mu_t := (\boldsymbol{r}_t)_\# \mu_0. \tag{2.109}$$

A functional $\phi : \mathcal{P}_2(\Omega) \to (-\infty, +\infty]$ is λ-*displacement convex* if for every $\mu_0, \mu_1 \in D(\phi) \subset \mathcal{P}_2^r(\Omega)$ the map $t \in [0,1] \mapsto \phi(\mu_t)$ is λ-convex, i.e.

$$\phi(\mu_t) \le (1-t)\phi(\mu_0) + t\phi(\mu_1) - \frac{\lambda}{2} t(1-t) W^2(\mu_0, \mu_1) \quad \forall\, t \in [0,1]. \tag{2.110}$$

Observe that

$$W(\mu_s, \mu_t) = (t-s) W(\mu_0, \mu_1) \quad \forall\, 0 \le s \le t \le 1. \tag{2.111}$$

Proposition 2.15 (**Entropy and Fisher information**). *Let $\gamma := e^{-V}\mathcal{L}^d$ for a λ-convex function V as in (2.8).*

i) ***λ-convexity of the Entropy:*** *The functional $\mu \in \mathcal{P}_2^r(\Omega) \mapsto \mathscr{H}(\mu|\,\gamma)$ is* λ-displacement convex.

ii) ***Regularity and limiting subdifferential of the Entropy:*** *the Entropy functional $\mathscr{H}(\cdot|\,\gamma)$ is regular and has a regular limiting subdifferential according to Definition 2.10. A measure $\mu = v\cdot\gamma$ belongs to $D(\partial_\ell \mathscr{H})$ iff $v \in W^{1,1}_\gamma(\Omega)$ and $\mathrm{D}v/v \in L^2_\mu(\Omega;\mathbb{R}^d)$, i.e. $\mathscr{I}_2(\mu|\,\gamma) < +\infty$; in this case*

$$\boldsymbol{\eta} = \partial_\ell \mathscr{H}(\mu|\,\gamma) \iff \boldsymbol{\eta} = \frac{\mathrm{D}v}{v} \in L^2_\mu(\Omega;\mathbb{R}^d), \qquad \text{so that} \quad \mathscr{I}_2(\mu|\,\gamma) = \int_\Omega |\boldsymbol{\eta}|^2 \,\mathrm{d}\mu. \tag{2.112}$$

iii) ***Variational inequality for the subdifferential:*** *if $\mathscr{I}_2(\mu|\,\gamma) < +\infty$, the subdifferential $\boldsymbol{\eta} = \frac{\mathrm{D}v}{v}$ satisfies the variational inequality*

$$\int_\Omega \Big((\boldsymbol{r} - x)\cdot\boldsymbol{\eta} + \frac{\lambda}{2}|\boldsymbol{r} - x|^2 \Big)\,\mathrm{d}\mu \le \mathscr{H}(\nu|\,\gamma) - \mathscr{H}(\mu|\,\gamma) \quad \forall\, \nu \in \mathcal{P}_2^r(\Omega),\ \boldsymbol{r} = T_o(\mu,\nu). \tag{2.113}$$

In particular for every sequence $\mu_n \in \mathcal{P}_2^r(\Omega)$

$$\lim_{n\to\infty} W(\mu_n, \mu) = 0, \quad \sup_n \mathscr{I}_2(\mu_n|\,\gamma) < +\infty \implies \lim_{n\to\infty} \mathscr{H}(\mu_n|\,\gamma) = \mathscr{H}(\mu|\,\gamma). \tag{2.114}$$

iv) ***Log-Sobolev and Talagrand inequality:*** *if $\lambda > 0$ and $\gamma \in \mathcal{P}(\Omega)$ then*

$$\frac{\lambda}{2} W^2(\mu,\gamma) \le \mathscr{H}(\mu|\,\gamma) \le \frac{1}{2\lambda}\mathscr{I}_2(\mu|\,\gamma) \qquad \forall\, \mu \in \mathcal{P}_2^r(\Omega). \tag{2.115}$$

More generally, if $\lambda \le 0$ and γ_ϑ is the Probability measure (2.24) we have

$$\mathscr{H}(\mu|\,\gamma) \le \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) + \frac{1-\lambda}{2} W^2(\mu,\gamma_\vartheta) + \mathscr{H}(\gamma_\vartheta|\,\gamma) \quad \forall\, \mu \in \mathcal{P}_2^r(\Omega). \tag{2.116}$$

v) ***Derivative of the Entropy along curves:*** *let $\mu : t \in [0,T] \mapsto \mu_t = v_t\cdot\gamma \in \mathcal{P}_2^r(\Omega)$ be an (absolutely) continuous curve satisfying the continuity equation*

$$\partial_t \mu + \operatorname{div}(\boldsymbol{v}\,\mu) = 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (0,T)) \tag{2.117}$$

for a Borel vector field $\boldsymbol{v}$ with

$$\int_0^T \int_\Omega |\boldsymbol{v}_t(x)|^2 \,\mathrm{d}\mu_t(x)\,\mathrm{d}t < +\infty, \qquad \int_0^T \mathscr{I}_2(\mu_t|\,\gamma)\,\mathrm{d}t < +\infty. \tag{2.118}$$

Then the map $t \mapsto \mathscr{H}(\mu_t|\,\gamma)$ is absolutely continuous in $[0,T]$ and

$$\frac{d}{dt}\mathscr{H}(\mu_t|\,\gamma) = \int_\Omega \boldsymbol{v}_t\cdot\boldsymbol{\eta}_t \,\mathrm{d}\mu_t \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in (0,T), \quad \text{where} \quad \boldsymbol{\eta}_t = \partial_\ell \mathscr{H}(\mu_t|\,\gamma) = \frac{\mathrm{D}v_t}{v_t}. \tag{2.119}$$

Proof. i) is due to R.J. McCann [McC97] (see e.g. [AGS05, Prop. 9.3.2 and Thm. 9.4.12]).
ii) and *iii)* have been proved in [AGS05, Sect. 10.4.4, 10.1.1].
(2.115) is a particular case of (2.113): the first (Talagrand [Tal96, OV00]) inequality follows by choosing $\mu = \gamma, \boldsymbol{\eta} = 0$, the second (Logarithmic-Sobolev) inequality can be obtained by choosing $\nu := \gamma$ and applying Cauchy-Schwarz inequality. (2.116) follows by the same method by choosing $\nu := \gamma_\vartheta$. The relationship between Displacement Convexity, Talagrand and Logarithmic-Sobolev inequalities has been deeply investigated by [OV00]
v) has been proved in [AGS05, Sec. 10.1.2 (E)]. $\square$

Remark 2.16. When $\gamma = \mathcal{L}^d$ is the Lebesgue measure, the convergence of the Entropy stated in (2.114) holds without assuming the convergence of the quadratic moments, i.e. if

$$\mu_n = u_n\mathcal{L}^d, \qquad \mu_n \rightharpoonup \mu \quad \text{narrowly in } \mathcal{P}(\Omega), \quad \sup_n \Big(\mathsf{m}_2(\mu_n) + \mathscr{I}_2(\mu_n|\,\mathcal{L}^d)\Big) < +\infty \tag{2.120}$$

then

$$\lim_{n\to\infty} \mathscr{H}(\mu_n|\,\mathcal{L}^d) = \mathscr{H}(\mu|\,\mathcal{L}^d), \quad u_n \to u \quad \text{strongly in } L^p(\Omega) \quad \text{for } 1 \le p < \tfrac{1}{2}\,2^\star, \tag{2.121}$$

where $2^\star$ is the Sobolev exponent recalled in (1.74). For, the boundedness of the quadratic moments of μ_n and Rellich-Sobolev embedding Theorem for $r_n = \sqrt{u_n}$ yield strong convergence of r_n in $L^{2p}(\Omega)$ (and of u_n in $L^p(\Omega)$) for every $2p < 2^\star$.

The next result collects the main results for the Wasserstein approach to Fokker-Planck equations, introduced by [JKO98].

Theorem 2.17 (**Fokker-Planck equation as gradient flow of the Entropy**). *Let $\nu_0 \in \mathcal{P}_2^r(\Omega)$ with $\mathscr{I}_2(\nu|\,\gamma) < +\infty$; there exists a unique solution $\nu_t = \rho_t \cdot \gamma \in C^0([0,+\infty); \mathcal{P}_2(\Omega))$, with $\sup_t \mathscr{I}_2(\nu_t|\,\gamma) < +\infty$, of the Fokker-Planck equation with homogeneous Neumann boundary conditions*

$$\partial_t\nu_t - \operatorname{div}\Big(\nu_t\frac{\mathrm{D}\rho_t}{\rho_t}\Big) = 0 \quad \text{in } \Omega_\infty, \quad \partial_{\boldsymbol{n}}\rho_t = 0 \quad \text{on } (\partial\Omega)_\infty, \tag{2.122}$$

according to the following weak formulation

$$\int_{\Omega_T} \Big(-\rho_t\,\partial_t\zeta + \mathrm{D}\zeta\cdot\mathrm{D}\rho_t\Big)\,\mathrm{d}\gamma(x)\,\mathrm{d}t = \int_\Omega \zeta(0,x)\,\mathrm{d}\nu_0(x), \quad \forall\,\zeta\in\mathcal{D}(\mathbb{R}^d\times\mathbb{R}). \tag{2.123}$$

The map $t \mapsto e^{2\lambda t}\mathscr{I}_2(\nu_t|\,\gamma)$ is right continuous and not increasing; the optimal transport maps $\boldsymbol{r}_{t,h} = T_o(\nu_t,\nu_{t+h})$ satisfy

$$\mathscr{I}_2(\nu_t|\,\gamma) = \lim_{h\downarrow 0}\frac{W(\nu_t,\nu_{t+h})^2}{h^2}, \quad \frac{\mathrm{D}\rho_t}{\rho_t} = \lim_{h\downarrow 0}\frac{\boldsymbol{r}_{t,h}-x}{h} \quad \textit{strongly in } L^2(\nu_t;\mathbb{R}^d) \quad \forall\,t\ge 0. \tag{2.124}$$

Moreover if $\nu \in \mathcal{P}_2(\Omega)$ with $\mathscr{H}(\nu|\,\gamma) < +\infty$ and $\boldsymbol{r}_t = T_o(\nu_t,\nu)$ is the optimal transportation map between ν_t and ν, then

$$\tfrac{1}{2}\tfrac{d}{dt}W^2(\nu_t,\nu) = \int_\Omega \big(\boldsymbol{r}_t(x) - x\big)\cdot\mathrm{D}\rho\,\mathrm{d}\gamma(x) \tag{2.125}$$

$$\le \mathscr{H}(\nu|\,\gamma) - \mathscr{H}(\nu_t|\,\gamma) - \frac{\lambda}{2}W^2(\nu_t,\nu). \tag{2.126}$$

Finally, if $\rho_0 \in L^2_\gamma(\Omega)$ then $\rho \in C^0([0,+\infty); L^2_\gamma(\Omega))$ and

$$\rho\in C^\infty([\delta,+\infty); W^{1,2}_\gamma(\Omega)), \quad \partial_t\rho, \Delta_\gamma\rho \in C^\infty([\delta,+\infty); L^2_\gamma(\Omega)) \quad \forall\,\delta>0. \tag{2.127}$$

ρ is (unique) variational solution of

$$\partial_t\rho_t - \Delta_\gamma\rho_t = 0 \quad \text{in } \Omega_\infty, \qquad \partial_{\boldsymbol{n}}\rho = 0 \quad \text{on } (\partial\Omega)_\infty, \tag{2.128}$$

which is associate to the closed and symmetric Dirichlet form in $L^2_\gamma(\Omega)$ with domain $W^{1,2}_\gamma(\Omega)$

$$a(\rho,\eta) := \int_\Omega \mathrm{D}\rho\cdot\mathrm{D}\eta\,\mathrm{d}\gamma; \tag{2.129}$$

in particular, it satisfies

$$\text{ess-inf}\,\rho_0 \le \text{ess-inf}\,\rho_t, \quad \text{ess-sup}\,\rho_t \le \text{ess-sup}\,\rho_0 \qquad \forall\, t \ge 0. \tag{2.130}$$

Proof. The lemma is a particular case (for the Relative Entropy functional) of the general result [AGS05, Theorems 11.2.1]; the link with the standard theory of variational parabolic problems in Hilbert spaces, analytic semigroups (yielding (2.127), see e.g. [Bre83]), and Dirichlet forms has also been discussed in [AS06]. □

3. Second order functionals and Sobolev inequalities

3.1. **A Sobolev inequality for the square root.** The next theorem shows a regularity result for the square root z of a *nonnegative* function s in $W^{2,2}_{\rm loc}(\Omega)$: the first derivatives of z gain an extra (local) L^4 summability. A global result also holds, if s satisfies an homogeneous Neumann condition on $\partial\Omega$: in this case a general weight measure γ as in § 2.2 is involved and the estimates are explicit. Observe that, in order to obtain stability constants *independent of the dimension d and of γ*, we should use a second order tensor A which involves $\mathrm{D}^2 s$ and $\Delta_\gamma s$ (the Laplace operator induced by γ, see (1.10)). Recall that $|\mathsf{A}|$ and $\|\mathsf{A}\|$ denotes the Euclidean and the operator norm of A (see the notation at the beginning of § 2.

Theorem 3.1 (**A Sobolev inequality for the square root**). *Let γ be a reference Radon measure as in § 2.2 and let $s\in W^{2,2}_{\rm loc}(\Omega)$ be a* nonnegative *function; then $z:=\sqrt{s}\in W^{1,4}_{\rm loc}(\Omega)$ and*

$$4|\mathrm{D}z|^4+\mathrm{div}_\gamma\big(|\mathrm{D}z|^2\mathrm{D}s\big)=\mathsf{A}(s)\mathrm{D}z\cdot\mathrm{D}z\qquad\textit{in the sense of distributions of }\mathcal{D}'(\Omega)\tag{3.1}$$

where

$$\mathsf{A}(s):=2\mathrm{D}^2s+(\Delta_\gamma s)\,\mathsf{Id},\qquad \mathsf{A}_{ij}:=2\partial^2_{ij}s+\Big(\sum_k\partial^2_{kk}s-\partial_kV\,\partial_ks\Big)\delta_{ij},\tag{3.2}$$

in particular, for each couple of open set $\Omega'\subset\subset\Omega''\subset\subset\Omega$ there exists a constant $C=C(\Omega',\Omega'')>0$ such that

$$\int_{\Omega'}|\mathrm{D}z|^4\,\mathrm{d}x\le C\int_{\Omega''}\Big(\|\mathsf{A}(s)\|^2+|\mathrm{D}s|^2+s^2\Big)\,\mathrm{d}x.\tag{3.3}$$

Moreover, if Ω is Lipschitz, $s\in W^{2,2}_\gamma(\Omega)$, $\Delta_\gamma s\in L^2_\gamma(\Omega)$, and $\partial_{\boldsymbol n}s=0$, then $z=\sqrt{s}\in W^{1,4}_\gamma(\Omega)$ and

$$4\int_\Omega|\mathrm{D}z|^4\,\mathrm{d}\gamma=\int_\Omega\mathsf{A}(s)\mathrm{D}z\cdot\mathrm{D}z\,\mathrm{d}\gamma\le\frac14\int_\Omega\|\mathsf{A}(s)\|^2\,\mathrm{d}\gamma.\tag{3.4}$$

Proof. Let us first suppose that the nonnegative function s belongs to $C^2(\Omega)$, so that $z_\varepsilon=\sqrt{s+\varepsilon}-\varepsilon\in C^2(\Omega)$ for $\varepsilon>0$. Since

$$s=(z_\varepsilon+\varepsilon)^2-\varepsilon,\quad \partial_is=2(z_\varepsilon+\varepsilon)\partial_iz_\varepsilon,\quad \partial^2_{ij}s=2\big((z_\varepsilon+\varepsilon)\partial^2_{ij}z_\varepsilon+\partial_iz_\varepsilon\,\partial_jz_\varepsilon\big),\tag{3.5}$$

a simple calculation shows that

$$\begin{aligned}4|\mathrm{D}z_\varepsilon|^4+\mathrm{div}_\gamma\big(|\mathrm{D}z_\varepsilon|^2\mathrm{D}s\big)&=\sum_{i,j}4(\partial_iz_\varepsilon)^2(\partial_jz_\varepsilon)^2+\partial_i\big((\partial_jz_\varepsilon)^2\partial_is\big)-\partial_iF\partial_is(\partial_jz_\varepsilon)^2\\&\overset{(3.5)}{=}\sum_{i,j}4(\partial_iz_\varepsilon\,\partial_jz_\varepsilon)^2+4(z_\varepsilon+\varepsilon)\partial^2_{ij}z_\varepsilon\,\partial_iz_\varepsilon\,\partial_jz_\varepsilon+\partial^2_{ii}s(\partial_jz_\varepsilon)^2-\partial_iF\partial_is(\partial_jz_\varepsilon)^2\\&\overset{(3.5)}{=}\sum_{i,j}2\partial^2_{ij}s\partial_iz_\varepsilon\partial_jz_\varepsilon+(\partial^2_{ii}s-\partial_iF\partial_is)(\partial_jz_\varepsilon)^2=\mathsf{A}(s)\mathrm{D}z_\varepsilon\cdot\mathrm{D}z_\varepsilon.\end{aligned}$$

If $\Omega''\subset\Omega$ is a bounded and regular open subset, $s\in C^2(\overline{\Omega''})$, and $\partial_{\boldsymbol n}s=0$ on $\partial\Omega''$, we can integrate the previous identity on Ω'' obtaining

$$4\int_{\Omega''}|\mathrm{D}z_\varepsilon|^4\,\mathrm{d}x=\int_{\Omega''}\mathsf{A}(s)\mathrm{D}z_\varepsilon\cdot\mathrm{D}z_\varepsilon\,\mathrm{d}\gamma\le\Big(\int_{\Omega''}\|\mathsf{A}(s)\|^2\,\mathrm{d}\gamma\Big)^{1/2}\Big(\int_{\Omega''}|\mathrm{D}z_\varepsilon|^4\,\mathrm{d}\gamma\Big)^{1/2},$$

and therefore

$$\Big(\int_{\Omega''}|\mathrm{D}z_\varepsilon|^4\,\mathrm{d}\gamma\Big)^{1/2}\le\frac14\Big(\int_{\Omega''}\|\mathsf{A}(s)\|^2\,\mathrm{d}\gamma\Big)^{1/2},\tag{3.6}$$

which holds also for $\varepsilon=0$, simply by passing to the limit as $\varepsilon\downarrow0$. In particular, choosing open sets $\Omega'\subset\subset\Omega''\subset\subset\Omega$ and applying (3.6) to $s\zeta$ where ζ is a nonnegative smooth cutoff function such that $\zeta_{|\Omega'}\equiv1$, $\zeta_{|\Omega\setminus\Omega''}\equiv0$, we get (3.3) for an arbitrary (nonnegative) function $s\in C^2(\Omega)$. Passing to the limit as $\varepsilon\downarrow0$ in the identity

$$4|\mathrm{D}z_\varepsilon|^4+\mathrm{div}_\gamma\big(|\mathrm{D}z_\varepsilon|^2\mathrm{D}s\big)=\mathsf{A}(s)\mathrm{D}z_\varepsilon\cdot\mathrm{D}z_\varepsilon,\tag{3.7}$$

we obtain (3.1) for a C^2 function s.

A standard approximation argument by convolution yields (3.3) for $s \in W^{2,2}_{\rm loc}(\Omega)$; the next auxiliary Lemma and a localization-regularization technique yield (3.1) for $s \in W^{2,2}_{\rm loc}(\Omega)$ (being (3.1) nonlinear with respect to z, a proof by regularization should guarantee the *strong* convergence of the regularized functions).

The global identity and the inequality of (3.4) are consequences of (3.1), whenever one knows that $\mathrm{D}z \in L^4_\gamma(\Omega)$. If Ω is bounded and Lipschitz, then this global regularity can be simply obtained by extending s outside Ω preserving the summability of its derivatives.

When Ω is unbounded, we can approximate s by the sequence $s_k := s\,\vartheta_k$ where $\vartheta_k(x) = \big(\zeta_k(x)\big)^2 = \big(\zeta(|x|/k)\big)^2$ for a nonincreasing and smooth function $\zeta : [0,+\infty) \to [0,+\infty)$ satisfying $\zeta(x) \equiv 1$ if $|x| \le 1$ and $\zeta(x) \equiv 0$ if $|x| \ge 2$. Setting $z_k = \sqrt{s_k} = z\,\zeta_k$ and applying (3.1) we obtain

$$4\big|\mathrm{D}z_k\big|^4 + \mathrm{div}_\gamma\Big(\big|\mathrm{D}z_k\big|^2\vartheta_k\,\mathrm{D}s\Big) = \vartheta_k\mathsf{A}(s)\mathrm{D}z_k\cdot\mathrm{D}z_k + \mathsf{B}_k\,\mathrm{D}z_k\cdot\mathrm{D}z_k - R_k, \tag{3.8}$$

where

$$\mathsf{B}_k = s\Big(2\mathrm{D}^2\vartheta_k + \big(\Delta\vartheta_k\big)\mathsf{Id}\Big) + 2\mathrm{D}s\otimes\mathrm{D}\vartheta_k + 2\mathrm{D}\vartheta_k\otimes\mathrm{D}s + 2\big(\mathrm{D}s\cdot\mathrm{D}\vartheta_k\big)\,\mathsf{Id}, \tag{3.9}$$

and

$$R_k = \mathrm{div}\,\Big(\big|\mathrm{D}z_k\big|^2 s\,\mathrm{D}\vartheta_k\Big) = \mathrm{div}\,\Big(\big|\zeta_k\mathrm{D}z + z\mathrm{D}\zeta_k\big|^2 z^2\,\mathrm{D}\vartheta_k\Big) = \mathrm{div}\,\Big(\big|\tfrac12\zeta_k\mathrm{D}s + s\mathrm{D}\zeta_k\big|^2\,\mathrm{D}\vartheta_k\Big). \tag{3.10}$$

It is not difficult to check that there exists a constant C depending only on ζ such that

$$\big\|\mathsf{B}_k\big\|^2 \le Ck^{-2}\Big(s^2 + \big|\mathrm{D}s\big|^2\Big),\quad |R_k| \le Ck^{-1}\Big(s^2 + \big|\mathrm{D}s\big|^2 + \|\mathrm{D}^2 s\|^2\Big). \tag{3.11}$$

Integrating in Ω and arguing as before, we obtain a uniform upper bound for $\mathrm{D}z_k$ in $L^4_\gamma(\Omega)$; passing to the limit as $k \uparrow +\infty$ we get (3.4). □

Lemma 3.2. *Let $\rho_\delta := \delta^{-d}\rho(\cdot/\delta)$ be a standard family of nonnegative, C^∞ mollifiers with* $\operatorname{supp}\rho \subset B_1(0)$ *and $\int_{\mathbb{R}^d}\rho_\delta(x)\,\mathrm{d}x = 1$. If $z \in W^{1,4}(\mathbb{R}^d)$ is nonnegative and $z_\delta := \big(z^2 * \rho_\delta\big)^{1/2}$, then $z_\delta \in W^{1,4}(\mathbb{R}^d)$ converges* strongly *to z in $W^{1,4}(\mathbb{R}^d)$.*

Proof. Let us set $s := z^2$, $s_\delta := s * \rho_\delta$; since s_δ converges strongly to s in $L^2(\mathbb{R}^d)$ it is immediate to check that z_δ converges strongly to z in $L^4(\mathbb{R}^d)$. Since $\partial_i s = 2z\partial_i z$ we know that

$$I := \int_{\mathbb{R}^d}\big|\mathrm{D}z\big|^4\,\mathrm{d}x = \frac14\int_{\mathbb{R}^d}\frac{\big|\mathrm{D}s\big|^4}{s^2}\,\mathrm{d}x < +\infty. \tag{3.12}$$

The convexity of the function $Q : (u,v) \in (0,+\infty)\times\mathbb{R}^d \to G(u,v) := \frac{|v|^4}{u^2}$, whose l.s.c. envelope in $[0,+\infty)\times\mathbb{R}^d$ is

$$\bar Q(u,v) = \begin{cases} \frac{|v|^4}{u^2} & \text{if } u > 0,\\ +\infty & \text{if } u = 0,\ v \ne 0,\\ 0 & \text{if } u = 0,\ v = 0,\end{cases} \tag{3.13}$$

and Jensen inequality yield

$$I_\delta := \int_{\mathbb{R}^d}|\mathrm{D}z_\delta|^4\,\mathrm{d}x = \frac14\int_{\mathbb{R}^d}\frac{\big|\mathrm{D}s_\delta\big|^4}{s_\delta^2}\,\mathrm{d}x = \int_{\mathbb{R}^d}\bar Q(s*\rho_\delta, \mathrm{D}s*\rho_\delta)\,\mathrm{d}x \le \int_{\mathbb{R}^d}\bar Q(s,\mathrm{D}s)\,\mathrm{d}x \le I < +\infty,$$

so that $\limsup_{\delta\downarrow0} I_\delta \le I$.

We therefore deduce that $z_\delta \in W^{1,4}(\mathbb{R}^d)$ and it converges to z in $W^{1,4}(\mathbb{R}^d)$. □

3.2. **Second order functionals.** If we want to introduce a second order functional analogous to the first order Fisher information (for $p = 2$), we have at least three natural choices, which in the case $\mu = v\cdot\gamma$ with $v \in C^2(\Omega)$ and $v > 0$ read as

$$\mathscr{K}_0(\mu|\,\gamma) := 4\sum_{i,j}\int_\Omega \Big|\partial^2_{ij}\sqrt{v(x)}\Big|^2\, d\gamma(x), \tag{3.14a}$$

$$\mathscr{K}_1(\mu|\,\gamma) := \sum_{i,j}\int_\Omega \Big|\frac{\partial^2_{ij}v(x)}{v(x)}\Big|^2\,\mathrm{d}\mu(x) = \sum_{i,j}\int_\Omega \frac{|\partial^2_{ij}v(x)|^2}{v(x)}\,\mathrm{d}\gamma(x), \tag{3.14b}$$

$$\mathscr{K}_{-1}(\mu|\,\gamma) := \sum_{i,j}\int_\Omega \big|\partial^2_{ij}\log(v(x))\big|^2\,\mathrm{d}\mu(x). \tag{3.14c}$$

The main purpose of this section is to show that all these (suitably defined) functionals provide a control (at least locally) of the 4th order relative information (recall Lemma 2.4)

$$\mathscr{I}_4(\mu|\,\gamma) := 4^4\int_\Omega \big|\mathrm{D}z\big|^4\,\mathrm{d}\gamma = \int_\Omega\Big|\frac{\mathrm{D}v}{v}\Big|^4\,\mathrm{d}\mu, \qquad z := \sqrt{s} = \sqrt[4]{v}, \tag{3.15}$$

and, moreover, that they differ (locally) by a multiple of $\mathscr{I}_4(\mu|\,\gamma)$.

This kind of information is particular important when we try to extend these functional to a weaker setting (where the density v could vanish on sets of positive γ-measure) and to study their lower semicontinuity w.r.t. weak convergence of measures. In particular, the lower semicontinuity of $\mathscr{K}_{-1}$ (which is the main motivation of the present discussion) is far to be obvious starting from the definition (3.14c)

It is not difficult to extend the previous formulae for $\mathscr{K}_j$ to Sobolev functions. A first possibility, as in the definition of Fisher information (2.35a,b), is to start from (3.14b) and to ask that the (distributional) second derivatives of v admits a square integrable density with respect to the measure μ; we discuss this point of view in Corollary 3.4.

A second possibility is to exploit the distinguished role of $s = \sqrt{v}$ starting from (3.14a); here we follow this approach and we thus introduce the natural domain $D(\mathscr{K}_0)$ of $\mathscr{K}_0$, which is obviously related to the weighted Sobolev space $W^{2,2}_\gamma(\Omega)$ (2.12)

$$D(\mathscr{K}_0) := \big\{\mu = s^2\,\gamma \in \mathcal{P}_2^r(\Omega) : s \in W^{2,2}_\gamma(\Omega)\big\}. \tag{3.16}$$

$\mathscr{K}_1$ and $\mathscr{K}_{-1}$ can also be written in terms of $s = \sqrt{v}$, since, at least in the smooth and strictly positive case,

$$\partial^2_{ij}v = 2\big(s\partial^2_{ij}s + \partial_i s\,\partial_j s\big), \qquad \mathrm{D}^2 v = 2\big(s\mathrm{D}^2 s + \mathrm{D}s\otimes\mathrm{D}s\big), \tag{3.17}$$

and

$$v\partial^2_{ij}\log v = 2s^2\partial^2_{ij}\log s = 2\big(s\partial^2_{ij}s - \partial_i s\partial_j s\big), \qquad v\mathrm{D}^2\log v = 2\big(s\mathrm{D}^2 s - \mathrm{D}s\otimes\mathrm{D}s\big). \tag{3.18}$$

In order to unify our discussion, it is then natural to introduce a real parameter α and, for $s \in W^{2,2}_{\gamma,\mathrm{loc}}(\Omega) = W^{2,2}_{\mathrm{loc}}(\Omega)$, the matrices

$$\mathsf{S}^{(\alpha)} := s\mathrm{D}^2 s + \alpha\mathrm{D}s\otimes\mathrm{D}s, \quad \mathsf{S}^{(\alpha)}_{ij} = s\partial^2_{ij}s + \alpha\partial_i s\,\partial_j s \in L^1_{\mathrm{loc}}(\Omega). \tag{3.19}$$

For measures $\mu = s^2\cdot\gamma$ with $s \in W^{2,2}_{\mathrm{loc}}(\Omega)$ we can thus define

$$\mathscr{K}_\alpha(\mu|\,\gamma) := 4\int_\Omega \frac{|\mathsf{S}^{(\alpha)}|^2}{v}\,\mathrm{d}\gamma(x) = 4\mathscr{L}_2\big(\mathsf{S}^{(\alpha)}\,\gamma|\mu\big) = 4\sum_{i,j}\int_\Omega\Big|\frac{s\partial^2_{ij}s + \alpha\partial_i s\,\partial_j s}{s}\Big|^2\,\mathrm{d}\gamma, \tag{3.20}$$

whose natural domain is

$$D(\mathscr{K}_\alpha) := \Big\{\mu = s^2\,\gamma \in \mathcal{P}_2^r(\Omega) : s \in W^{2,2}_{\gamma,\mathrm{loc}}(\Omega), \quad \mathscr{K}_\alpha(\mu|\,\gamma) < +\infty\Big\}. \tag{3.21}$$

Since we are dealing with a general measure γ, it is also important to consider the functionals

$$\hat{\mathscr{K}_\alpha}(\mu|\,\gamma) := 4\int_\Omega\Big|\frac{s\Delta_\gamma s + \alpha|Ds|^2}{s}\Big|^2\,\mathrm{d}\gamma, \tag{3.22}$$

whose domains are defined as in (3.21). Observe that

$$\hat{\mathscr{K}_0}(\mu|\,\gamma) := 4\int_\Omega \Big|\Delta_\gamma \sqrt{v(x)}\Big|^2 \,\mathrm{d}\gamma(x), \tag{3.23a}$$

$$\hat{\mathscr{K}_1}(\mu|\,\gamma) := \int_\Omega \Big|\frac{\Delta_\gamma v(x)}{v(x)}\Big|^2 \,\mathrm{d}\mu(x) = \int_\Omega \frac{|\Delta_\gamma v(x)|^2}{v(x)} \,\mathrm{d}\gamma(x), \tag{3.23b}$$

$$\hat{\mathscr{K}_{-1}}(\mu|\,\gamma) := \int_\Omega |\Delta_\gamma \log(v(x))|^2 \,\mathrm{d}\mu(x), \tag{3.23c}$$

and they are "controlled" by the corresponding functionals $\mathscr{K}_\alpha$ only if V is Lipschitz (e.g. when Ω is bounded, see next Lemma 3.5).

Theorem 3.3 (**A general identity for second order functionals**). *Let Ω be a Lipschitz open set, γ a reference measure as in § 2.2, and let $s \in W^{2,2}_\gamma(\Omega)$ be a nonnegative function with $\Delta_\gamma s \in L^2_\gamma(\Omega)$ and $\partial_{\boldsymbol{n}} s = 0$ on $\partial\Omega$. Then*

$$2\mathscr{K}_\alpha(\mu|\,\gamma) + \hat{\mathscr{K}_\alpha}(\mu|\,\gamma) = 2\mathscr{K}_0(\mu|\,\gamma) + \hat{\mathscr{K}_0}(\mu|\,\gamma) + \frac{1}{4}\alpha(3\alpha+2)\mathscr{I}_4(\mu|\,\gamma) \tag{3.24}$$

$$\geq \frac{(3\alpha+1)^2}{12}\mathscr{I}_4(\mu|\,\gamma) = (3\alpha+1)^2\frac{64}{3}\int_\Omega |\mathrm{D}z|^4 \,\mathrm{d}\gamma. \tag{3.25}$$

Moreover, even if $\alpha \in [-2/3, 0]$ (in this case the coefficient $\alpha(3\alpha+2)$ in (3.24) *is non-positive), we have*

$$2\mathscr{K}_\alpha(\mu|\,\gamma) + \hat{\mathscr{K}_\alpha}(\mu|\,\gamma) \geq (3\alpha+1)^2\Big(2\mathscr{K}_0(\mu|\,\gamma) + \hat{\mathscr{K}_0}(\mu|\,\gamma)\Big) \quad \forall\, \alpha \in [-2/3, 0]. \tag{3.26}$$

Proof. Observe that the integrands of $\mathscr{K}_\alpha$, $\hat{\mathscr{K}_\alpha}$ in (3.20) and (3.22) are (up to the coefficient 4)

$$I_\alpha := \sum_{i,j}\big(\partial^2_{ij}s + 4\alpha\partial_i z\partial_j z\big)^2, \qquad \hat{I}_\alpha := \big(\Delta_\gamma s + 4\alpha|\mathrm{D}z|^2\big)^2; \tag{3.27}$$

developing the squares and summing all the contributions we have

$$\begin{aligned}
2I_\alpha + \hat{I}_\alpha &= 2\sum_{i,j}\Big((\partial^2_{ij}s)^2 + 16\alpha^2(\partial_i z\partial_j z)^2 + 8\alpha\partial^2_{ij}s\,\partial_i z\,\partial_j z\Big) + \big(\Delta_\gamma s\big)^2 + 16\alpha^2|\mathrm{D}z|^4 + 8\alpha\Delta_\gamma s\,|\mathrm{D}z|^2\\
&= 2|\mathrm{D}^2 s|^2 + (\Delta_\gamma s)^2 + 3\cdot 16\alpha^2|\mathrm{D}z|^4 + 8\alpha\mathsf{A}(s)\mathrm{D}z\cdot\mathrm{D}z\\
&\overset{(3.1)}{=} 2|\mathrm{D}^2 s|^2 + (\Delta_\gamma s)^2 + 16(3\alpha^2+2\alpha)|\mathrm{D}z|^4 + 8\alpha\,\mathrm{div}_\gamma\big(|\mathrm{D}z|^2\mathrm{D}s\big)\\
&= 2I_0 + \hat{I}_0 + \frac{16}{3}\big((3\alpha+1)^2-1\big)|\mathrm{D}z|^4 + 8\alpha\,\mathrm{div}_\gamma\big(|\mathrm{D}z|^2\mathrm{D}s\big).
\end{aligned}$$

Integrating the above identity in Ω we obtain (3.24). Observe now that

$$\big\|\mathsf{A}(s)\big\|^2 = 9\,\big\|\tfrac{2}{3}\mathrm{D}^2 s + \tfrac{1}{3}(\Delta_\gamma s)\mathsf{Id}\big\|^2 \leq 6\big|\mathrm{D}^2 s\big|^2 + 3\big|\Delta_\gamma s\big|^2 = 6I_0 + 3\hat{I}_0 \tag{3.28}$$

recalling (3.4), we obtain

$$\frac{16}{3}\int_\Omega |\mathrm{D}z|^4 \,\mathrm{d}\gamma \leq \int_\Omega \big(2I_0 + \hat{I}_0\big)\,\mathrm{d}\gamma, \qquad \frac{1}{12}\mathscr{I}_4(\mu|\,\gamma) \leq 2\mathscr{K}_0(\mu|\,\gamma) + 2\hat{\mathscr{K}_0}(\mu|\,\gamma);$$

substituting this inequality in (3.24) we obtain (3.25) and (3.26). $\square$

The next corollary shows that a completely equivalent result could be obtained starting from the density v instead of its square root s:

Corollary 3.4. *Let Ω be a Lipschitz open set, γ be a reference measure as in § 2.2, and let $s \in W^{2,2}_\gamma(\Omega)$ be a nonnegative function with $\Delta_\gamma s \in L^2_\gamma(\Omega)$ and $\partial_{\boldsymbol{n}} s = 0$ on $\partial\Omega$.*
Then $v = s^2 \in W^{2,1}_\gamma(\Omega)$, $\Delta_\gamma v \in L^1_\gamma(\Omega)$, and

$$\begin{aligned}
2\mathscr{K}_1(\mu|\,\gamma) + \hat{\mathscr{K}_1}(\mu|\,\gamma) &= \int_\Omega \frac{2|\mathrm{D}^2 v|^2 + \big(\Delta_\gamma v\big)^2}{v}\,\mathrm{d}\gamma\\
&= 4\int_\Omega \Big(2|\mathrm{D}^2 s|^2 + \big(\Delta_\gamma s\big)^2 + 5\cdot 16|\mathrm{D}z|^4\Big)\,\mathrm{d}\gamma = 2\mathscr{K}_0(\mu|\,\gamma) + \hat{\mathscr{K}_0}(\mu|\,\gamma) + \frac{5}{4}\mathscr{I}_4(\mu|\,\gamma).
\end{aligned} \tag{3.29}$$

Conversely, if $v \in W^{2,1}_\gamma(\Omega)$ with $\Delta_\gamma v \in L^1_\gamma(\Omega)$, $\partial_{\boldsymbol n} v = 0$ on $\partial\Omega$, and

$$\int_\Omega \frac{|\mathrm{D}^2 v|^2}{v}\,\mathrm{d}\gamma + \int_\Omega \frac{|\Delta_\gamma v|^2}{v}\,\mathrm{d}\gamma < +\infty, \tag{3.30}$$

then $s = \sqrt{v} \in W^{2,2}_\gamma(\Omega)$, $\Delta_\gamma v \in L^2_\gamma(\Omega)$, and (3.4), (3.29) hold. In both cases, for every $\alpha, \beta \in \mathbb{R} \setminus \{-1/3\}$ there exist constants $C_{\alpha\beta} > 0$ such that

$$\mathscr{K}_\alpha(\mu|\,\gamma) + \hat{\mathscr{K}_\alpha}(\mu|\,\gamma) \le C_{\alpha\beta}\Big(\mathscr{K}_\beta(\mu|\,\gamma) + \hat{\mathscr{K}_\beta}(\mu|\,\gamma)\Big). \tag{3.31}$$

Finally, if V is L-Lipschitz continuous in Ω we have the estimate

$$\mathscr{K}_\alpha(\mu|\,\gamma) + \hat{\mathscr{K}_\alpha}(\mu|\,\gamma) \le C_{\alpha\beta}\Big(\mathscr{K}_\beta(\mu|\,\gamma) + L^2\mathscr{I}_2(\mu|\,\gamma)\Big). \tag{3.32}$$

Proof. (3.29) follows easily by the previous Theorem, simply by (3.24) for $\alpha = 1$.

The second part of the theorem is still a direct consequence of (3.29) if $\inf_\Omega v > 0$; in the general case it is sufficient to replace v by $v + \varepsilon$ and to pass to the limit as $\varepsilon \downarrow 0$. Finally, (3.32) follows directly from (3.34) of the next Lemma. □

Lemma 3.5. *For every $\alpha \in \mathbb{R}$ we have*

$$\hat{\mathscr{K}_\alpha}(\mu|\,\mathcal{L}^d) \le d\,\mathscr{K}_\alpha(\mu|\,\mathcal{L}^d), \tag{3.33}$$

and, if $\sup_{x\in\Omega} |\mathrm{D}V(x)| \le L$,

$$\hat{\mathscr{K}_\alpha}(\mu|\,\gamma) \le 2d\mathscr{K}_\alpha(\mu|\,\gamma) + 2L^2\mathscr{I}_2(\mu|\,\gamma). \tag{3.34}$$

Proof. We simply observe that for $\mu = s^2\gamma$ the integrand of $\hat{\mathscr{K}_\alpha}(\mu|\,\mathcal{L}^d)$ is the square of

$$s^{-1}\Big(s\Delta_\gamma s - \alpha|\mathrm{D}s|^2\Big) = s^{-1}\,\mathrm{tr}\,\Big(s\mathrm{D}^2 s - \alpha \mathrm{D}s \otimes \mathrm{D}s\Big) - \mathrm{D}s\cdot \mathrm{D}V \overset{(3.19)}{=} s^{-1}\,\mathrm{tr}\,\big(\mathsf{S}^\alpha\big) - \mathrm{D}s\cdot \mathrm{D}V,$$

so that (3.33) (corresponding to $V \equiv 0$) follows by the well known inequality for the trace of a symmetric matrix

$$\big|\,\mathrm{tr}\,\mathsf{S}^\alpha\big|^2 \le d\sum_i \big|\mathsf{S}^\alpha_{ii}\big|^2 \le d\big|\mathsf{S}^\alpha\big|^2.$$

A simple application of Cauchy-Schwarz inequality yields (3.34). □

Choosing $V \equiv 0, \gamma = \mathcal{L}^d$, and combining (3.24) with (3.33) we easily get:

Corollary 3.6 (**Second order estimates w.r.t. the Lebesgue measure**). *If Ω is Lipschitz and $\mu = r^2\mathcal{L}^d$ with $r \in W^{2,2}(\Omega)$ with $\partial_{\boldsymbol n} r = 0$ on $\partial\Omega$, then $\sqrt{r} \in W^{1,4}(\Omega)$ and*

$$\int_\Omega \Big(8\big|\mathrm{D}^2 r\big|^2 + 4\big|\Delta r\big|^2 + 64\big|\mathrm{D}\sqrt{r}\big|^4\Big)\,\mathrm{d}x \le (2+d)\mathscr{K}_{-1}(\mu|\,\mathcal{L}^d). \tag{3.35}$$

For $\Omega' \subset\subset \Omega$ we denote by $\mathscr{K}_{\alpha,\Omega'}(\mu|\,\gamma)$, $\hat{\mathscr{K}_{\alpha,\Omega'}}(\mu|\,\gamma)$ the corresponding *localized* functionals restricted to Ω'.

Corollary 3.7 (**Local bounds**). *For each $\Omega' \subset\subset \Omega'' \subset \Omega$ and $\alpha, \beta \in \mathbb{R} \setminus \{-1/3\}$ there exist constants $c_{\alpha\beta}, c_\beta > 0$, which also depends on $\sup_{\Omega''} |\mathrm{D}V|$ and on the dimension d, such that*

$$\mathscr{K}_{\alpha,\Omega'}(\mu|\,\gamma) + \hat{\mathscr{K}_{\alpha,\Omega'}}(\mu|\,\gamma) \le c_{\alpha\beta}\Big(\mathscr{K}_{\beta,\Omega''}(\mu|\,\gamma) + \mathscr{I}_2(\mu|\,\gamma)\Big) \quad \forall\,\mu \in D(\mathscr{K}_\beta), \tag{3.36}$$

$$\int_{\Omega'} \Big(|\mathrm{D}^2 s|^2 + |\Delta_\gamma s|^2 + |D\sqrt{s}|^4\Big)\,\mathrm{d}\gamma \le c_\beta\Big(\mathscr{K}_{\beta,\Omega''}(\mu|\,\gamma) + \mathscr{I}_2(\mu|\,\gamma)\Big) \quad \forall\,\mu = s^2\gamma \in D(\mathscr{K}_\beta). \tag{3.37}$$

The same result holds even for bounded *open subsets $\Omega' \subset \Omega'' \equiv \Omega$, provided the square root s of the density of μ belongs to $W^{2,2}_{\mathrm{loc}}(\overline{\Omega})$ and satisfies $\partial_{\boldsymbol n} s = 0$ on $\partial\Omega$.*

Proof. We use a standard localization argument as in the proof of Theorem 3.1. We can then apply the estimates of Theorem 3.3 and (3.34) relative to the open set Ω'', where L denotes the Lipschitz constant of V on Ω''. □

Corollary 3.8 (Lower semicontinuity). *Let $\alpha \neq -1/3$, and let $\mu_n = s_n^2\gamma \in D(\mathscr{K}_\alpha)$ be a sequence of probability measures satisfying*

$$\mu_n \rightharpoonup \mu \quad \textit{narrowly in } \mathcal{P}(\Omega), \quad \sup_n \Big(\mathscr{K}_\alpha(\mu_n|\,\gamma) + \mathscr{I}_2(\mu_n|\,\gamma)\Big) < +\infty. \tag{3.38}$$

Then $\mu = s^2\gamma \in \mathcal{P}_2^r(\Omega)$, $s \in W^{2,2}_{\rm loc}(\Omega)$, and

$$\mathscr{K}_\alpha(\mu|\,\gamma) \leq \liminf_{n\to\infty} \mathscr{K}_\alpha(\mu_n|\,\gamma) < +\infty. \tag{3.39}$$

If moreover $\partial_{\boldsymbol{n}} s_n = 0$ on $\partial\Omega$ and V is Lipschitz on bounded sets, then

$$s \in W^{2,2}(\Omega') \quad \partial_{\boldsymbol{n}} s = 0 \textit{ on } \partial\Omega \quad \textit{for every bounded subset } \Omega' \subset \Omega. \tag{3.40}$$

Proof. By the local bounds of Corollary 3.7 we can extract a subsequence, still labeled by s_n, strongly converging to s in $W^{1,2}_{\rm loc}(\Omega)$ with $\mathrm{D}^2 s_n \rightharpoonup \mathrm{D}^2 s$ in $L^2_{\rm loc}(\Omega)$, so that the tensors $\mathsf{S}_n^{(\alpha)} = s_n\mathrm{D}^2 s_n + \alpha \mathrm{D}s_n \otimes \mathrm{D}s_n$, defined as in (3.19), weakly converge to $\mathsf{S}^{(\alpha)} = s\mathrm{D}^2 s + \alpha\mathrm{D}s \otimes \mathrm{D}s$ in $L^1_{\rm loc}(\Omega)$. Applying Proposition 2.2 we conclude. □

4. First variation of the Fisher information

In this section we find a first expression for the strong "*Wasserstein*" subdifferential of the Fisher information functionals

$$\mathscr{G}(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) = 2\int_\Omega \Big|\mathrm{D}\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\gamma}}\Big|^2\,\mathrm{d}\gamma, \quad \mathscr{F}^f(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \langle f,\mu\rangle, \tag{4.1}$$

according to Definition 2.10; recall that $\mathscr{G}$ (resp. $\mathscr{F}^f$) is finite if and only if $\mu \ll \gamma$ and $s := \sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\gamma}} \in W^{1,2}_\gamma(\Omega)$ (resp. $r := \sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}} \in W^{1,2}(\Omega)$).

We first consider the case of a general measure γ, which also cover the particular case of the Lebesgue measure $\mathcal{L}^d$. Recalling the "smooth perturbation" property S3 of the subdifferential discussed in §2.5, we will immediately obtain the subdifferential of $\mathscr{F}^f(\mu)$.

Following the approach originally introduced by [JKO98], we take inner variations of the functional $\mathscr{G}$ at a point $\mu = s^2\gamma \in D(\partial_s\mathscr{G})$ along the curves originated by flowing the density μ through a smooth vector field. Let us first collect some preliminary and elementary properties.

4.1. **Flows and vector fields.** Let Ω be a convex subset of $\mathbb{R}^d$ and $\boldsymbol{\zeta}:\mathbb{R}^d\to\mathbb{R}^d$ be a vector field such that

$$\boldsymbol{\zeta}\in C^2(\mathbb{R}^d;\mathbb{R}^d), \quad \|\mathrm{D}\boldsymbol{\zeta}(x)\| + \|\mathrm{D}^2\boldsymbol{\zeta}(x)\| \le A \quad \forall\, x\in\mathbb{R}^d, \quad \boldsymbol{\zeta}\cdot\boldsymbol{n} = 0 \quad \mathscr{H}^{d-1}\text{-a.e. on } \partial\Omega, \tag{4.2}$$

for a suitable constant $A > 0$. We denote by $\boldsymbol{X}_t : \mathbb{R}^d \to \mathbb{R}^d$ the flow associated to $\boldsymbol{\zeta}$, i.e. the family of diffeomorphisms (which are globally defined, being $\boldsymbol{\zeta}$ Lipschitz) satisfying the system of ODE for $(x,t)\in\mathbb{R}^d\times\mathbb{R}$

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}_t(x) = \boldsymbol{\zeta}(\boldsymbol{X}_t(x)), \\ \quad \boldsymbol{X}_0(x) = x. \end{cases} \tag{4.3}$$

By Nagumo's Theorem [Nag42, Bre70] the open set Ω is an invariant region for $\boldsymbol{X}_t$, i.e.

$$\Omega_t := \boldsymbol{X}_t(\Omega) = \Omega \quad \forall t\in\mathbb{R}, \tag{4.4}$$

and the map $x\mapsto \boldsymbol{X}_t(x)$ is a C^2 diffeomorphism with inverse $\boldsymbol{X}_{-t}$. Setting

$$\mathsf{D}_t(x) := \mathrm{D}_x\boldsymbol{X}_t(x), \quad \mathsf{G}_t(x) := \mathsf{D}_t^{-1}(x), \quad J_t(x) := \det\mathsf{D}_t(x), \quad \ell_t(x) := \log J_t(x), \tag{4.5}$$

we have

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathsf{D}_t(x) = \mathrm{D}\boldsymbol{\zeta}(\boldsymbol{X}_t(x))\mathsf{D}_t(x) \\ \quad \mathrm{D}_0(x) = I, \end{cases} \qquad \begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}J_t(x) = \operatorname{div}\boldsymbol{\zeta}(\boldsymbol{X}_t(x))J_t(x) \\ \quad J_0(x) = 1, \end{cases} \tag{4.6}$$

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\ell_t(x) = \operatorname{div}\boldsymbol{\zeta}(\boldsymbol{X}_t(x)) \\ \quad \ell_0(x) = 0, \end{cases} \qquad \begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{D}\ell_t(x) = \mathrm{D}\operatorname{div}\boldsymbol{\zeta}(\boldsymbol{X}_t(x))\mathsf{D}_t(x) \\ \quad \mathrm{D}\ell_0(x) = 0. \end{cases} \tag{4.7}$$

Moreover, a constant $C > 0$ depending on A exists such that the uniform bounds hold:

$$\begin{cases} |\boldsymbol{X}_t(x)| + |\dot{\boldsymbol{X}}_t(x)| \le C(1+|x|), \\ \|\mathsf{D}_t\| + \|\dot{\mathsf{D}}_t\| + \|\mathsf{G}_t\| + \|\dot{\mathsf{G}}_t\| + \|\mathrm{D}\ell_t\| + \|\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{D}\ell_t\| \le C \end{cases} \quad \text{in } \mathbb{R}^d\times(-1,1). \tag{4.8}$$

For $\mu_0 = v_0\cdot\gamma \in \mathcal{P}_2^r(\Omega)$ we can consider the curve in $\mathcal{P}_2^r(\Omega)$ given by

$$\mu_t = v_t\cdot\gamma := (\boldsymbol{X}_t)_\#\mu_0, \tag{4.9}$$

with

$$v_t(\boldsymbol{X}_t(x))\exp\Big(\ell_t(x) - V(\boldsymbol{X}_t(x))\Big) = v_0(x)\exp\Big(-V(x)\Big) \tag{4.10}$$

and, for $s_t := \sqrt{v_t}$,

$$s_t(\boldsymbol{X}_t(x))\exp\Big(\tfrac{1}{2}\ell_t(x) - \tfrac{1}{2}V(\boldsymbol{X}_t(x))\Big) = s_0(x)\exp\Big(-\tfrac{1}{2}V(x)\Big) \qquad \forall\, x\in\Omega. \tag{4.11}$$

Since $\ell_0(x)=0$ and $\boldsymbol{X}_0(x)=x$, (4.11) can be written as

$$s_t(\boldsymbol{X}_t(x))\mathsf{e}_t(x)=s_0(x)\mathsf{e}_0(x)\quad\text{where}\quad \mathsf{e}_t(x):=\exp\Big(\tfrac12\ell_t(x)-\tfrac12 V(\boldsymbol{X}_t(x))\Big). \tag{4.12}$$

Remark 4.1 (The case of the dilation group in a convex cone). If Ω is a (convex) cone, i.e.

$$\alpha\Omega=\Omega\quad\forall\,\alpha>0, \tag{4.13}$$

then we can always choose the vector field

$$\boldsymbol{\zeta}(x):=x,\quad\text{corresponding to the flow}\quad \boldsymbol{X}_t(x)=e^t x. \tag{4.14}$$

4.2. **The derivative of the Fisher information along smooth curves.** The next result provides the main characterization of the "Wasserstein" subdifferential of $\mathscr{G}$, which is strictly related to the formulation ($\mathrm{EE}_{3,\gamma}$) we discussed in Definition 1.4: according to that definition, ($\mathrm{EE}_{3,\gamma}$) can be written as

$$\partial_t v+\mathrm{div}_\gamma\tilde{\boldsymbol{q}}=0,\quad \tilde{\boldsymbol{q}}_j=\tilde\partial_j\Delta_\gamma v-\sum_i\tilde\partial_i\Big(\frac{\partial_i v\,\partial_j v}{v}\Big)\quad\text{in }\mathscr{D}'(\Omega\times(0,+\infty)), \tag{4.15}$$

and it is interesting to compare (4.15) with the next (4.17).

Theorem 4.2 (**First variation of $\mathscr{G}$**). *Let us suppose that $s\in W^{1,2}_\gamma(\Omega)$, $\mu=v\gamma=s^2\gamma\in\mathcal{P}_2^r(\Omega)$, $\boldsymbol{\xi}\in\partial_s\mathscr{G}(\mu)$, and let $\boldsymbol{\zeta}:\mathbb{R}^d\to\mathbb{R}^d$ be a vector field satisfying* (4.2)*. If either $\boldsymbol{\zeta}$ has compact support or the potential V has bounded second order derivatives, then*

$$\int_\Omega\boldsymbol{\xi}(x)\cdot\boldsymbol{\zeta}(x)\,\mathrm{d}\mu(x)=-\int_\Omega\Big(4\mathrm{D}\boldsymbol{\zeta}(x)\mathrm{D}s(x)\cdot\mathrm{D}s(x)+2s(x)\mathrm{D}\big(\mathrm{div}_\gamma\boldsymbol{\zeta}(x)\big)\cdot\mathrm{D}s(x)\Big)\,\mathrm{d}\gamma(x). \tag{4.16}$$

In particular the vector $\tilde{\boldsymbol{q}}:=-v\boldsymbol{\xi}\in L^1_\gamma(\Omega;\mathbb{R}^d)$ satisfies

$$\begin{aligned}\tilde{\boldsymbol{q}}_j&=\partial_j\Delta_\gamma v-4\sum_i\tilde\partial_i(\partial_i s\,\partial_j s)=\partial_j\Delta_\gamma v-\sum_i\tilde\partial_i\Big(\frac{\partial_i v\,\partial_j v}{v}\Big),\\ \mathrm{div}_\gamma\tilde{\boldsymbol{q}}&=\Delta_\gamma^2 v-\sum_{i,j}\tilde\partial^2_{ij}\Big(\frac{\partial_i v\,\partial_j v}{v}\Big)\end{aligned} \tag{4.17}$$

in the sense of distributions in $\mathscr{D}'(\Omega)$.

Proof. Let $\boldsymbol{X}_t$, $t\ge0$, be the flow associated to $\boldsymbol{\zeta}$ as in (4.3) and let us set $\mu_t:=(\boldsymbol{X}_t)_\#\mu$. (4.16) is an immediate consequence of the Definition (2.68) and of (2.74) if we show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{G}(\mu_t)|_{t=0}=-\int_\Omega\Big(4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}s\cdot\mathrm{D}s+2s\,\mathrm{D}\big(\mathrm{div}_\gamma\boldsymbol{\zeta}\big)\cdot\mathrm{D}s\Big)\,\mathrm{d}\gamma. \tag{4.18}$$

Being $s_0\in W^{1,2}_\gamma(\Omega)$, by (4.10) and (4.11) we know that $s_t\in W^{1,2}_\gamma(\Omega)$ and $v_t=s_t^2\in W^{1,1}_\gamma(\Omega)$; keeping the same notation of (4.5), and setting

$$\mathbf{e}_t(x):=\frac12\mathrm{D}\ell_t(x)-\frac12\mathrm{D}V(\boldsymbol{X}_t(x))\mathsf{D}_t(x)\quad\text{so that}\quad \mathrm{D}\mathsf{e}_t(x)=\mathsf{e}_t(x)\mathbf{e}_t(x),$$

a differentiation with respect to x of (4.12) yields

$$\mathrm{D}s_t(\boldsymbol{X}_t)\mathsf{e}_t\mathsf{D}_t+s_t(\boldsymbol{X}_t)\mathsf{e}_t\mathbf{e}_t=\mathrm{D}s_0\mathsf{e}_0+s_0\mathsf{e}_0\mathbf{e}_0=\mathsf{e}_0\big(\mathrm{D}s_0+s_0\mathbf{e}_0\big), \tag{4.19}$$

so that, by (4.5),

$$\mathrm{D}s_t(\boldsymbol{X}_t)\mathsf{e}_t=\Big(\mathsf{e}_0\big(\mathrm{D}s_0+s_0\mathbf{e}_0\big)-s_t(\boldsymbol{X}_t)\mathsf{e}_t\mathbf{e}_t\Big)\mathsf{G}_t\overset{(4.12)}{=}\mathsf{e}_0\Big(\mathrm{D}s_0+s_0\big(\mathbf{e}_0-\mathbf{e}_t\big)\Big)\mathsf{G}_t=\mathsf{e}_0\boldsymbol{a}_t, \tag{4.20}$$

where we set

$$\boldsymbol{a}_t:=\Big(\mathrm{D}s_0+s_0\big(\mathbf{e}_0-\mathbf{e}_t\big)\Big)\mathsf{G}_t=\Big(\mathrm{D}s_0-\tfrac12 s_0\big(\mathrm{D}\ell_t-\mathrm{D}V(\boldsymbol{X}_t)\mathsf{D}_t+\mathrm{D}V\big)\Big)\mathsf{G}_t. \tag{4.21}$$

The time derivative of $\boldsymbol{a}_t$ is

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}_t=\Big(\mathrm{D}s_0+s_0\big(\mathbf{e}_0-\mathbf{e}_t\big)\Big)\dot{\mathsf{G}}_t-s_0\dot{\mathbf{e}}_t\mathsf{G}_t, \tag{4.22}$$

with

$$\dot{\mathsf{e}}_t = \frac{1}{2}\Big(\mathrm{D}\,\mathrm{div}\,\boldsymbol{\zeta}(\boldsymbol{X}_t) - \mathrm{D}^2V(\boldsymbol{X}_t)\boldsymbol{\zeta}(\boldsymbol{X}_t) - \mathrm{D}V(\boldsymbol{X}_t)\mathrm{D}\boldsymbol{\zeta}(\boldsymbol{X}_t)\Big)\mathsf{D}_t. \tag{4.23}$$

Substituting in (4.22) and observing that

$$\dot{\mathsf{G}}_t = -\mathsf{G}_t\dot{\mathsf{D}}_t\mathsf{G}_t = -\mathsf{G}_t\mathrm{D}\boldsymbol{\zeta}(\boldsymbol{X}_t)\mathsf{D}_t\mathsf{G}_t = -\mathsf{G}_t\mathrm{D}\boldsymbol{\zeta}(\boldsymbol{X}_t), \tag{4.24}$$

we get

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}_t = -\Big(\mathrm{D}s_0 + \tfrac{1}{2}s_0\mathrm{D}\ell_t - \tfrac{1}{2}s_0\mathrm{D}V\Big)\mathsf{G}_t\mathrm{D}\boldsymbol{\zeta}(\boldsymbol{X}_t) - \tfrac{1}{2}s_0\Big(\mathrm{D}\,\mathrm{div}\,\boldsymbol{\zeta}(\boldsymbol{X}_t) - \mathrm{D}^2V(\boldsymbol{X}_t)\boldsymbol{\zeta}(\boldsymbol{X}_t)\Big). \tag{4.25}$$

Let us check that there exists a suitable constant C such that in $\Omega\times(-1,1)$

$$|\boldsymbol{a}_t(x)| \le C\big(|\mathrm{D}s_0| + |s_0|(1+|x|)\big), \tag{4.26}$$

$$\Big|\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}_t\Big| \le C\big(|\mathrm{D}s_0| + |s_0|(1+|x|)\big). \tag{4.27}$$

For, when $\boldsymbol{\zeta}$ has a compact support $K \subset\subset \Omega$ then $\boldsymbol{X}_t(x) = x,\ \mathsf{G}_t(x) = I,\ \dot{\mathsf{G}}_t(x) = 0$ in $\Omega\setminus K$, so that $\mathrm{D}^2V(\boldsymbol{X}_t) - \mathrm{D}^2V \equiv 0$ in $\Omega\setminus K$ and (4.26), (4.27) hold by (4.8) since $\mathrm{D}V$ and D^2V are locally bounded.

In the second case, when V has bounded second derivatives, then $\mathrm{D}V$ has a linear growth, and (4.26), (4.27) still follow from (4.8).

An integration in Ω and the change of variable formula yield

$$\int_\Omega |\mathrm{D}s_t(y)|^2\,\mathrm{d}\gamma(y) = \int_\Omega \Big|\mathrm{D}s_t(y)\exp\big(-\tfrac{1}{2}V(y)\big)\Big|^2\,\mathrm{d}y = \int_\Omega |\mathrm{D}s_t(\boldsymbol{X}_t(x))\mathsf{e}_t(x)|^2\,\mathrm{d}x \overset{(4.20)}{=} \int_\Omega \big|\boldsymbol{a}_t\big|^2\,\mathrm{d}\gamma.$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{G}_t|_{t=0} \overset{(4.6)}{=} -\mathrm{D}\boldsymbol{\zeta},\quad \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{D}\ell_t\Big)_{|t=0} \overset{(4.7)}{=} \mathrm{D}\,\mathrm{div}\,\boldsymbol{\zeta},\quad \mathsf{D}_0 = \mathsf{G}_0 = \mathsf{Id},\quad \mathrm{D}\ell_0 = 0,\quad \boldsymbol{a}_0 = \mathrm{D}s_0,$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}_t|_{t=0} &\overset{(4.25)}{=} -\mathrm{D}s_0\mathrm{D}\boldsymbol{\zeta} + \tfrac{1}{2}s_0\mathrm{D}V\mathrm{D}\boldsymbol{\zeta} - \tfrac{1}{2}s_0\mathrm{D}\,\mathrm{div}\,\boldsymbol{\zeta} + \tfrac{1}{2}s_0\mathrm{D}^2V\boldsymbol{\zeta}\\ &= -\mathrm{D}s_0\mathrm{D}\boldsymbol{\zeta} - \tfrac{1}{2}s_0\mathrm{D}\,\mathrm{div}\,\boldsymbol{\zeta} + \tfrac{1}{2}s_0\mathrm{D}\big(\mathrm{D}V\cdot\boldsymbol{\zeta}\big) \overset{(1.10)}{=} -\mathrm{D}s_0\mathrm{D}\boldsymbol{\zeta} - \tfrac{1}{2}s_0\mathrm{D}\big(\mathrm{div}_\gamma\boldsymbol{\zeta}\big),\end{aligned}$$

thanks to (4.27) and Lebesgue Dominated Convergence Theorem we have

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |\mathrm{D}s_t(y)|^2\,\mathrm{d}\gamma(y)\Big)_{|t=0} = -2\int_\Omega \mathrm{D}s_0\cdot\Big(\mathrm{D}s_0\mathrm{D}\boldsymbol{\zeta} + \tfrac{1}{2}s_0\mathrm{D}\big(\mathrm{div}_\gamma\boldsymbol{\zeta}\big)\Big)\,\mathrm{d}\gamma(x). \qquad \square$$

Combining Theorem 4.2 with the smooth perturbation argument of §2.5 we immediately obtain a characterization of the "Wasserstein" subdifferential of $\mathscr{F}^f$, which is strictly related to equation (EE$_3$)

$$\partial_t u + \mathrm{div}\,\boldsymbol{q} = 0,\quad \boldsymbol{q}_j = \partial_j\Delta u - \sum_i \partial_i\Big(\frac{\partial_i u\,\partial_j u}{u}\Big) - u\partial_j f. \tag{4.28}$$

Corollary 4.3 (First variation of $\mathscr{F}^f$). *Let us suppose that $\mu = r^2\mathscr{L}^d \in \mathscr{P}_2^r(\Omega)$ for $r\in W^{1,2}(\Omega)$, let*

$$\mathscr{F}^f(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\mathscr{L}^d) + \langle f,\mu\rangle, \tag{4.29}$$

with f satisfying (1.4), let $\boldsymbol{\xi}\in\partial_s\mathscr{F}^f(\mu)$, and let $\boldsymbol{\zeta}:\mathbb{R}^d\to\mathbb{R}^d$ be a vector field satisfying (4.2). Then

$$\int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\int_\Omega\Big(4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}r\cdot\mathrm{D}r + 2r\mathrm{D}\big(\mathrm{div}\,\boldsymbol{\zeta}\big)\cdot\mathrm{D}r\Big)\,\mathrm{d}x + \int_\Omega \mathrm{D}f\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu. \tag{4.30}$$

and $\boldsymbol{q} := -u\boldsymbol{\xi}\in L^1_\gamma(\Omega;\mathbb{R}^d)$ satisfies

$$\boldsymbol{q}_j = \partial_j\Delta_\gamma u - 4\sum_i\partial_i(\partial_j r\,\partial_i r) - u\partial_j f = \partial_j\Delta_\gamma u - \sum_i\partial_i\Big(\frac{\partial_j u\,\partial_i u}{u}\Big) - u\partial_j f. \tag{4.31}$$

Observe that under the assumptions (1.5a,b,c,d,e) which yield

$$\mathscr{G}(\mu) = \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) = \frac{1}{2}\mathscr{I}_2(\mu_t|\,\mathcal{L}^d) + \int_\Omega f(x)\,\mathrm{d}\mu_t(x) = \mathscr{F}^f(\mu), \tag{4.32}$$

the expressions of the subdifferential of $\mathscr{G}$ and $\mathscr{F}^f$ given by (4.16) (in terms of $s := (\mathrm{d}\mu/\mathrm{d}\gamma)^{1/2}$) and (4.30) (in terms of $r = (\mathrm{d}\mu/\mathrm{d}\mathcal{L}^d)^{1/2}$) coincide.

A particular but interesting case of the previous formulae is provided by the next corollary: we keep the same notation of the previous Theorem.

Corollary 4.4. *Suppose that Ω is a (convex) cone, V has bounded second order derivatives, f satisfies (1.4), $\mu = v\gamma = u\mathcal{L}^d \in \mathcal{P}_2^r(\Omega)$, and that $\boldsymbol{\xi}_1 \in \partial_s\mathscr{G}(\mu)$, $\boldsymbol{\xi}_2 \in \partial_s\mathscr{F}^f(\mu)$. Then*

$$-\int_\Omega \boldsymbol{\xi}_1(x)\cdot x\,\mathrm{d}\mu = \mathscr{I}_2(\mu|\,\gamma) - \int_\Omega \Big(\mathrm{D}v\cdot\mathrm{D}V + \mathrm{D}^2V\mathrm{D}v\cdot x\Big)\,\mathrm{d}\gamma, \tag{4.33}$$

$$-\int_\Omega \boldsymbol{\xi}_2(x)\cdot x\,\mathrm{d}\mu = \mathscr{I}_2(\mu|\,\mathcal{L}^d) - \int_\Omega \mathrm{D}f\cdot x\,\mathrm{d}\mu(x). \tag{4.34}$$

In the particular case (1.6) of the Lebesgue Measure $\gamma = \mathcal{L}^d$, $f \equiv 0$ (4.33) and (4.34) read

$$\boldsymbol{\xi} = \partial_s\Big(\tfrac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d)\Big) \quad\Rightarrow\quad -\int_{\mathbb{R}^d}\boldsymbol{\xi}(x)\cdot x\,\mathrm{d}\mu = \mathscr{I}_2(\mu|\,\mathcal{L}^d). \tag{4.35}$$

When $\gamma = \gamma_\lambda$ is the centered Gaussian measure with variance λ^{-1} of (1.8), (4.33) becomes

$$\boldsymbol{\xi} = \partial_s\Big(\tfrac{1}{2}\mathscr{I}_2(\mu|\,\gamma_\lambda)\Big) \quad\Rightarrow\quad \begin{cases} -\displaystyle\int_{\mathbb{R}^d}\boldsymbol{\xi}(x)\cdot x\,\mathrm{d}\mu = \mathscr{I}_2(\mu|\,\gamma) + 2\lambda d - 2\lambda^2\mathsf{m}_2^2(\mu) \\ \qquad\qquad\qquad\qquad = \mathscr{I}_2(\mu|\,\mathcal{L}^d) - \lambda^2\mathsf{m}_2^2(\mu). \end{cases} \tag{4.36}$$

5. First variation of the Fisher information w.r.t. the Fokker-Planck equation

In this section we derive further information on the Wasserstein-subdifferential of the Fisher Information functional $\mathscr{I}_2(\cdot|\,\gamma)$ by evaluating its derivative along the flow $\nu_t = \rho_t\cdot\gamma$ generated by the Fokker-Planck equation starting from a given $\nu_0=\rho_0\cdot\gamma$ with $\mathscr{I}_2(\nu_0|\,\gamma)<+\infty$ (see §2.6). We will show that this derivative provides an explicit control of the functionals $\mathscr{K}_{-1}(\nu_t|\,\gamma)$ and $\mathscr{I}_2^V(\nu_t|\,\gamma)$ we introduced in (3.14a,b,c), (3.20), and (2.53). It is interesting to compare the present estimates with the results of [DPL04a, DPL04b].

We set for $\nu=\rho\cdot\gamma$ with $\sigma=\sqrt{\rho}\in W^{2,2}_{\rm loc}(\Omega)$

$$\mathscr{P}(\nu|\,\gamma)=\mathscr{K}_{-1}(\nu|\,\gamma)+\mathscr{I}_2^V(\nu|\,\gamma)=4\int_\Omega\Big(\Big|\frac{\sigma\mathrm{D}^2\sigma-\mathrm{D}\sigma\otimes\mathrm{D}\sigma}{\sigma}\Big|^2+\mathrm{D}^2V\mathrm{D}\sigma\cdot\mathrm{D}\sigma\Big)\,\mathrm{d}\gamma \tag{5.1a}$$

$$=\int_\Omega\Big(\big|\mathrm{D}^2\log\rho\big|^2+\mathrm{D}^2V\mathrm{D}\log\rho\cdot\mathrm{D}\log\rho\Big)\,\mathrm{d}\nu, \tag{5.1b}$$

recalling that the last expression (5.1b) is meaningful only when ess-inf $\rho>0$. Let us first consider the case of a bounded smooth domain.

5.1. The case of a bounded smooth domain.

Theorem 5.1 (**Fisher information and Fokker-Planck equation**). *Let us suppose that Ω is a bounded, smooth, and convex open set, that V,γ satisfy* (2.8)*, let $\nu_0=\rho_0\gamma\in\mathcal{P}_2^r(\Omega)$ with*

$$\rho_0\in L^2(\Omega),\quad \rho_{\rm inf}:=\text{ess-inf}\,\rho_0>0,\quad \mathscr{I}_2(\nu_0|\,\gamma)<+\infty, \tag{5.2}$$

and let $\nu_t=\rho_t\gamma$ be the solution of the Fokker-Planck equation with Neumann boundary conditions

$$\partial_t\rho-\Delta_\gamma\rho=0\quad\text{in }\Omega_T,\quad \partial_{\boldsymbol n}\rho=0\quad\text{on }(\partial\Omega)_T, \tag{5.3}$$

given by Theorem 2.17. Then the map $t\mapsto\mathscr{I}_2(\nu_t|\,\gamma)$ is absolutely continuous and

$$-\frac{d}{dt}\frac12\mathscr{I}_2(\nu_t|\,\gamma)\ge\mathscr{P}(\nu_t|\,\gamma)\quad\text{for a.e. }t>0. \tag{5.4}$$

In particular, for $h>0$ we have

$$\frac12\mathscr{I}_2(\nu_h|\,\gamma)+\int_0^h\mathscr{P}(\nu_t|\,\gamma)\,\mathrm{d}t\le\frac12\mathscr{I}_2(\nu_0|\,\gamma), \tag{5.5}$$

$$\frac{e^{2\lambda h}}{2}\mathscr{I}_2(\nu_h|\,\gamma)+\int_0^h e^{2\lambda t}\Big(\mathscr{P}(\nu_t|\,\gamma)-\lambda\mathscr{I}_2(\nu_t|\,\gamma)\Big)\,\mathrm{d}t\le\frac12\mathscr{I}_2(\nu_0|\,\gamma). \tag{5.6}$$

Proof. Let us first recall that the norms of $W^{k,p}_\gamma(\Omega)$ and $W^{k,p}(\Omega)$ are equivalent. Being $\mathrm{D}V\in L^\infty(\Omega)$ and Ω smooth, standard estimates in $W^{2,2}(\Omega)$ for Neumann problems, (2.127), and the homogeneous Neumann boundary conditions on $\partial\Omega$ yield

$$\rho\in C^\infty([\delta,+\infty);W^{2,2}_\gamma(\Omega))\quad\forall\,\delta>0. \tag{5.7}$$

Since ess-inf $\rho_t\ge\rho_{\rm inf}>0$ by (2.130), we deduce that

$$\sigma=\sqrt{\rho}\in C^\infty([\delta,+\infty);W^{2,2}_\gamma(\Omega)),\quad \Delta_\gamma\sigma\in C^\infty([\delta,+\infty);L^2_\gamma(\Omega))\quad\forall\,\delta>0. \tag{5.8}$$

Simple calculations for $t>0, x\in\Omega$ show

$$2\sigma\partial_t\sigma=2\sigma\Delta_\gamma\sigma+2|\mathrm{D}\sigma|^2,\qquad \partial_t\sigma=\Delta_\gamma\sigma+\frac{|\mathrm{D}\sigma|^2}{\sigma}=\Delta_\gamma\sigma+4|\mathrm{D}z|^2\quad z:=\sqrt{\sigma}. \tag{5.9}$$

We thus have enough regularity to evaluate the (opposite of the) time derivative of $\mathscr{I}_2(\nu_t|\,\gamma)$ for $t>0$ (which is a convex functional w.r.t. σ)

$$-\frac12\frac{d}{dt}\mathscr{I}_2(\nu_t|\,\gamma)=-\frac{d}{dt}2\int_\Omega\big|\mathrm{D}\sigma\big|^2\,\mathrm{d}\gamma=-4\int_\Omega\mathrm{D}\sigma\cdot\mathrm{D}(\partial_t\sigma)\,\mathrm{d}\gamma\overset{(1.11)}{=}4\int_\Omega\Delta_\gamma\sigma\partial_t\sigma\,\mathrm{d}\gamma \tag{5.10}$$

$$\overset{(5.9)}{=}4\int_\Omega\Delta_\gamma\sigma\Big(\Delta_\gamma\sigma+\frac{|\mathrm{D}\sigma|^2}{\sigma}\Big)\,\mathrm{d}\gamma=4\int_\Omega|\Delta_\gamma\sigma|^2\,\mathrm{d}\gamma+4^2\int_\Omega\Delta_\gamma\sigma\,|\mathrm{D}z|^2\,\mathrm{d}\gamma; \tag{5.11}$$

recall that by (5.8) and (5.9) (or by the general results of Theorem 3.1) $\mathrm{D}z\in C^0(\delta,+\infty;W^{1,4}_\gamma(\Omega))$.

We want to exploit now the convexity of Ω to prove that $-\frac{1}{2}\frac{d}{dt}\mathscr{I}_2(\nu_t|\,\gamma) \ge \mathscr{P}(\nu_t|\,\gamma)$. Applying the next Lemma 5.2 we obtain

$$-\frac{1}{2}\frac{d}{dt}\mathscr{I}_2(\nu_t|\,\gamma) \ge 4\int_\Omega \big|\mathrm{D}^2\sigma\big|^2\,\mathrm{d}\gamma + 4^2\int_\Omega \Delta_\gamma\sigma\,|\mathrm{D}z|^2\,\mathrm{d}\gamma + 4\int_\Omega \mathrm{D}^2V\mathrm{D}\sigma\cdot\mathrm{D}\sigma\,\mathrm{d}\gamma. \tag{5.12}$$

We now observe that

$$\begin{aligned}4\int_\Omega \big|\mathrm{D}^2\sigma\big|^2\,\mathrm{d}\gamma &= 4\int_\Omega \Big(\big|\mathrm{D}^2\sigma - 4\mathrm{D}z\otimes\mathrm{D}z\big|^2 - 4^2|\mathrm{D}z|^4 + 2\cdot 4\mathrm{D}^2\sigma\mathrm{D}z\cdot\mathrm{D}z\Big)\,\mathrm{d}\gamma\\ &= \mathscr{K}_{-1}(\nu_t|\,\gamma) - 4^3\int_\Omega |\mathrm{D}z|^4\,\mathrm{d}\gamma + 2\cdot 4^2\int_\Omega \mathrm{D}^2\sigma\mathrm{D}z\cdot\mathrm{D}z\,\mathrm{d}\gamma.\end{aligned}$$

Substituting this last expression in (5.12) and recalling the definition (3.2) of the tensor $\mathsf{A}(\sigma) := 2\mathrm{D}^2\sigma + \Delta_\gamma\sigma\mathsf{Id}$, we get

$$\begin{aligned}-\frac{1}{2}\frac{d}{dt}\mathscr{I}_2(\gamma|\,\nu) &\ge \mathscr{K}_{-1}(\nu_t|\,\gamma) + 4\int_\Omega \mathrm{D}^2V\mathrm{D}\sigma\cdot\mathrm{D}\sigma\,\mathrm{d}\gamma + 4^2\int_\Omega \mathsf{A}(\sigma)\mathrm{D}z\cdot\mathrm{D}z\,\mathrm{d}\gamma - 4^3\int_\Omega |\mathrm{D}z|^4\,\mathrm{d}\gamma\\ &\overset{(3.4)}{=} \mathscr{K}_{-1}(\nu_t|\,\gamma) + \mathscr{I}_2^V(\nu_t|\,\gamma),\end{aligned}$$

which yields (5.4). (5.5) simply follows by integrating (5.4) between 0 and $h > 0$, recalling that the map $t \mapsto \mathscr{I}_2(\nu_t|\,\gamma)$ is continuous at $t = 0$. In order to get (5.6) we multiply (5.4) by $e^{2\lambda t}$: since

$$-\frac{1}{2}e^{2\lambda t}\frac{d}{dt}\mathscr{I}_2(\nu_t|\,\gamma) = -\frac{1}{2}\frac{d}{dt}\Big(e^{2\lambda t}\mathscr{I}_2(\nu_t|\,\gamma)\Big) + \lambda e^{2\lambda t}\mathscr{I}_2(\nu_t|\,\gamma),$$

we get

$$\frac{1}{2}\frac{d}{dt}\Big(e^{2\lambda t}\mathscr{I}_2(\nu_t|\,\gamma)\Big) + e^{2\lambda t}\Big(\mathscr{P}(\nu_t|\,\gamma) - \lambda\mathscr{I}_2(\nu_t|\,\gamma)\Big) \le 0,$$

and therefore (5.6). □

The next lemma is well known when $V \equiv 0$, $\gamma = \mathscr{L}^d$ (see e.g. [Gri85]).

Lemma 5.2. *Let Ω be a smooth convex set, let V, γ be as in (2.8), and let $\sigma \in C^3(\overline{\Omega})$ with $\partial_{\boldsymbol{n}}\sigma = 0$ on $\partial\Omega$. Then*

$$\mathrm{D}^2\sigma\mathrm{D}\sigma\cdot\boldsymbol{n} = \sum_{i,j}\partial^2_{ij}\sigma\,\partial_i\sigma\,\boldsymbol{n}_j \le 0 \quad \text{on } \partial\Omega. \tag{5.13}$$

If $\sigma \in W^{2,2}_\gamma(\Omega)$ with $\partial_{\boldsymbol{n}}\sigma = 0$ on $\partial\Omega$, then

$$\int_\Omega \big|\Delta_\gamma\,\sigma\big|^2\,\mathrm{d}\gamma \ge \int_\Omega \Big(\big|\mathrm{D}^2\sigma\big|^2 + \mathrm{D}^2V\mathrm{D}\sigma\cdot\mathrm{D}\sigma\Big)\,\mathrm{d}\gamma. \tag{5.14}$$

Proof. Let $d_\Omega(x) := d(x,\Omega) - d(x,\mathbb{R}^d\setminus\Omega)$ be the "signed distance" function from the boundary of Ω; since Ω is convex, d_Ω is a convex function in $\mathbb{R}^d$ and it is also smooth in a suitable neighborhood of $\partial\Omega$, being Ω smooth. Moreover,

$$\mathrm{D}d_\Omega = \boldsymbol{n} \quad \text{on } \partial\Omega. \tag{5.15}$$

We thus evaluate the expression of (5.13) on $\partial\Omega$

$$\begin{aligned}\mathrm{D}^2\sigma\mathrm{D}\sigma\cdot\boldsymbol{n} &= \sum_{i,j}\partial^2_{ij}\sigma\,\partial_i\sigma\,\partial_j d_\Omega = \sum_{i,j}\Big(\partial_i\big(\partial_j\sigma\,\partial_j d_\Omega\big)\partial_i\sigma - \partial^2_{ij}d_\Omega\,\partial_i\sigma\,\partial_j\sigma\Big)\\ &\le \sum_{i,j}\partial_i\big(\partial_j\sigma\,\partial_j d_\Omega\big)\partial_i\sigma = \mathrm{D}\big(\mathrm{D}\sigma\cdot\mathrm{D}d_\Omega\big)\cdot\mathrm{D}\sigma,\end{aligned} \tag{5.16}$$

being $\mathrm{D}^2 d_\Omega$ a positive semi-definite symmetric matrix.

Observe now that $\mathrm{D}\sigma$ is a tangential vector field on $\partial\Omega$; thus for the C^1 function $\zeta := \mathrm{D}\sigma\cdot\mathrm{D}d_\Omega$ the restriction of $\mathrm{D}\zeta\cdot\mathrm{D}\sigma$ on $\partial\Omega$ depends only on the tangential gradient of ζ on $\partial\Omega$, i.e. on the restriction of ζ on $\partial\Omega$. Being $\zeta = \mathrm{D}\sigma\cdot\boldsymbol{n} = 0$ on $\partial\Omega$, we conclude that

$$\mathrm{D}\big(\mathrm{D}\sigma\cdot\mathrm{D}d_\Omega\big)\cdot\mathrm{D}\sigma = 0 \quad \text{on } \partial\Omega,$$

thus obtaining (5.13).

In order to prove (5.14), by standard extension and regularization techniques we can find a sequence $\sigma_n \in C^\infty_c(\mathbb{R}^d)$ preserving the boundary condition $\partial_{\boldsymbol n}\sigma_n = 0$ on $\partial\Omega$ such that $\sigma_n \to \sigma$ in $W^{2,2}(\Omega)$ as $n \to +\infty$; thus it is not restrictive to assume $\sigma \in C^3(\overline{\Omega})$. We recall the commutator identity

$$\partial_i\tilde\partial_j\sigma = \partial_i\big(\partial_j\sigma - \sigma\partial_j V\big) = \partial^2_{ij}\sigma - \partial_i\sigma\partial_j V - \sigma\partial^2_{ij}V = \tilde\partial_j\partial_i\sigma - \sigma\partial^2_{ij}V, \tag{5.17}$$

and we therefore obtain

$$\begin{aligned}
\int_\Omega \big|\Delta_\gamma\sigma\big|^2\,\mathrm{d}\gamma &= \sum_{i,j}\int_\Omega \tilde\partial_i\partial_i\sigma\tilde\partial_j\partial_j\sigma\,\mathrm{d}\gamma \overset{(1.11)}{=} -\sum_{i,j}\int_\Omega \partial_i\sigma\,\partial_i\tilde\partial_j\partial_j\sigma\,\mathrm{d}\gamma + \int_{\partial\Omega}\tilde\partial_j\partial_j\sigma\,\partial_i\sigma\,\boldsymbol n_i\,e^{-V}\,\mathrm{d}\mathcal{H}^{d-1}\\
&= -\sum_{i,j}\int_\Omega \partial_i\sigma\,\partial_i\tilde\partial_j\partial_j\sigma\,\mathrm{d}\gamma \overset{(5.17)}{=} \sum_{i,j}\int_\Omega \partial_i\sigma\Big(-\tilde\partial_j\partial^2_{ij}\sigma + \partial^2_{ij}V\partial_j\sigma\Big)\,\mathrm{d}\gamma\\
&\overset{(1.11)}{=} \sum_{i,j}\int_\Omega\Big(\big(\partial^2_{ij}\sigma\big)^2 + \partial^2_{ij}V\,\partial_i\sigma\,\partial_j\sigma\Big)\,\mathrm{d}\gamma - \sum_{i,j}\int_{\partial\Omega}\partial^2_{ij}\sigma\,\partial_i\sigma\,\boldsymbol n_j\,e^{-V}\,\mathrm{d}\mathcal{H}^{d-1}\\
&\overset{(5.13)}{\geq} \int_\Omega\Big(\big|\mathrm{D}^2\sigma\big|^2 + \mathrm{D}^2V\mathrm{D}\sigma\cdot\mathrm{D}\sigma\Big)\,\mathrm{d}\gamma. \qquad\square
\end{aligned}$$

5.2. **The case of a general convex domain.** Let now Ω be an arbitrary convex open subset of $\mathbb{R}^d$; we consider a monotonically increasing family of smooth, convex, bounded, and open sets Ω^k, $k \in \mathbb{N}$, such that (see e.g. [Gri85]) $\bigcup_{k\in\mathbb{N}}\Omega^k = \Omega$, and a decreasing vanishing sequence $\varepsilon^k > 0$ such that

$$\lim_{k\to+\infty}\varepsilon^k\log\varepsilon^k\int_{\Omega^k}\big(1+|x|^2\big)\,\mathrm{d}\gamma(x) = 0. \tag{5.18}$$

We denote by $\gamma^k = \chi_{\Omega^k}\cdot\gamma$ the restriction of γ to Ω^k and we associate to an initial datum $\nu_0 = \rho_0\,\gamma \in \mathcal{P}^r_2(\Omega)$ its normalized approximations $\nu^k_0 = \rho^k_0\gamma = \rho^k_0\,\gamma^k$ defined as

$$\rho^k_0(x) := \begin{cases}\frac{1}{a^k}\big(\varepsilon^k + \rho_0\wedge k\big) & \text{if } x\in\Omega^k,\\ 0 & \text{if } x \in \Omega\setminus\Omega^k,\end{cases}\qquad a^k := \int_{\Omega^k}\Big(\varepsilon^k + \rho_0\wedge k\Big)\,\mathrm{d}\gamma, \tag{5.19}$$

where $a\wedge b := \min\{a,b\}$. We recall in the next lemma some preliminary properties of this approximation we will need in the sequel.

Lemma 5.3 (**Inner approximation of convex sets**). *Let $\Omega,\Omega^k,\gamma,\gamma^k,\nu_0,\nu^k_0$ be defined as above. Then*

$$\nu^k_0 \to \nu_0 \quad \text{in } \mathcal{P}_2(\mathbb{R}^d),\quad \mathscr{H}(\nu^k_0|\,\gamma^k) \to \mathscr{H}(\nu_0|\,\gamma),\quad \mathscr{I}_2(\nu^k_0|\,\gamma^k) \to \mathscr{I}_2(\nu_0|\,\gamma)\qquad \text{as } k\to\infty. \tag{5.20}$$

Moreover, for every sequence $\mu^k \in \mathcal{P}_2(\mathbb{R}^d)$

$$\mu^k \rightharpoonup \mu \quad \text{in } \mathcal{P}(\mathbb{R}^d),\quad \sup_k \mathsf{m}_2(\mu^k) < +\infty \qquad\Rightarrow\qquad \liminf_{k\to+\infty}\mathscr{H}(\mu^k|\,\gamma^k) \geq \mathscr{H}(\mu|\,\gamma). \tag{5.21}$$

Proof. Since $\rho_0\cdot\gamma$ is a probability measures, let us first observe that by (5.18)

$$\lim_{k\to+\infty} a^k = 1. \tag{5.22}$$

If $\zeta \in C^0(\mathbb{R}^d)$ satisfies the quadratic growth condition $|\zeta(x)| \leq A(1+|x|^2)$ for every $x \in \mathbb{R}^d$, then

$$|\rho^k_0(x)\zeta(x)| \leq \frac{A}{a^k}(1+|x|^2)(\varepsilon^k + |\rho_0(x)|),\qquad \lim_{k\to+\infty}\frac{A}{a^k}\int_\Omega \varepsilon^k(1+|x|^2)\,\mathrm{d}\gamma(x) \overset{(5.18)}{=} 0;$$

since $\rho^k_0(x) \to \rho_0(x)$ for every $x\in\Omega$, Lebesgue Dominated Convergence Theorem yields

$$\lim_{k\to\infty}\int_{\mathbb{R}^d}\zeta(x)\,\mathrm{d}\nu^k_0(x) = \lim_{k\to\infty}\int_\Omega\zeta(x)\rho^k_0(x)\,\mathrm{d}\gamma(x) = \int_\Omega \zeta(x)\rho_0(x)\,\mathrm{d}\gamma(x) = \int_{\mathbb{R}^d}\zeta(x)\,\mathrm{d}\nu_0(x),$$

which proves the first assertion of (5.20).

In order to prove the convergence of the Entropy in (5.20), we observe that for every $r > 0$ and $0 < \varepsilon < 1$

$$(\varepsilon + r)\log(\varepsilon + r) = \varepsilon\log(\varepsilon + r) + r\log(\varepsilon + r) \geq \varepsilon\log\varepsilon + r\log r \geq \varepsilon\log\varepsilon - (r\log r)^-, \tag{5.23}$$

and

$$(\varepsilon + r)\log(\varepsilon + r) \le (\varepsilon + r)\log(2(r \vee 1)) \le \log 2(\varepsilon + r) + 2r(\log r)^+. \tag{5.24}$$

Combining (5.23) and (5.24) with $\varepsilon := \varepsilon^k$ and $r := \rho_0 \wedge k$, we get

$$\varepsilon^k \log \varepsilon^k - (\rho_0 \log \rho_0)^- \le (\varepsilon^k + \rho_0 \wedge k)\log(\varepsilon^k + \rho_0 \wedge k) \le \log 2(\varepsilon^k + \rho_0) + 2\rho_0(\log \rho_0)^+; \tag{5.25}$$

since $\varepsilon^k \log \varepsilon^k \gamma(\Omega^k) \to 0$ as $k \to \infty$ by (5.18), Lebesgue Dominated Convergence Theorem yields

$$\lim_{k\to\infty} \mathscr{H}(\nu_0^k|\,\gamma^k) \overset{(5.22)}{=} \lim_{k\to\infty} \int_{\Omega^k} \log(\varepsilon^k + \rho_0 \wedge k)(\varepsilon^k + \rho_0 \wedge k)\,\mathrm{d}\gamma = \int_\Omega \rho_0 \log \rho_0 \,\mathrm{d}\gamma = \mathscr{H}(\nu_0|\,\gamma).$$

It is easier to get the third limit in (5.20): setting $R^k := \{x \in \Omega : \rho_0 \le k\}$, Stampacchia's truncation Theorem yields

$$\mathscr{I}_2(\nu_0^k|\,\gamma^k) = \frac{1}{a^k}\int_{\Omega^k} \frac{\big|\mathrm{D}(\rho_0 \wedge k)\big|^2}{\varepsilon^k + \rho_0 \wedge k}\,\mathrm{d}\gamma = \frac{1}{a^k}\int_{\Omega^k \cap R^k} \frac{\big|\mathrm{D}\rho_0\big|^2}{\varepsilon^k + \rho_0}\,\mathrm{d}\gamma; \tag{5.26}$$

we can then apply Lebesgue Monotone Convergence Theorem and (5.22). In particular

$$\mathscr{I}_2(\nu_0^k|\,\gamma^k) \le \frac{1}{a^k}\mathscr{I}_2(\nu_0|\,\gamma). \tag{5.27}$$

Finally, (5.21) is an immediate consequence of Lemma 2.3, recalling that $\inf_{\mathbb{R}^d} V > -\infty$. □

We denote by e^k the operator trivially extending a function to 0 outside Ω^k.

Lemma 5.4 (**Inner approximation of the Fokker-Planck flow**). *Let $\nu_t = \rho_t\gamma = (\sigma_t)^2\gamma$ be the solution of* (5.3) *with respect to the initial datum $\nu_0 = \rho_0\gamma \in \mathcal{P}_2^r(\Omega)$ with $\mathscr{I}_2(\nu_0|\,\gamma) < +\infty$, and let $\nu_t^k = \rho_t^k\gamma = (\sigma_t^k)^2\gamma$ be the corresponding solutions of* (5.3) *in $\Omega^k \times (0, +\infty)$, originating from ν_0^k defined as in* (5.19) *for the approximating family Ω^k. Then for every $t > 0$*

$$\nu_t^k \to \nu_t \quad \text{in } \mathcal{P}_2(\mathbb{R}^d), \qquad \begin{aligned} \mathsf{e}^k\big[\sigma_t^k\big] \to \sigma_t, \quad \mathsf{e}^k\big[\mathrm{D}\sigma_t^k\big] \rightharpoonup \mathrm{D}\sigma_t \quad \text{in } L^2_\gamma(\Omega),\\ \mathsf{e}^k\big[\rho_t^k\big] \to \rho_t, \quad \mathsf{e}^k\big[\mathrm{D}\rho_t^k\big] \rightharpoonup \mathrm{D}\rho_t \quad \text{in } L^1_\gamma(\Omega), \end{aligned} \tag{5.28}$$

and for every bounded open set $\Omega' \subset \Omega$ and every $T > 0$

$$\mathsf{e}^k\big[\mathrm{D}^2\sigma_t^k\big] \rightharpoonup \mathrm{D}^2\sigma_t \quad \text{in } L^2(0,T;L^2_\gamma(\Omega')), \qquad \mathsf{e}^k\big[\mathrm{D}^2\rho_t^k\big] \rightharpoonup \mathrm{D}^2\rho_t \quad \text{in } L^1(0,T;L^1_\gamma(\Omega')), \tag{5.29}$$

$$\sigma \in L^2(0,T;W^{2,2}_\gamma(\Omega')), \quad \Delta_\gamma\sigma \in L^2(0,T;L^2_\gamma(\Omega')), \quad \sqrt{\sigma} \in L^4(0,T;L^4_\gamma(\Omega'), \tag{5.30}$$

with $\partial_{\boldsymbol{n}}\sigma_t = 0$ on $\partial\Omega$ for a.e. $t > 0$. Moreover, for every $h > 0$

$$\frac{1}{2}\mathscr{I}_2(\nu_h|\,\gamma) + \int_0^h \mathscr{P}(\nu_t|\,\gamma)\,\mathrm{d}t \le \frac{1}{2}\mathscr{I}_2(\nu_0|\,\gamma) \qquad \text{if } \lambda \ge 0, \tag{5.31}$$

$$\frac{e^{2\lambda h}}{2}\mathscr{I}_2(\nu_h|\,\gamma) + \int_0^h e^{2\lambda t}\Big(\mathscr{P}(\nu_t|\,\gamma) - \lambda\mathscr{I}_2(\nu_t|\,\gamma)\Big)\,\mathrm{d}t \le \frac{1}{2}\mathscr{I}_2(\nu_0|\,\gamma) \qquad \text{for every } \lambda \in \mathbb{R}. \tag{5.32}$$

where $\mathscr{P}(\mu|\,\gamma)$ denotes the "Fisher dissipation functional" introduced in (5.1a,b).

Proof. The curves $t \mapsto \nu_t^k$ are the gradient flows in $\mathcal{P}_2(\mathbb{R}^d)$ of the relative entropy functionals $\mu \mapsto \mathscr{H}(\mu|\,\gamma^k)$ with respect to the Wasserstein distance (see Theorem 2.17). Thanks to Lemma 5.3 we can apply the stability result of [AGS05, Theorem 11.2.1] which shows that

$$\nu_t^k \to \nu_t \quad \text{in } \mathcal{P}_2(\mathbb{R}^d) \quad \forall\, t \ge 0. \tag{5.33}$$

Since $t \mapsto e^{2\lambda t}\mathscr{I}_2(\nu_t^k|\,\gamma^k)$ is non-increasing (see Theorem 2.17 or (5.6)), we get the uniform estimate

$$\sup_t 4\int_{\Omega^k} |\mathrm{D}\sigma_t^k|^2\,\mathrm{d}\gamma = \sup_t \mathscr{I}_2(\nu_t^k|\,\gamma^k) \le e^{-2\lambda t}\mathscr{I}_2(\nu_0^k|\,\gamma^k) \overset{(5.27)}{\le} \frac{e^{-2\lambda t}}{a^k}\mathscr{I}_2(\nu_0|\,\gamma). \tag{5.34}$$

Since every compact subset of Ω is definitely included in Ω^k for sufficiently big $k \in \mathbb{N}$, (5.34) yields $\mathsf{e}^k\big[\sigma_t^k\big] \to \sigma_t$ in $L^2_{\gamma,\mathrm{loc}}(\Omega)$, and therefore the convergence in $L^2_\gamma(\Omega)$ as

$$\int_{\Omega^k} |\sigma_t^k|^2\,\mathrm{d}\gamma = \int_\Omega \big|\mathsf{e}^k\big[\sigma_t^k\big]\big|^2\,\mathrm{d}\gamma = \int_\Omega |\sigma_t|^2\,\mathrm{d}\gamma = 1.$$

Observe now that for every vector field $\boldsymbol{\zeta} \in C^\infty_c(\Omega)$

$$\lim_{k\uparrow+\infty} \int_\Omega \mathsf{e}^k \left[\mathrm{D}\sigma^k_t(x)\right] \cdot \boldsymbol{\zeta}(x)\,\mathrm{d}\gamma(x) = -\lim_{k\uparrow+\infty} \int_\Omega \sigma^k_t(x)\,\mathrm{div}_\gamma \boldsymbol{\zeta}(x)\,\mathrm{d}\gamma(x) = -\int_\Omega \sigma_t(x)\,\mathrm{div}_\gamma \boldsymbol{\zeta}(x)\,\mathrm{d}\gamma(x)$$

so that

$$-\int_\Omega \sigma_t(x)\,\mathrm{div}_\gamma \boldsymbol{\zeta}(x)\,\mathrm{d}\gamma(x) \le \|\boldsymbol{\zeta}\|_{L^2_\gamma(\Omega;\mathbb{R}^d)} \liminf_{k\uparrow+\infty} \Big(\int_\Omega \left|\mathsf{e}^k\left[\mathrm{D}\sigma^k_t\right](x)\right|^2 \mathrm{d}\gamma(x)\Big)^{1/2},$$

and

$$\sigma_t \in W^{1,2}_\gamma(\Omega), \quad \int_\Omega \left|\mathrm{D}\sigma_t(x)\right|^2 \mathrm{d}\gamma \le \liminf_{k\to+\infty} \int_\Omega \left|\mathsf{e}^k\left[\mathrm{D}\sigma^k_t\right](x)\right|^2 \mathrm{d}\gamma.$$

In order to prove (5.29) let us observe that by Corollary 3.7 for every bounded open set $\Omega' \subset \Omega$ and $k \in \mathbb{N}$

$$\begin{aligned}\int_0^T \int_{\Omega^k\cap\Omega'} \Big(\left|\mathrm{D}^2\sigma^k_t(x)\right|^2 + \left|\Delta_\gamma \sigma^k\right|^2 + \left|\mathrm{D}\sqrt{\sigma^k}\right|^4\Big)\,\mathrm{d}\gamma\,\mathrm{d}t &\overset{(3.37)}{\le} c_{-1} \int_0^T \Big(\mathscr{K}_{-1}(\nu^k_t|\,\gamma^k) + \mathscr{I}_2(\nu^k|\,\gamma^k)\Big)\,\mathrm{d}t \\ &\overset{(5.6)}{\le} c_{-1}(\tfrac12 + T)e^{-2\lambda^- T}\mathscr{I}_2(\nu^k_0|\,\gamma^k) \overset{(5.27)}{\le} c_{-1}(\tfrac12+T)\frac{e^{-2\lambda^- T}}{a^k}\mathscr{I}_2(\nu_0|\,\gamma);\end{aligned}$$

(5.22), Fatou's Lemma, and the previous weak convergence result yields (5.29) and (5.30).

Let us now fix $t > 0$ such that $\sigma_t \in W^{2,2}_\gamma(\Omega)$ and, for a suitable subsequence still labeled σ^k_t, $\mathsf{e}^k\left[\mathrm{D}^2\sigma^k_t\right] \rightharpoonup \mathrm{D}^2\sigma_t$ in $L^2_\gamma(\Omega\cap B_h(0))$ for every $h \in \mathbb{N}$. In order to check that $\partial_{\boldsymbol{n}}\sigma_t = 0$ on $\partial\Omega$ we must show

$$-\int_\Omega \Delta_\gamma \sigma_t \zeta\,\mathrm{d}\gamma = \int_\Omega \mathrm{D}\sigma_t\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma \quad \forall\,\zeta \in C^1_c(\mathbb{R}^d). \tag{5.35}$$

Choosing $h \in \mathbb{N}$ such that $\operatorname{supp}\zeta \subset\subset B_h(0)$, since $\partial_{\boldsymbol{n}}\sigma^k_t = 0$ on $\partial\Omega^k \cap B_h(0)$, we easily obtain

$$\begin{aligned}-\int_\Omega \Delta_\gamma\sigma_t\zeta\,\mathrm{d}\gamma &= -\int_{\Omega\cap B_h(0)} \Delta_\gamma\sigma_t\zeta\,\mathrm{d}\gamma = -\lim_{k\uparrow+\infty}\int_{\Omega^k\cap B_h(0)} \Delta\sigma^k_t\,\zeta\,\mathrm{d}\gamma = \lim_{k\uparrow+\infty}\int_{\Omega^k\cap B_h(0)} \mathrm{D}\sigma^k_t\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma \\ &= \int_{\Omega\cap B_h(0)} \mathrm{D}\sigma_t\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma = \int_\Omega \mathrm{D}\sigma_t\cdot\mathrm{D}\zeta\,\mathrm{d}\gamma.\end{aligned}$$

Finally, (5.31) (resp. (5.32)) follows from (5.5) (resp. (5.6)) by Fatou's Lemma and the lower semicontinuity of the Fisher information (Lemma 2.4) and of $\mathscr{K}_{-1}(\cdot|\,\gamma)$ (Corollary 3.8). Observe that when λ is negative $\mathscr{I}^V_2(\cdot|\,\gamma)$ is not convex but the integrand in (5.32) is still l.s.c. by (2.54). $\square$

5.3. **Estimates and characterization of the limiting subdifferential of $\mathscr{I}_2(\cdot|\,\gamma)$.** The next result provides the main estimate of the strong subdifferential of the Relative Fisher information

$$\mathscr{G}(\mu) = \frac12 \mathscr{I}_2(\mu|\,\gamma), \quad \mu \in \mathcal{P}^r_2(\Omega). \tag{5.36}$$

Recall that the "Fisher dissipation functional" $\mathscr{P}(\mu|\,\gamma)$ is defined by

$$\mathscr{P}(\mu|\,\gamma) = \mathscr{K}_{-1}(\mu|\,\gamma) + \mathscr{I}^V_2(\mu|\,\gamma), \quad \mathscr{P}(\mu|\,\mathcal{L}^d) = \mathscr{K}_{-1}(\mu|\,\mathcal{L}^d) \quad \forall\,\mu\in\mathcal{P}^r_2(\Omega). \tag{5.37}$$

Theorem 5.5 (**A priori estimates on the strong subdifferential**). *Let us suppose that $\mu = v\gamma = s^2\gamma \in D(\partial_s\mathscr{F})$ and $\boldsymbol{\xi} \in \partial_s\mathscr{G}(\mu)$ according to Definition 2.10 and Theorem 4.2. Then*

$$\mathscr{P}(\mu|\,\gamma) \le \int_\Omega \boldsymbol{\xi}(x)\cdot\frac{\mathrm{D}v(x)}{v(x)}\,\mathrm{d}\mu(x) = \int_\Omega \boldsymbol{\xi}(x)\cdot\mathrm{D}v(x)\,\mathrm{d}\gamma(x) \le \|\boldsymbol{\xi}\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\Big(\mathscr{I}_2(\mu|\,\gamma)\Big)^{1/2}. \tag{5.38}$$

In particular, for every bounded *open subset $\Omega' \subset \Omega$*

$$s \in W^{2,2}_\gamma(\Omega'), \quad \sqrt{s} \in W^{1,4}_\gamma(\Omega'), \quad \partial_{\boldsymbol{n}} s = 0 \quad \textit{on } \partial\Omega, \tag{5.39}$$

and there exists a constant $C_{\Omega'} > 0$ such that

$$\int_{\Omega'} \Big(|\mathrm{D}^2 s|^2 + |\Delta_\gamma s|^2 + |\mathrm{D}\sqrt{s}|^4\Big)\,\mathrm{d}\gamma \le C'_\Omega\Big[\|\boldsymbol{\xi}\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\Big(\mathscr{I}_2(\mu|\,\gamma)\Big)^{1/2} + \mathscr{I}_2(\mu|\,\gamma)\Big]. \tag{5.40}$$

Finally, if $\lambda > 0$,

$$\lambda^2 \mathscr{I}_2(\mu|\,\gamma) \le \lambda \mathscr{I}_2^V(\mu|\,\gamma) \le \int_\Omega |\boldsymbol{\xi}|^2 \,\mathrm{d}\mu. \tag{5.41}$$

Proof. For the sake of simplicity we assume here $\lambda \ge 0$ (the case $\lambda < 0$ is completely analogous, by invoking (5.32) instead of (5.31)) and we argue as in the proof of Theorem 4.2 choosing now as flowing curve the solution ν_t of the Fokker Planck equation with Neumann boundary conditions (2.123) starting from $\nu_0 := \mu$. Setting $\boldsymbol{t}_h = T_o(\mu, \nu_h)$, by the very definition (2.68) of strong subdifferential and (5.31), we know that

$$\int_0^h \mathscr{P}(\nu_t|\,\gamma)\,\mathrm{d}t \overset{(5.31)}{\le} \mathscr{G}(\mu) - \mathscr{G}(\nu_h) \overset{(2.68)}{\le} \int_\Omega \boldsymbol{\xi}(x)\cdot\big(x - \boldsymbol{t}_h(x)\big)\,\mathrm{d}\mu(x) + o\big(\|\boldsymbol{t}_h - \boldsymbol{i}\|_{L^2_\mu(\Omega)}\big). \tag{5.42}$$

Since

$$\nu_h \to \nu_0 = \mu \text{ in } \mathcal{P}_2(\Omega), \qquad \frac{\boldsymbol{i} - \boldsymbol{t}_h}{h} \to \frac{\mathrm{D}v}{v} \text{ in } L^2_\mu(\Omega;\mathbb{R}^d), \qquad \mathscr{I}_2(\nu_h|\,\gamma) \to \mathscr{I}_2(\nu|\,\gamma) \quad \text{as } h\downarrow 0, \tag{5.43}$$

Fatou's Lemma and the lower semicontinuity properties stated in (2.54) and in Corollary 3.8 yield

$$\mathscr{P}(\mu|\,\gamma) \le \liminf_{h\downarrow 0} \frac{1}{h}\int_0^h \mathscr{P}(\nu_t|\,\gamma)\,\mathrm{d}t \le \int_\Omega \boldsymbol{\xi}(x)\cdot\frac{\mathrm{D}v(x)}{v(x)}\,\mathrm{d}\mu(x),$$

which concludes the proof of (5.38). Since $\mathscr{P}(\mu|\,\gamma) \ge \mathscr{K}_{-1}(\mu|\,\gamma)$, Lemma 3.7 yields (5.39) and (5.40); (5.41) follows from (5.38) and

$$\lambda\mathscr{I}_2(\mu|\,\gamma) \le \mathscr{I}_2^V(\mu|\,\gamma) \le \mathscr{P}(\mu|\,\gamma).$$

□

Thanks to the previous Theorem 5.5 we can now study the closure properties of the Wasserstein subdifferential of $\mathscr{G}$.

Lemma 5.6 (**Closure properties for** $\partial\mathscr{G}$)**.** *Let* $\mu_k = s_k^2\gamma \in D(\partial_s\mathscr{G})$, $\mu = s^2\gamma \in D(\mathscr{G})$, *and* $\boldsymbol{\xi}_k \in \partial_s\mathscr{G}(\mu_k)$ *be satisfying* (2.69), *i.e.*

$$\mu_k \rightharpoonup \mu \quad \text{in } \mathcal{P}(\Omega), \quad \sup_k \Big(\mathscr{G}(\mu_k) + \mathsf{m}_2(\mu_k) + \int_\Omega |\boldsymbol{\xi}_k(x)|^2\,\mathrm{d}\mu_k(x)\Big) \le S < +\infty. \tag{5.44}$$

Then for every bounded set $\Omega' \subset \Omega$

$$s_k \to s \quad \text{strongly in } W^{1,2}_\gamma(\Omega'), \qquad \sqrt{s_k} \rightharpoonup \sqrt{s} \quad \text{weakly in } W^{1,4}_\gamma(\Omega'), \tag{5.45}$$

$$\Delta_\gamma s_k \rightharpoonup \Delta_\gamma s, \quad \partial^2_{ij} s_k \rightharpoonup \partial^2_{ij} s \quad \text{weakly in } L^2_\gamma(\Omega'), \tag{5.46}$$

and $\boldsymbol{\nu}_k = \boldsymbol{\xi}_k\mu_k = -\tilde{\boldsymbol{q}}_k\gamma \rightharpoonup \boldsymbol{\nu} = \boldsymbol{\xi}\mu = -\tilde{\boldsymbol{q}}\gamma$ *in* $\big[\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)\big]^d$, *where* $\boldsymbol{\xi}$ *and* $\tilde{\boldsymbol{q}} = -v\boldsymbol{\xi}$ *are characterized by*

$$\begin{gathered}-\int_\Omega \tilde{\boldsymbol{q}}\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma = \int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\int_\Omega \Big(4\mathrm{D}\boldsymbol{\zeta}\,\mathrm{D}s\cdot\mathrm{D}s + 2s\mathrm{D}\big(\operatorname{div}_\gamma\boldsymbol{\zeta}\big)\cdot\mathrm{D}s\Big)\,\mathrm{d}\gamma \\ \forall\,\boldsymbol{\zeta} \in C^\infty_{\mathrm{c}}(\mathbb{R}^d;\mathbb{R}^d),\ \boldsymbol{\zeta}\cdot\boldsymbol{n} = 0 \text{ on } \partial\Omega.\end{gathered} \tag{5.47}$$

Moreover,

$$\text{if} \quad \liminf_{k\to\infty} \mathscr{I}_2(\mu_k|\,\gamma) = \mathscr{I}_2(\mu|\,\gamma) \quad \text{then} \quad \mu, \boldsymbol{\xi} \text{ satisfy (5.38)}. \tag{5.48}$$

Finally, when Ω *is* bounded *or, more generally, when* V *satisfies a* global Lipschitz condition *on* Ω, *then the functional* $\mathscr{G} = \frac{1}{2}\mathscr{I}_2(\cdot|\,\gamma)$ *is* regular, *according to Definition 2.10, i.e.*

$$\lim_{k\to\infty} \mathscr{I}_2(\mu_k|\,\gamma) = \mathscr{I}_2(\mu|\,\gamma), \tag{5.49}$$

Proof. Since $\boldsymbol{\xi}_k \in \partial_s\mathscr{G}(\mu_k)$, Theorem 4.2 yields

$$-\int_\Omega \boldsymbol{q}_k\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma = -\int_\Omega \Big(4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}s_k\cdot\mathrm{D}s_k + 2s_k\mathrm{D}\big(\operatorname{div}_\gamma\boldsymbol{\zeta}\big)\cdot\mathrm{D}s_k\Big)\,\mathrm{d}\gamma \tag{5.50}$$

for every vector field $\boldsymbol{\zeta} \in C^2_{\mathrm{c}}(\mathbb{R}^d;\mathbb{R}^d)$ with $\boldsymbol{\zeta}\cdot\boldsymbol{n} = 0$ on $\partial\Omega$. Thanks to (5.44) and the a priori estimate (5.40), it is immediate to get (5.45), (5.46), and to obtain (5.47) by passing to the limit in (5.50) as $k \to +\infty$. In particular the limit vector field $\boldsymbol{\xi}$ (which belongs to $L^2_\mu(\Omega;\mathbb{R}^d)$ thanks to (5.44) and Proposition 2.2 is uniquely determined.

When Ω is bounded (5.45) immediately yields (5.49). In the unbounded case, being V Lipschitz continuous, we notice that (5.45) and (5.46) hold even for $\Omega' = \Omega$, since we can apply (3.32) with $\beta = -1$. Since

$$\int_\Omega |\mathrm{D}s_k|^2 \,\mathrm{d}\gamma = -\int_\Omega s_k \Delta_\gamma s_k \,\mathrm{d}\gamma, \tag{5.51}$$

strong converge of s_k to s and weak convergence of $\Delta_\gamma s_k$ to $\Delta_\gamma s$ in $L^2_\gamma(\Omega)$ are enough to pass to the limit in the Dirichlet integral, obtaining

$$\begin{gathered}\lim_{k\uparrow+\infty} \mathscr{I}_2(\mu_k|\,\gamma) = \lim_{k\uparrow+\infty} 4\int_\Omega |\mathrm{D}s_k|^2 \,\mathrm{d}\gamma = 4\int_\Omega |\mathrm{D}s|^2 \,\mathrm{d}\gamma = \mathscr{I}_2(\mu|\,\gamma),\\ \mathrm{D}s_k \to \mathrm{D}s \quad \text{strongly in } L^2_\gamma(\Omega;\mathbb{R}^d).\end{gathered} \tag{5.52}$$

In order to prove (5.48), let us observe that the convergence of the relative Fisher information (up to the extraction of a suitable subsequence, still denoted by μ_k) and (2.54) yield

$$\liminf_{k\to\infty} \mathscr{I}_2^V(\mu_k|\,\gamma) \ge \mathscr{I}_2^V(\mu|\,\gamma), \quad \liminf_{k\to\infty} \mathscr{P}(\mu_k|\,\gamma) \overset{(3.39)}{\ge} \mathscr{P}(\mu|\,\gamma). \tag{5.53}$$

It is then sufficient to prove that

$$\int_\Omega \boldsymbol{\xi}_k \cdot \mathrm{D}v_k \,\mathrm{d}\gamma \to \int_\Omega \boldsymbol{\xi} \cdot \mathrm{D}v \,\mathrm{d}\gamma. \tag{5.54}$$

(5.54) follows by the next Lemma, where $\boldsymbol{\eta}_k = \frac{\mathrm{D}v_k}{v_k}$ denotes the logarithmic gradient of μ_k. □

The next technical lemma extends the classical convergence result for the L^2-scalar product of two sequences $\boldsymbol{\xi}_k, \boldsymbol{\eta}_k$ under weak-strong convergence of the factors. Here we consider the case when the underlying measures μ_k still depend on the index k and narrowly converge to μ as $k \to \infty$. The proof of this result relies on the theory (and the notation) developed in [AGS05, Chap. 5], to which we refer for more details.

Lemma 5.7. *Let $\mu_k \in \mathcal{P}_2(\Omega)$ be a sequence of measures narrowly converging to $\mu \in \mathcal{P}_2(\Omega)$, and let $\boldsymbol{\xi}_k, \boldsymbol{\eta}_k \in L^2_{\mu_k}(\Omega;\mathbb{R}^d)$, $\boldsymbol{\xi}, \boldsymbol{\eta} \in L^2_\mu(\Omega;\mathbb{R}^d)$, satisfying*

$$\lim_{k\to\infty} \int_\Omega \boldsymbol{\xi}_k \cdot \boldsymbol{\zeta} \,\mathrm{d}\mu_k \to \int_\Omega \boldsymbol{\xi} \cdot \boldsymbol{\zeta} \,\mathrm{d}\mu, \quad \lim_{k\to\infty} \int_\Omega \boldsymbol{\eta}_k \cdot \boldsymbol{\zeta} \,\mathrm{d}\mu_k \to \int_\Omega \boldsymbol{\eta} \cdot \boldsymbol{\zeta} \,\mathrm{d}\mu \tag{5.55}$$

for every bounded and continuous function $\boldsymbol{\zeta} \in C_b(\Omega;\mathbb{R}^d)$. If

$$\limsup_{k\to\infty} \int_\Omega |\boldsymbol{\xi}_k|^2 \,\mathrm{d}\mu_k < +\infty, \qquad \limsup_{k\to\infty} \int_\Omega |\boldsymbol{\eta}_k|^2 \,\mathrm{d}\mu_k \le \int_\Omega |\boldsymbol{\eta}|^2 \,\mathrm{d}\mu, \tag{5.56}$$

then

$$\lim_{k\to\infty} \int_\Omega \boldsymbol{\xi}_k \cdot \boldsymbol{\eta}_k \,\mathrm{d}\mu_k = \int_\Omega \boldsymbol{\xi} \cdot \boldsymbol{\eta} \,\mathrm{d}\mu. \tag{5.57}$$

Proof. Let us introduce the measures $\boldsymbol{\mu}_k := (\boldsymbol{i}, \boldsymbol{\xi}_k, \boldsymbol{\zeta}_k)_{\#}\mu_k \in \mathcal{P}_2(\Omega \times \mathbb{R}^d \times \mathbb{R}^d)$, which satisfy

$$\int_\Omega \boldsymbol{\xi}_k \cdot \boldsymbol{\zeta}_k \,\mathrm{d}\mu_k = \int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d} x_2 \cdot x_3 \,\mathrm{d}\boldsymbol{\mu}_k(x_1, x_2, x_3). \tag{5.58}$$

Since the marginals of the sequence $\boldsymbol{\mu}_k$ are narrowly relatively compact, $(\boldsymbol{\mu}_k)$ is narrowly relatively compact in $\mathcal{P}(\Omega\times\mathbb{R}^d\times\mathbb{R}^d)$ [AGS05, Lemma 5.2.2] and we can assume that a suitable subsequence, still denoted by $\boldsymbol{\mu}_k$, narrowly converges to $\boldsymbol{\mu} \in \mathcal{P}_2(\Omega \times \mathbb{R}^d \times \mathbb{R}^d)$. We denote by π^j the projection map on the j-th coordinate in $\Omega \times \mathbb{R}^d \times \mathbb{R}^d$, $\pi^j(x_1, x_2, x_3) := x_j$, and by $\boldsymbol{\mu}^j = (\pi^j)_{\#}\boldsymbol{\mu}$ the j-th marginal of $\boldsymbol{\mu}$; observe that $\boldsymbol{\mu}^1 = \mu$.

[AGS05, Thm. 5.4.4] and (5.55) yield

$$\begin{aligned}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}\boldsymbol{\zeta}(x_1)\cdot x_2\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2,x_3) &= \int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}\boldsymbol{\zeta}(x_1)\cdot\boldsymbol{\xi}(x_1)\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2,x_3)\\ &= \int_{\mathbb{R}^d}\boldsymbol{\zeta}(x)\cdot\boldsymbol{\xi}(x)\,\mathrm{d}\boldsymbol{\mu}^1(x),\end{aligned}\tag{5.59}$$

for every vector field $\boldsymbol{\zeta}\in L^2_{\boldsymbol{\mu}^1}(\Omega;\mathbb{R}^d)$.

Moreover, the marginals of $\boldsymbol{\mu}_k$ also satisfy

$$\boldsymbol{\mu}_k^{1,3} = (\boldsymbol{i},\boldsymbol{\zeta}_k)_{\#}\mu_k = (\pi^1,\pi^3)_{\#}\boldsymbol{\mu}_k \rightharpoonup \boldsymbol{\mu}^{1,3} = (\boldsymbol{i},\boldsymbol{\zeta})_{\#}\mu = (\pi^1,\pi^3)_{\#}\boldsymbol{\mu}\quad\text{in }\mathcal{P}(\omega\times\mathbb{R}^d),\tag{5.60}$$

thanks to (5.56) and [AGS05, Thm. 5.4.4]; in particular [AGS05, Lemma 5.3.2], for every function $\theta\in L^1_{\boldsymbol{\mu}}(\Omega\times\mathbb{R}^d\times\mathbb{R}^d)$ we have

$$\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}\theta(x_1,x_2,x_3)\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2,x_3) = \int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}\theta(x_1,x_2,\boldsymbol{\zeta}(x_1))\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2,x_3).\tag{5.61}$$

Applying Lemma 5.2.4 of [AGS05] we obtain that

$$\lim_{k\to\infty}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}x_2\cdot x_3\,\mathrm{d}\boldsymbol{\mu}_k(x_1,x_2,x_3) = \int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}x_2\cdot x_3\,\mathrm{d}\boldsymbol{\mu}(x_1,x_2,x_3)\tag{5.62}$$

and therefore

$$\begin{aligned}\lim_{k\to\infty}\int_\Omega\boldsymbol{\xi}_k\cdot\boldsymbol{\eta}_k\,\mathrm{d}\mu_k &\overset{(5.58)}{=}\lim_{k\to\infty}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}x_2\cdot x_3\,\mathrm{d}\boldsymbol{\mu}_k \overset{(5.62)}{=}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}x_2\cdot x_3\,\mathrm{d}\boldsymbol{\mu}\\ &\overset{(5.61)}{=}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}x_2\cdot\boldsymbol{\zeta}(x_1)\,\mathrm{d}\boldsymbol{\mu}\overset{(5.59)}{=}\int_{\mathbb{R}^d\times\mathbb{R}^d}\boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu.\end{aligned}$$

□

Combining the previous two results in the case $V\equiv 0$ and the smooth perturbation argument of §2.5 we easily get

Corollary 5.8 (Closure properties for $\partial\mathscr{F}^f$). *Let $\mathscr{F}^f(\cdot) := \frac{1}{2}\mathscr{I}_2(\cdot|\,\mathcal{L}^d) + \langle f,\cdot\rangle$ for a given function f satisfying (1.4), let $\mu_k = r_k^2\mathcal{L}^d\in D(\partial_s\mathscr{F}^f)$, $\mu = r^2\mathcal{L}^d\in D(\mathscr{F}^f)$, and $\boldsymbol{\xi}_k\in\partial_s\mathscr{F}^f(\mu_k)$ be satisfying (2.69), i.e.*

$$\mu_k\rightharpoonup\mu\quad\textit{in }\mathcal{P}(\Omega),\quad\sup_k\Big(\mathscr{I}_2(\mu_k|\,\mathcal{L}^d) + \mathsf{m}_2(\mu_k) + \int_\Omega|\boldsymbol{\xi}_k(x)|^2\,\mathrm{d}\mu_k(x)\Big)\le S<+\infty.\tag{5.63}$$

Then $r_k, r\in W^{2,2}(\Omega)$, $\sqrt{r}_k,\sqrt{r}\in W^{1,4}(\Omega)$ with $\partial_{\boldsymbol{n}}r_k = \partial_{\boldsymbol{n}}r = 0$ on $\partial\Omega$,

$$\lim_{k\to\infty}\mathscr{I}_2(\mu_k|\,\mathcal{L}^d) = \mathscr{I}_2(\mu|\,\mathcal{L}^d),\tag{5.64}$$

$$r_k\to r\quad\textit{strongly in }W^{1,2}(\Omega),\qquad\sqrt{r}_k\rightharpoonup\sqrt{r}\quad\textit{weakly in }W^{1,4}(\Omega),\tag{5.65}$$

$$\Delta r_k\rightharpoonup\Delta r,\quad\partial^2_{ij}r_k\rightharpoonup\partial^2_{ij}r\quad\textit{weakly in }L^2(\Omega),\tag{5.66}$$

and $\boldsymbol{\nu}_k = \boldsymbol{\xi}_k\mu_k = -\boldsymbol{q}_k\mathcal{L}^d\rightharpoonup\boldsymbol{\nu} = \boldsymbol{\xi}\mu = -\boldsymbol{q}\mathcal{L}^d$ in $\big[\mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d)\big]^d$, where $\boldsymbol{q},\boldsymbol{\xi}$ are characterized by

$$\begin{aligned}-\int_\Omega\boldsymbol{q}\cdot\boldsymbol{\zeta}\,\mathrm{d}x = \int_\Omega\boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\int_\Omega\Big(4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}r\cdot\mathrm{D}r + 2r\mathrm{D}\big(\operatorname{div}\boldsymbol{\zeta}\big)\cdot\mathrm{D}r - r^2\mathrm{D}f\cdot\boldsymbol{\zeta}\Big)\,\mathrm{d}\gamma\\ \forall\,\boldsymbol{\zeta}\in C^\infty_{\mathrm{c}}(\mathbb{R}^d;\mathbb{R}^d),\ \boldsymbol{\zeta}\cdot\boldsymbol{n} = 0\ \textit{on }\partial\Omega.\end{aligned}\tag{5.67}$$

Finally, $\mu, u = r^2$, and $\boldsymbol{\xi}$ satisfy

$$\begin{aligned}\frac{12}{2+d}\int_\Omega\big|\mathrm{D}^2r\big|^2\,\mathrm{d}x + \frac{64}{(2+d)}\int_\Omega\big|\mathrm{D}\sqrt{r}\big|^4\,\mathrm{d}x\le\mathscr{K}_{-1}(\mu|\,\mathcal{L}^d)\le\int_\Omega(\boldsymbol{\xi}-\mathrm{D}f)\cdot\mathrm{D}u\,\mathrm{d}x,\\ \le\|\boldsymbol{\xi}-\mathrm{D}f\|_{L^2_\mu(\Omega;\mathbb{R}^d)}\Big(\mathscr{I}_2(\mu|\,\mathcal{L}^d)\Big)^{1/2}\end{aligned}\tag{5.68}$$

$$\Big(\mathscr{I}_2(\mu|\,\mathcal{L}^d)\Big)^2\le 2(2+d)\mathscr{K}_{-1}(\mu|\,\mathcal{L}^d).\tag{5.69}$$

Proof. (5.68) follows from 3.6 and Lemma 5.2. Let us check (5.69). Since $r \in W^{2,2}(\Omega)$ with $\|r\|_{L^2(\Omega)} = 1$ and $\partial_{\boldsymbol{n}} r = 0$ on Ω, we get

$$\mathscr{I}_2(\mu|\,\mathcal{L}^d) = 4\int_\Omega \big|\mathrm{D}r\big|^2 = -4\int_\Omega r\Delta r\,\mathrm{d}x \le 4\|\Delta r\|_{L^2(\Omega)}, \tag{5.70}$$

$$\Big(\mathscr{I}_2(\mu|\,\mathcal{L}^d)\Big)^2 \overset{(2.39)}{\le} \mathscr{I}_4(\mu|\,\mathcal{L}^d) \overset{(2.40)}{=} 4^4\int_\Omega \big|\mathrm{D}\sqrt{r}\big|^4\,\mathrm{d}x. \tag{5.71}$$

A convex combination of the previous inequalities yields

$$\Big(\mathscr{I}_2(\mu|\,\mathcal{L}^d)\Big)^2 \le 2\int_\Omega \Big(4\big|\Delta r\big|^2 + 64\big|\mathrm{D}\sqrt{r}\big|^4\Big)\,\mathrm{d}x \overset{(3.35)}{\le} 2(2+d)\mathscr{K}_{-1}(\mu|\,\mathcal{L}^d). \qquad \square$$

Lemma 5.6 and some computations allow to obtain various differential expressions for the *limiting subdifferential* of $\mathscr{G} = \frac{1}{2}\mathscr{I}_2(\cdot|\,\gamma)$; besides the identity (5.47) attached to $\mathrm{EE}_{3,\gamma}$, we also obtain the differential characterization (1.57b) we adopted for the Definition 1.5 of weak solutions of $(\mathrm{EE}_{1,\gamma})$.

Theorem 5.9 (**Limiting subdifferential of $\mathscr{G}$**). *The functional $\mathscr{G} := \frac{1}{2}\mathscr{I}_2(\cdot|\,\gamma)$ has a regular subdifferential, according to Definition 2.10. If $\mu = v\gamma = s^2\gamma \in D(\partial_\ell\mathscr{G})$ then s satisfies the regularity properties* (5.39) *and $\boldsymbol{\xi} = \partial_\ell\mathscr{G}(\mu)$ is characterized by*

$$\int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\int_\Omega \Big(4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}s\cdot\mathrm{D}s + 2s\mathrm{D}\big(\operatorname{div}_\gamma\boldsymbol{\zeta}\big)\cdot\mathrm{D}s\Big)\,\mathrm{d}\gamma \quad \forall\,\boldsymbol{\zeta}\in C^2_{\mathrm{c}}(\mathbb{R}^d;\mathbb{R}^d),\ \ \boldsymbol{\zeta}\cdot\boldsymbol{n} = 0, \tag{5.72}$$

$$= \sum_{i,j}\int_\Omega \Big(\mathsf{L}_{ij}\partial_i\boldsymbol{\zeta}_j + \partial^2_{ij}V\,\partial_i v\,\boldsymbol{\zeta}_j\Big)\,\mathrm{d}\gamma \quad \forall\,\boldsymbol{\zeta}\in C^2_{\mathrm{c}}(\Omega;\mathbb{R}^d) \tag{5.73}$$

$$= -\int_\Omega \Big(2\mathrm{D}(s\Delta_\gamma s) - 4\Delta_\gamma s\mathrm{D}s\Big)\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma \quad \forall\,\boldsymbol{\zeta}\in C^2_{\mathrm{c}}(\Omega;\mathbb{R}^d), \tag{5.74}$$

where $\mathsf{L}_{ij} = 2\big(s\partial^2_{ij}s - \partial_i s\partial_j s\big)$; in particular

$$s\Delta_\gamma s \in W^{1,1}_{\gamma,\mathrm{loc}}(\Omega) \quad \textit{and} \quad \tilde{\boldsymbol{q}} = -v\boldsymbol{\xi} = 2\mathrm{D}\big(s\Delta_\gamma s\big) - 4\Delta_\gamma s\mathrm{D}s. \tag{5.75}$$

Proof. (5.72) is an immediate consequence of the previous Lemma 5.6, thus we simply have to prove (5.73) and (5.74).

In order to prove (5.73) we split (5.72) in two integrals

$$\int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\int_\Omega 4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}s\cdot\mathrm{D}s\,\mathrm{d}\gamma - \int_\Omega 2s\mathrm{D}\big(\operatorname{div}_\gamma\boldsymbol{\zeta}\big)\cdot\mathrm{D}s\,\mathrm{d}\gamma = A + B, \tag{5.76}$$

and we integrate by parts the second term B recalling the identity (5.17) $\partial_i\tilde{\partial}_j\zeta = \tilde{\partial}_j\partial_i\zeta - \zeta\partial^2_{ij}V$
Observing that

$$B = -\int_\Omega 2s\mathrm{D}\big(\operatorname{div}_\gamma\boldsymbol{\zeta}\big)\cdot\mathrm{D}s\,\mathrm{d}\gamma = -\sum_{i,j}\int_\Omega 2s\partial_i\tilde{\partial}_j\boldsymbol{\zeta}_j\,\partial_i s\,\mathrm{d}\gamma \overset{(5.17)}{=} \sum_{i,j}\int_\Omega \Big(-2s\tilde{\partial}_j\partial_i\boldsymbol{\zeta}_j\,\partial_i s + 2s\partial^2_{ij}V\,\partial_i s\,\boldsymbol{\zeta}_j\Big)\,\mathrm{d}\gamma$$

$$= \sum_{i,j}\int_\Omega \Big(2\partial_i\boldsymbol{\zeta}_j\partial_j\big(s\,\partial_i s\big) + 2s\partial^2_{ij}V\,\partial_i s\,\boldsymbol{\zeta}_j\Big)\,\mathrm{d}\gamma = \sum_{i,j}\int_\Omega \Big(\partial_i\boldsymbol{\zeta}_j\big(2s\,\partial^2_{ij}s + 2\partial_i s\,\partial_j s\big) + 2s\partial^2_{ij}V\,\partial_i s\,\boldsymbol{\zeta}_j\Big)\,\mathrm{d}\gamma,$$

we obtain

$$\int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\sum_{i,j}\int_\Omega 4\partial_i\,\boldsymbol{\zeta}_j\,\partial_i s\,\partial_j s\,\mathrm{d}\gamma + B = \sum_{i,j}\int_\Omega \Big(\partial_i\,\boldsymbol{\zeta}_j\big(-2\partial_i s\,\partial_j s + 2s\,\partial^2_{ij}s\big) + 2s\partial^2_{ij}V\,\partial_i s\,\boldsymbol{\zeta}_j\Big)\,\mathrm{d}\gamma$$

which yields (5.73).

In order to prove (5.76) we proceed in a slightly different way, starting again from the decomposition (5.76)

$$B = -2\int_\Omega s\mathrm{D}\operatorname{div}_\gamma\boldsymbol{\zeta}\cdot\mathrm{D}s\,\mathrm{d}\gamma = 2\int_\Omega \operatorname{div}_\gamma\boldsymbol{\zeta}\ \operatorname{div}_\gamma\big(s\mathrm{D}s\big)\,\mathrm{d}\gamma = 2\int_\Omega \operatorname{div}_\gamma\boldsymbol{\zeta}\,\big(s\,\Delta_\gamma s + |\mathrm{D}s|^2\big)\,\mathrm{d}\gamma \tag{5.77}$$

$$A = -4\int_\Omega \mathrm{D}\boldsymbol{\zeta}\mathrm{D}s\cdot\mathrm{D}s\,\mathrm{d}\gamma = -4\sum_{i,j}\int_\Omega \partial_i\boldsymbol{\zeta}_j\,\partial_i s\,\partial_j s\,\mathrm{d}\gamma = 4\sum_{i,j}\int_\Omega \boldsymbol{\zeta}_j\tilde\partial_i\big(\partial_i s\partial_j s\big)\,\mathrm{d}\gamma$$
$$= 4\sum_{i,j}\int_\Omega \boldsymbol{\zeta}_j\Big(\tilde\partial_i\partial_i s\,\partial_j s + \partial^2_{ij}s\partial_i s\Big)\,\mathrm{d}\gamma = \int_\Omega\Big(4\Delta_\gamma s\mathrm{D}s + 2\mathrm{D}\big(|\mathrm{D}s|^2\big)\Big)\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma$$
$$= \int_\Omega 4\Delta_\gamma s\mathrm{D}s\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma - 2\int_\Omega |\mathrm{D}s|^2\,\mathrm{div}_\gamma\boldsymbol{\zeta}\,\mathrm{d}\gamma. \tag{5.78}$$

Summing up the two contributions (5.77) and (5.78) we end up with

$$\int_\Omega v\boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma = \int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = 2\int_\Omega s\Delta_\gamma s\;\mathrm{div}_\gamma\boldsymbol{\zeta}\,\mathrm{d}\gamma + 4\int_\Omega \Delta_\gamma s\mathrm{D}s\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma \quad \forall\,\boldsymbol{\zeta}\in C^\infty_c(\Omega;\mathbb{R}^d), \tag{5.79}$$

which shows that $s\Delta_\gamma s\in W^{1,1}_{\gamma,\mathrm{loc}}(\Omega)$. Integrating by parts the left hand side of (5.79) once more we obtain (5.74). □

In a similar way way, combining Corollary 5.8, Theorem 5.9 with $\gamma=\mathcal{L}^d$, and the perturbation of §2.5, we obtain the analogous differential characterizations of the limiting subdifferential of $\mathscr{F}^f$, which in particular apply to (EE$_1$) and Definition 1.5.

Corollary 5.10 (**Limiting subdifferential of $\mathscr{F}^f$**). *The functional $\mathscr{F}^f := \frac12\mathscr{I}_2(\cdot|\,\mathcal{L}^d)+\langle f,\cdot\rangle$ has a regular subdifferential, according to Definition 2.10. If $\mu=u\mathcal{L}^d=r^2\mathcal{L}^d\in D(\partial_\ell\mathscr{F}^f)$ then $r\in W^{2,2}(\Omega)$ with $\partial_{\boldsymbol{n}}r=0$, $\sqrt{r}\in W^{1,4}(\Omega)$, and $\boldsymbol{\xi}=\partial_\ell\mathscr{F}^f(\mu)$ is characterized by*

$$\int_\Omega \boldsymbol{\xi}\cdot\boldsymbol{\zeta}\,\mathrm{d}\mu = -\int_\Omega\Big(4\mathrm{D}\boldsymbol{\zeta}\mathrm{D}r\cdot\mathrm{D}r + \mathrm{D}\big(\,\mathrm{div}\,\boldsymbol{\zeta}\big)\cdot\mathrm{D}u - u\mathrm{D}f\cdot\boldsymbol{\zeta}\Big)\,\mathrm{d}x \quad \forall\,\boldsymbol{\zeta}\in C^2_{\mathrm c}(\mathbb{R}^d;\mathbb{R}^d),\ \boldsymbol{\zeta}\cdot\boldsymbol{n}=0, \tag{5.80}$$

$$= \int_\Omega\Big(\sum_{i,j}\mathsf{L}_{ij}\partial_i\boldsymbol{\zeta}_j + u\sum_i\partial_i f\,\boldsymbol{\zeta}_i\Big)\,\mathrm{d}x \quad \forall\,\boldsymbol{\zeta}\in C^2_{\mathrm c}(\Omega;\mathbb{R}^d) \tag{5.81}$$

$$= -\int_\Omega\Big(2\mathrm{D}(r\Delta r) - 4\Delta r\mathrm{D}r - u\mathrm{D}f\Big)\cdot\boldsymbol{\zeta}\,\mathrm{d}\gamma, \quad \forall\,\boldsymbol{\zeta}\in C^2_{\mathrm c}(\Omega;\mathbb{R}^d) \tag{5.82}$$

where $\mathsf{L}_{ij}=2\big(r\partial^2_{ij}r-\partial_i r\partial_j r\big)$; in particular

$$r\Delta r\in W^{1,1}(\Omega) \quad \textit{and} \quad \boldsymbol{q}=-u\boldsymbol{\xi}=2\mathrm{D}\big(r\Delta r\big)-4\Delta r\mathrm{D}r-u\mathrm{D}f. \tag{5.83}$$

Corollary 5.11. *Suppose that Ω is a (convex)* cone*, V has bounded second order derivatives, f satisfies* (1.4)*, $\mu=v\gamma=u\mathcal{L}^d\in\mathcal{P}^r_2(\mathbb{R}^d)$, and that $\boldsymbol{\xi}_1\in\partial_\ell\mathscr{G}(\mu)$, $\boldsymbol{\xi}_2\in\partial_\ell\mathscr{F}^f(\mu)$. Then*

$$-\int_{\mathbb{R}^d}\boldsymbol{\xi}_1(x)\cdot x\,\mathrm{d}\mu = \mathscr{I}_2(\mu|\,\gamma) - \int_{\mathbb{R}^d}\Big(\mathrm{D}v\cdot\mathrm{D}V + \mathrm{D}^2V\mathrm{D}v\cdot x\Big)\,\mathrm{d}\gamma, \tag{5.84}$$

$$-\int_{\mathbb{R}^d}\boldsymbol{\xi}_2(x)\cdot x\,\mathrm{d}\mu = \mathscr{I}_2(\mu|\,\mathcal{L}^d) - \int_{\mathbb{R}^d}\mathrm{D}f\cdot x\,\mathrm{d}\mu(x). \tag{5.85}$$

Proof. Choose a nonnegative cutoff function $\vartheta\in C^\infty_{\mathrm c}(\mathbb{R}^d)$ such that $\vartheta\equiv 1$ in the unit ball of $\mathbb{R}^d$ and set $\vartheta_n(x):=\vartheta(x/n)$, $\boldsymbol{\zeta}_n(x):=x\vartheta_n(x)$. Since

$$\begin{gathered}\mathrm{D}\boldsymbol{\zeta}_n=\vartheta_n\mathsf{Id}+x\otimes\mathrm{D}\vartheta_n,\quad \mathrm{div}\,\boldsymbol{\zeta}=d\,\vartheta_n+\mathrm{D}\vartheta_n\cdot x,\quad \mathrm{D}(\mathrm{div}\,\boldsymbol{\zeta})=(d+1)\mathrm{D}\vartheta_n+\mathrm{D}^2\vartheta_n x,\\ \mathrm{div}_\gamma\boldsymbol{\zeta}=d\,\vartheta_n+\mathrm{D}\vartheta_n\cdot x-\vartheta_n\mathrm{D}V\cdot x,\\ \mathrm{D}(\mathrm{div}_\gamma\boldsymbol{\zeta})=(d+1)\mathrm{D}\vartheta_n+\mathrm{D}^2\vartheta_n x-\mathrm{D}\vartheta_n(\mathrm{D}V\cdot x)-\vartheta_n\big(\mathrm{D}^2V\,x+\mathrm{D}V\big),\end{gathered} \tag{5.86}$$

and

$$|\mathrm{D}\vartheta_n|\le C/n,\quad |\mathrm{D}^2\vartheta_n x|\le C/n,\quad |x\otimes\mathrm{D}\vartheta_n|\le C,\quad \lim_{n\to\infty}|x\otimes\mathrm{D}\vartheta_n|=0\quad\forall\,x\in\mathbb{R}^d,\ n\in\mathbb{N},$$

substituting in (5.72) and (5.80) we can pass to the limit as $n\to+\infty$ thanks to Lebesgue Dominated Convergence Theorem, obtaining (5.84) and (5.85) respectively. □

In the particular cases of the Lebesgue or the Gaussian Measure $\gamma=\mathcal{L}^d$ we thus have the same formulae (4.35) and (4.36) we proved for the strong subdifferential.

6. Proofs of the main results of §1.9: gradient flow of $\mathscr{F}^f$

6.1. **Proof of Theorem 1 (the case $\mathscr{F}^f(\mu_0) < +\infty$).** Since the initial energy $\mathscr{F}^f(\mu_0)$ is finite, the main claims of Theorem 1 are a direct consequence of the abstract Theorem 2.13 for the functional

$$\phi(\mu) := \mathscr{F}^f(\mu) = \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) + \langle f, \mu\rangle,$$

and the results we presented in the previous sections. Let us quickly check the various statements.

Claim i) mainly follows from the properties **a)** and **b)** of Theorem 2.13. Observe that (1.4) and Lemma 2.4 ensure the narrow lower semicontinuity of $\mathscr{F}^f$ on sequences in $\mathcal{P}(\mathbb{R}^d)$ with bounded second order moment; moreover, for every $\tau_o > 0$ the function $x \mapsto f(x) + \frac{1}{2\tau_o}|x|^2$ is bounded from below so that assumptions (2.66a,b) are satisfied. (1.72) says that the initial data M^0_τ, μ_0 satisfy (2.96).

Concerning the strong convergence (1.74) of the densities in $L^p(\Omega)$ and the convergence of the Entropy (1.75), they follow from Remark 2.16 and the *a priori* estimate (2.97) which yields a time step τ_* such that for every $T > 0$

$$\sup_{t\in[0,T],\ \tau\le\tau_*}\Big(\mathsf{m}_2(\overline{M}_{\tau,t}) + \mathscr{I}_2(\overline{M}_{\tau,t}|\,\mathcal{L}^d)\Big) < +\infty, \quad \sup_{\tau\le\tau_*}\int_0^T\int_\Omega \big|\overline{\boldsymbol{v}}_{\tau,t}\big|^2\,\mathrm{d}\overline{M}_{\tau,t}\,\mathrm{d}t < +\infty. \tag{6.1}$$

By choosing $\boldsymbol{\xi} := -\overline{\boldsymbol{v}}_{\tau,t}$ in the second order estimate (5.68) (recall that $-\overline{\boldsymbol{v}}_{\tau,t} \in \partial_s\mathscr{F}^f(\overline{M}_{\tau,t})$ by (2.94)) and taking into account that

$$\|\mathrm{D}f\|^2_{L^2_{\overline{M}_{\tau,t}}(\Omega;\mathbb{R}^d)} \le 2C_f\big(1 + \mathsf{m}_2^2(\overline{M}_{\tau,t})\big) \quad \text{so that} \quad \sup_{t\in[0,T],\ \tau\le\tau_*}\|\mathrm{D}f\|^2_{L^2_{\overline{M}_{\tau,t}}(\Omega;\mathbb{R}^d)} < +\infty, \tag{6.2}$$

the integral estimate of (6.1) yields

$$\sup_{\tau<\tau_*}\int_0^T\int_\Omega\Big(\big|\mathrm{D}^2\overline{R}_{\tau,t}\big|^2 + \big|\mathrm{D}\sqrt{\overline{R}_{\tau,t}}\big|^4\Big)\,\mathrm{d}x\,\mathrm{d}t < +\infty, \quad \text{where} \quad \overline{M}_{\tau,t} = (\overline{R}_{\tau,t})^2\mathcal{L}^d. \tag{6.3}$$

Since $\overline{R}_{\tau_k,t} \to r_t = \sqrt{u_t}$ strongly in $L^2(\Omega)$ keeping the $L^2(\Omega)$-norm constantly equal to 1, we easily get the strong convergence in $L^2(0,T;L^2(\Omega))$; (6.3) yields the weak $L^2(0,T;W^{2,2}(\Omega))$ convergence of $\overline{R}_{\tau_k}$ and the strong convergence in $L^2(0,T;W^{1,2}(\Omega))$ (e.g. by proving the convergence of the $L^2(0,T;W^{1,2}(\Omega))$ norm, thanks to the identity of (5.70)). We thus obtain (1.77) and therefore (up to a further extraction of a subsequence) (1.76).

The regularity (1.78) of the density of μ still follows from the estimate (6.3) (the homogeneous Neumann boundary condition is a consequence of Corollary 5.10, since $\mu_t \in D(\partial_\ell\mathscr{F}^f)$ for a.e. $t > 0$). The property of $r\Delta r$ is a consequence of (5.83).

Claim ii) is a consequence of (2.103), point **c)** of Theorem 2.13, and Corollary 5.10: we simply write the distributional formulation of the continuity equation (2.103)

$$\iint_{\Omega_\infty}\Big(\partial_t\zeta + \boldsymbol{v}\cdot\mathrm{D}\zeta\Big)\,\mathrm{d}\mu_t\,\mathrm{d}t = 0 \quad \forall\,\zeta \in C^\infty_c(\mathbb{R}^d\times(0,+\infty)), \quad \partial_{\boldsymbol{n}}\zeta = 0 \text{ on } \partial\Omega, \tag{6.4}$$

and we use the description of $\boldsymbol{v}$ given by (5.82) choosing $\boldsymbol{\zeta} := \mathrm{D}\zeta$.

Claim iii): by Point *v)* of Proposition 2.15 we get the absolute continuity of the map $t \mapsto \mathscr{H}(\mu_t|\,\mathcal{L}^d)$ via uniform bound on $\mathscr{I}_2(\mu_t|\,\mathcal{L}^d)$ and the estimate (1.80) on the velocity. Applying (2.119) we get

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}(\mu_t|\,\mathcal{L}^d) = -\int_\Omega \boldsymbol{v}_t\cdot\mathrm{D}u_t\,\mathrm{d}x \le \|\boldsymbol{v}\|_{L^2_{\mu_t}(\Omega;\mathbb{R}^d)}\Big(\mathscr{I}_2(\mu_t|\,\mathcal{L}^d)\Big)^{1/2} \quad \text{for a.e. } t > 0. \tag{6.5}$$

Taking into account (5.68) (with $\boldsymbol{\xi} = -\boldsymbol{v}_t$) we thus get (1.81) and (1.82).

6.2. **Proof of Theorem 2 (Quadratic moments).** If a family of measure μ_t satisfies the continuity equation (6.4) with

$$\int_0^T\int_\Omega |\boldsymbol{v}_t|^2\,\mathrm{d}\mu_t\,\mathrm{d}t < +\infty \quad \forall\,T > 0, \tag{6.6}$$

then it is absolutely continuous in $\mathcal{P}_2(\Omega)$ [AGS05, Thm. 8.3.1] and in particular its quadratic moments are absolutely continuous, too. Moreover, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int_\Omega |x|^2\,\mathrm{d}\mu_t = \int_\Omega \boldsymbol{v}_t\cdot x\,\mathrm{d}\mu_t \quad \text{for a.e. } t>0. \tag{6.7}$$

Since Ω is a *cone*, (1.83) follows directly from (5.85), being $-\boldsymbol{v}\in\partial_\ell\mathscr{F}^f(\mu_t)$. In particular, multiplying (6.7) by the factor $e^{-2\beta t}$ and integrating it in the time interval $(0,t)$, we obtain

$$\frac{e^{-2\beta t}}{2}\mathsf{m}_2^2(\mu_t)+\int_0^t e^{-2\beta s}\int_\Omega \big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}\mu_s\,\mathrm{d}s = \frac{1}{2}\mathsf{m}_2^2(\mu_0)+\int_0^t e^{-2\beta s}\mathscr{I}_2(\mu_s|\,\mathcal{L}^d)\,\mathrm{d}s \tag{6.8}$$

for every choice of $t>0$ and $\beta\in\mathbb{R}$.

We choose the parameter β greater than $2C_f$ (the constant introduced in (1.4)), so that

$$\inf_{x\in\Omega}\frac{\big(\mathrm{D}f(x)+\beta x\big)\cdot x}{|x|^2}\geq -2C_f>-\infty. \tag{6.9}$$

We observe that at the discrete level

$$\frac{1}{2}\mathsf{m}_2^2(M_\tau^n)-\frac{1}{2}\mathsf{m}_2^2(M_\tau^{n-1})=\frac{1}{2}\int_\Omega\Big(|x|^2-|\boldsymbol{r}_\tau^n(x)|^2\Big)\,\mathrm{d}M_\tau^n(x)\leq\int_{\mathbb{R}^d}(x-\boldsymbol{r}_\tau^n)\cdot x\,\mathrm{d}M_\tau^n=\tau\int_{\mathbb{R}^d}\boldsymbol{v}_\tau^n\cdot x\,dM_\tau^n,$$

for every $n\in\mathbb{N}$, being $\boldsymbol{r}_\tau^n=\boldsymbol{i}-\tau\boldsymbol{v}_\tau^n=T_o(M_\tau^n,M_\tau^{n-1})$. Since $-\boldsymbol{v}_\tau^n\in\partial_s\mathscr{F}^f(M_\tau^n)$ by (2.76), (5.85) yields

$$\frac{1}{2}\mathsf{m}_2^2(M_\tau^n)-\frac{1}{2}\mathsf{m}_2^2(M_\tau^{n-1})+\tau\int_{\mathbb{R}^d}\mathrm{D}f\cdot x\,\mathrm{d}M_\tau^n\leq\tau\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d). \tag{6.10}$$

Since $\beta>0$ we have

$$e^{-2\beta\tau}\leq 1-2\beta\tau+\frac{(2\beta\tau)^2}{2}=1-2\beta\tau\big(1-\beta\tau\big)\quad\forall\,\tau>0,$$

so that

$$\frac{1-e^{-2\beta\tau}}{2}\mathsf{m}_2^2(M_\tau^n)\geq\tau\beta(1-\tau\beta)\mathsf{m}_2^2(M_\tau^n),$$

and therefore (6.10) yields

$$\frac{e^{-2\beta\tau}}{2}\mathsf{m}_2^2(M_\tau^n)-\frac{1}{2}\mathsf{m}_2^2(M_\tau^{n-1})+\tau\int_\Omega\big(\mathrm{D}f+\beta(1-\tau\beta)x\big)\cdot x\,\mathrm{d}M_\tau^n\leq\tau\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d). \tag{6.11}$$

Multiplying this inequality by $e^{-2\beta t_\tau^{n-1}}=e^{-2\beta(n-1)\tau}=e^{2\beta\tau}\,e^{-2\beta t_\tau^n}$ we obtain

$$\begin{aligned}\frac{e^{-2\beta t_\tau^n}}{2}\mathsf{m}_2^2(M_\tau^n)-\frac{e^{-2\beta t_\tau^{n-1}}}{2}\mathsf{m}_2^2(M_\tau^{n-1})+\tau e^{2\beta\tau}e^{-2\beta t_\tau^n}\int_{\mathbb{R}^d}\big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}M_\tau^n\\ \leq\tau e^{2\beta\tau}e^{-2\beta t_\tau^n}\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d)+\beta^2\tau^2\mathsf{m}_2^2(M_\tau^n).\end{aligned} \tag{6.12}$$

Let us fix the final point $t\in((N-1)\tau,N\tau]$ and let us recall that (see (2.90)) $s\mapsto\overline{s}_\tau$ is the piecewise constant function taking the value $t_\tau^n=n\tau$ in the interval $(n-1)\tau<s\leq n\tau$; summing the above inequalities for $n=1$ to N, we obtain

$$\begin{aligned}\frac{e^{-2\beta\bar{t}_\tau}}{2}\mathsf{m}_2^2(\overline{M}_{\tau,t})+e^{2\beta\tau}\int_0^{\bar{t}_\tau}e^{-2\beta\overline{s}_\tau}\int_{\mathbb{R}^d}\big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}\overline{M}_{\tau,s}\,\mathrm{d}s\\ \leq\frac{1}{2}\mathsf{m}_2^2(M_\tau^0)+e^{2\beta\tau}\int_0^{\bar{t}_\tau}e^{-2\beta\overline{s}_\tau}\mathscr{I}_2(\overline{M}_{\tau,s}|\,\mathcal{L}^d)\,\mathrm{d}s+\beta^2\tau\int_0^{\bar{t}_\tau}\mathsf{m}_2^2(\overline{M}_{\tau,s})\,\mathrm{d}s.\end{aligned} \tag{6.13}$$

Since $\mathrm{D}f$ is continuous, (6.9) and the the uniform boundedness of the second moment of $\overline{M}_{\tau,t}$ given by (6.1) yield

$$\liminf_{k\to\infty}\int_\Omega\big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}\overline{M}_{\tau_k,s}\geq\int_\Omega\big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}\mu_s\quad\forall\,s\geq 0;$$

Fatou's Lemma and the fact that $\bar{t}_\tau \to t$, $\bar{s}_\tau \to s$ uniformly as $\tau \downarrow 0$ yield

$$\liminf_{k\to\infty} e^{2\beta\tau_k}\int_0^{\bar{t}_{\tau_k}} e^{-2\beta\bar{s}_{\tau_k}}\int_\Omega \big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}\overline{M}_{\tau_k,s}\,\mathrm{d}s \ge \int_0^t\int_\Omega e^{-2\beta s}\big(\mathrm{D}f+\beta x\big)\cdot x\,\mathrm{d}\mu_s\,\mathrm{d}s; \tag{6.14}$$

(1.77) and (6.1) yield

$$\begin{gathered}\lim_{k\to\infty} e^{2\beta\tau_k}\int_0^{\bar{t}_{\tau_k}} e^{-2\beta\bar{s}_{\tau_k}}\mathscr{I}_2(\overline{M}_{\tau_k,s}|\,\mathcal{L}^d)\,\mathrm{d}s = \int_0^t \mathscr{I}_2(\mu_s|\,\mathcal{L}^d)\,\mathrm{d}s,\\ \lim_{\tau\downarrow 0}\beta^2\tau\int_0^{\bar{t}_\tau}\mathsf{m}_2^2(\overline{M}_{\tau,s})\,\mathrm{d}s = 0.\end{gathered} \tag{6.15}$$

Since assumption (1.84) ensures the convergence of $\mathsf{m}_2(M_\tau^0)$ to $\mathsf{m}_2(\mu_0)$, taking the "lim sup" of (6.13) as $\tau_k \downarrow 0$ and invoking (6.14), (6.15), and (6.8), we obtain

$$\limsup_{k\to\infty}\mathsf{m}_2^2(\overline{M}_{\tau_k,t}) \le \mathsf{m}_2^2(\mu_t)\quad \forall\, t>0, \tag{6.16}$$

which is equivalent to (1.85).

6.3. **Proof of Theorem 3 (Energy inequalities and asymptotic behavior).** The energy inequality (1.86) is an immediate consequence of (2.107) and (2.100), whenever one can show that

$$\lim_{k\to\infty}\int_\Omega f(x)\,\mathrm{d}\overline{M}_{\tau_k,t} = \int_\Omega f(x)\,\mathrm{d}\mu_t\quad \forall\, t\ge 0; \tag{6.17}$$

for, combining (6.17) and (1.76), we get

$$\lim_{k\to\infty}\phi(\overline{M}_{\tau_k,t}) = \lim_{k\to\infty}\mathscr{F}^f(\overline{M}_{\tau_k,t}) = \mathscr{F}^f(\mu_t) = \varphi(\mu_t)\quad \text{for a.e. } t>0. \tag{6.18}$$

When assumption **H1)** is satisfied, (6.17) follows directly from the narrow convergence of $\overline{M}_{\tau_k,t}$ (1.73) and the uniform estimate (6.1) on its quadratic moments, having f a sub-quadratic growth at ∞.

In the case of assumption **H2)**, we can invoke Theorem 2 which shows the convergence of $\overline{M}_{\tau_k,t}$ in $\mathcal{P}_2(\Omega)$ and thus yields (6.17) even for functions f with quadratic growth.

(1.87) follows by the same argument we used in the proof of *Claim iii)* of Theorem 1: here we invoke Remark 2.5 which shows that

$$\mathscr{F}^f(\mu) = \mathscr{G}(\mu) = \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma). \tag{6.19}$$

In particular, the velocity vector field $\boldsymbol{v}_t$ also satisfies $-\boldsymbol{v}_t \in \partial_\ell(\mathscr{G}(\mu_t))$ for a.e. $t>0$. Moreover, thanks to (6.18), for a.e. $t>0$ we have

$$\lim_{k\to\infty}\mathscr{I}_2(\overline{M}_{\tau_k,t}|\,\gamma) = \mathscr{I}_2(\mu_t|\,\gamma), \tag{6.20}$$

and there exist vectors $-\overline{\boldsymbol{v}}_{\tau_k,t} \in \partial_s\mathscr{F}(\overline{M}_{\tau_k,t})$ such that

$$\overline{M}_{\tau_k,t}\overline{\boldsymbol{v}}_{\tau_k,t} \rightharpoonup \mu_t\boldsymbol{v}_t\quad \text{in } \big[\mathcal{M}_{\mathrm{loc}}(\Omega)\big]^d,\quad \liminf_{k\to\infty}\int_\Omega\big|\overline{\boldsymbol{v}}_{\tau_k,t}\big|^2\,d\overline{M}_{\tau_k,t} < +\infty.$$

(5.48) of Theorem 5.5 thus yields that for a.e. $t>0$

$$\mathscr{K}_{-1}(\mu_t|\,\gamma) + \mathscr{I}_2^V(\mu_t|\,\gamma) = \mathscr{P}(\mu_t|\,\gamma) \le -\int_\Omega \boldsymbol{v}_t\cdot\frac{\mathrm{D}v_t}{v_t}\,\mathrm{d}\mu_t \overset{(2.119)}{=} -\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}(\mu_t|\,\gamma). \tag{6.21}$$

The first inequality of (1.87) follows now from Schwarz inequality, (1.86), and (6.19).

(1.89) follows now by neglecting the (nonnegative) contribution of $\mathscr{K}_{-1}(\mu_t|\,\gamma)$ and integrating (1.87); taking account of the Logarithmic Sobolev inequality (2.115), we still obtain (1.88) by integration.

6.4. **Proof of Theorem 4 (Regularizing effect).** In order to prove the theorem, we need other *a priori* estimates on the entropy of the discrete solutions of the Minimizing Movement Scheme. Let us keep the notation of section 2.5 for the discrete solution $\{M_\tau^n\}_{n\in\mathbb{N}}$ of (1.44) related to the functional

$$\mathscr{F}^f(\mu) := \mathscr{I}_2(\mu|\,\mathcal{L}^d) + \langle f, \mu\rangle.$$

In particular M_τ^n satisfies

$$\frac{W^2(M_\tau^n, M_\tau^{n-1})}{2\tau} + \mathscr{F}^f(M_\tau^n) \le \frac{W^2(\mu, M_\tau^{n-1})}{2\tau} + \mathscr{F}^f(\mu) \quad \forall\, \mu \in \mathcal{P}_2(\Omega). \tag{6.22}$$

and $\boldsymbol{r}_\tau^n, \boldsymbol{v}_\tau^n, U_\tau^n, \boldsymbol{\eta}_\tau^n$ are defined according to

$$\boldsymbol{r}_\tau^n = T_o(M_\tau^n, M_\tau^{n-1}), \quad -\boldsymbol{v}_\tau^n := \frac{\boldsymbol{r}_\tau^n - \boldsymbol{i}}{\tau} \in \partial_s\phi(M_\tau^n) \quad U_\tau^n := \frac{dM_\tau^n}{d\mathcal{L}^d}, \quad \boldsymbol{\eta}_\tau^n := \frac{\mathrm{D}U_\tau^n}{U_\tau^n}. \tag{6.23}$$

$\overline{M}_\tau, \overline{\boldsymbol{v}}_\tau$ are the corresponding piecewise constant functions

$$\overline{M}_{\tau,t} := M_\tau^n, \quad \overline{\boldsymbol{v}}_\tau := \boldsymbol{v}_\tau^n, \quad \overline{t}_\tau := n\tau = t_\tau^n \qquad \text{if} \quad (n-1)\tau < t \le n\tau = t_\tau^n. \tag{6.24}$$

$\boldsymbol{v}_\tau^n, \boldsymbol{\eta}_\tau^n$ are vector fields in $L^2_{M_\tau^n}(\Omega;\mathbb{R}^d)$ and we also set

$$|\!|\!|\boldsymbol{v}_\tau^n|\!|\!|^2 := \int_\Omega |\boldsymbol{v}_\tau^n|^2\,\mathrm{d}M_\tau^n \overset{(6.23)}{=} \frac{W^2(M_\tau^n, M_\tau^{n-1})}{\tau^2}, \quad |\!|\!|\boldsymbol{\eta}_\tau^n|\!|\!|^2 := \int_\Omega |\boldsymbol{\eta}_\tau^n|^2\,\mathrm{d}M_\tau^n = \mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d). \tag{6.25}$$

Proposition 6.1 (**Discrete estimates**)**.** *For each couple of integers* $0 \le m < k$ *we have*

$$\mathscr{F}^f(M_\tau^k) + \frac{\tau}{2}\sum_{n=m+1}^k |\!|\!|\boldsymbol{v}_\tau^n|\!|\!|^2 \le \inf_{\mu\in\mathcal{P}_2^r(\Omega)} \frac{1}{2\tau}W^2(\mu, M_\tau^m) + \mathscr{F}^f(\mu) \le \mathscr{F}^f(M_\tau^m), \tag{6.26}$$

$$\mathscr{H}(M_\tau^k|\,\mathcal{L}^d) + \frac{\tau}{2}\sum_{n=m+1}^k \mathscr{K}_{-1}(M_\tau^n|\,\mathcal{L}^d) \le \mathscr{H}(M_\tau^m|\,\mathcal{L}^d) + \tau\, C_{f,d}\sum_{n=m+1}^k \big(1 + \mathsf{m}_2^2(M_\tau^n)\big) \tag{6.27}$$

where $C_{f,d} := 6(2+d)^{1/3}C_f^{4/3}$*; moreover, when* Ω *is a (convex) cone, for every* $\varepsilon > 0$ *we have*

$$\begin{aligned}\frac{1}{2}\mathsf{m}_2^2(M_\tau^k) &\le \frac{1}{2}\mathsf{m}_2^2(M_\tau^m) + \tau\sum_{n=m+1}^k \Big(\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d) + 2C_f\big(1 + \mathsf{m}_2^2(M_\tau^n)\big)\Big)\\ &\le \frac{1}{2}\mathsf{m}_2^2(M_\tau^m) + \tau\sum_{n=m+1}^k \Big(\varepsilon\mathscr{K}_{-1}(M_\tau^n|\,\mathcal{L}^d) + 2C_f + \frac{d+2}{2\varepsilon} + 2C_f\mathsf{m}_2^2(M_\tau^n)\Big).\end{aligned} \tag{6.28}$$

Proof. The first inequality of (6.26) follows simply by choosing $\mu := M_\tau^{n-1}$ in (6.22) and summing up from $n = m+2$ to $n = k$, owing (6.25): we obtain

$$\mathscr{F}^f(M_\tau^k) + \frac{\tau}{2}\sum_{n=m+2}^k |\!|\!|\boldsymbol{v}_\tau^n|\!|\!|^2 \le \mathscr{F}^f(M_\tau^{m+1}).$$

Adding (6.22) for $n = m+1$ and taking the infimum with respect to μ we get (6.26).

In order to get (6.27), we apply the displacement convexity inequality (2.113) for $\gamma = \mathcal{L}^d$ with $\lambda = 0$, $\mu := M_\tau^n$, and $\nu := M_\tau^{n-1}$, obtaining

$$\mathscr{H}(M_\tau^n|\,\mathcal{L}^d) - \mathscr{H}(M_\tau^{n-1}|\,\mathcal{L}^d) \le \tau\int_\Omega \boldsymbol{v}_\tau^n\cdot\boldsymbol{\eta}_\tau^n\,\mathrm{d}M_\tau^n. \tag{6.29}$$

Since, by (2.76) and (2.81a), $-\boldsymbol{v}_\tau^n = \partial_s\mathscr{F}^f(M_\tau^n) = \frac{1}{2}\partial_s\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d) + \mathrm{D}f$, applying (5.68) of Corollary 5.8 we get

$$\int_\Omega \boldsymbol{v}_\tau^n\cdot\boldsymbol{\eta}_\tau^n\,\mathrm{d}M_\tau^n \le -\mathscr{K}_{-1}(M_\tau^n|\,\mathcal{L}^d) - \int_\Omega \mathrm{D}f\cdot\boldsymbol{\eta}_\tau^n\,\mathrm{d}M_\tau^n. \tag{6.30}$$

Hölder inequality and (1.4) yield for every $\varepsilon > 0$

$$
\begin{aligned}
-\int_\Omega \mathrm{D}f \cdot \boldsymbol{\eta}_\tau^n \,\mathrm{d}M_\tau^n &\le 2C_f |||\boldsymbol{\eta}_\tau^n||| \Big(\int_\Omega \big(1+|x|^2\big)\,\mathrm{d}M_\tau^n\Big)^{1/2} \le 2C_f \Big(\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d)\Big)^{1/2}\Big(1+\mathsf{m}_2(M_\tau^n)\Big)\\
&\le \frac{\varepsilon^4}{4}\Big(\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d)\Big)^2 + \frac{3}{4\varepsilon^{4/3}}\Big(2C_f\big(1+\mathsf{m}_2(M_\tau^n)\big)\Big)^{4/3}\\
&\le \frac{\varepsilon^4}{4}\Big(\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d)\Big)^2 + \frac{6}{\varepsilon^{4/3}}C_f^{4/3}\Big(1+\mathsf{m}_2^2(M_\tau^n)\Big).
\end{aligned}
$$

Choosing now $\varepsilon^4 = (2+d)^{-1}$ and recalling (5.69) we get

$$
-\int_\Omega \mathrm{D}f \cdot \boldsymbol{\eta}_\tau^n \,\mathrm{d}M_\tau^n \le \frac{1}{2}\mathscr{K}_{-1}(\mu|\,\mathcal{L}^d) + 6(2+d)^{1/3}C_f^{4/3}\Big(1+\mathsf{m}_2^2(M_\tau^n)\Big). \tag{6.31}
$$

Inserting (6.30) and (6.31) in (6.29) and summing it from $n := m+1$ to k we obtain (6.27).

(6.28) follows similarly from (6.10) and (1.4); the last inequality is a direct consequence of (5.69), which yields for every $\varepsilon > 0$

$$
\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d) \le \frac{2+d}{2\varepsilon} + \frac{\varepsilon}{2(2+d)}\big(\mathscr{I}_2(M_\tau^n|\,\mathcal{L}^d)\big)^2 \le \frac{2+d}{2\varepsilon} + \varepsilon\mathscr{K}_{-1}(M_\tau^n|\,\mathcal{L}^d). \qquad \square
$$

Before stating the next *a priori* bound, let us recall that by (2.26)

$$
\mathscr{H}(\mu|\,\mathcal{L}^d) + \pi\,\mathsf{m}_2^2(\mu) \ge 0 \quad \forall\,\mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{6.32}
$$

Corollary 6.2 (A priori bounds for discrete solutions)**.** *Let us suppose that*

$$
H_\tau^0 := \mathscr{H}(M_\tau^0|\,\mathcal{L}^d) + (\pi + \tfrac{1}{2})\mathsf{m}_2^2(M_\tau^0) < +\infty. \tag{6.33}
$$

There exists a time step $\tau_o > 0$ solely dependent on C_f and the dimension d such that for every $0 < \tau < \tau_o$ and every time $t > 0$

$$
\mathscr{H}(\overline{M}_{\tau,t}|\,\mathcal{L}^d) + (\pi+\tfrac{1}{2})\mathsf{m}_2^2(\overline{M}_{\tau,t}) + \frac{1}{4}\int_0^{\bar{t}_\tau} \mathscr{K}_{-1}(\overline{M}_{\tau,r}|\,\mathcal{L}^d)\,\mathrm{d}r \le C(M_\tau^0, t), \tag{6.34}
$$

where

$$
C(M_\tau^0, t) := \big(H_\tau^0 + C\bar{t}_\tau\big)e^{C(t+2\tau)}, \tag{6.35}
$$

and the constant C depends only on C_f and d. Moreover,

$$
\sqrt{\bar{t}_\tau}\,\mathscr{I}_2(\overline{M}_{\tau,t}|\,\mathcal{L}^d) \le D(M_\tau^0, t), \quad \textit{with} \quad D(M_\tau^0,t) = O(1) \quad \textit{as } t \downarrow 0 \quad \textit{uniformly w.r.t. } \tau, \tag{6.36}
$$

where

$$
D(M_\tau^0, t) := C\Big(C^{1/2}(M_\tau^0, t) + \big(1 + C(M_\tau^0,t)\big)\sqrt{\bar{t}_\tau}\Big). \tag{6.37}
$$

Proof. In order to keep simpler notation, let us set for a given $\tau > 0$

$$
H^k := \mathscr{H}(M_\tau^k|\,\mathcal{L}^d) + (\pi+\tfrac{1}{2})\mathsf{m}_2^2(M_\tau^k), \quad m^k := \tfrac{1}{2}\mathsf{m}_2^2(M_\tau^k), \quad K^k := \mathscr{K}_{-1}(M_\tau^n|\,\mathcal{L}^d),
$$

observing that

$$
H^k \ge m^k \quad \forall\, k \in \mathbb{N}.
$$

Summing up (6.27) and (the second inequality of) (6.28) multiplied by $1+2\pi$, and choosing the initial integer $m = 0$, $\varepsilon^{-1} := 4(1+2\pi)$, we obtain

$$
H^k + \frac{\tau}{4}\sum_{n=1}^k K^k \le H^0 + C_{f,d}^{(1)}\tau k + \tau C_{f,d}^{(2)}\sum_{n=1}^k m^n \tag{6.38}
$$

where $C_{f,d}^{(1)} := C_{f,d} + 2(1+2\pi)(C_f + 4(d+2)(1+2\pi))$ and $C_{f,d}^{(2)} := 2C_{f,d} + 16(1+2\pi)C_f$. Applying a discrete Gronwall Lemma (see e.g. [AGS05, Lemma 3.2.4]) we obtain for

$$
\tau \le \tau_o := \frac{1}{2C_{f,d}^{(2)}}, \quad \text{so that} \quad \frac{1}{1-\tau C_{f,d}^{(2)}} \le e^{2\log 2\, C_{f,d}^{(2)}\tau},
$$

the bound

$$H^k + \frac{\tau}{4}\sum_{n=1}^{k} K^k \le \Big(H^0 + C^{(1)}_{f,d}\tau k\Big)\exp\Big(2C^{(1)}_{f,d}k\tau + 2\log 2\, C^{(2)}_{f,d}\tau\Big). \tag{6.39}$$

Choosing $C := 2C^{(1)}_{f,d} + 2\log 2\, C^{(2)}_{f,d}$ we obtain (6.34).

In order to prove (6.36) let us recall that by (1.4)

$$\mathscr{F}^f(\mu) - C_f(1+\mathsf{m}_2^2(\mu)) \le \frac{1}{2}\mathscr{I}_2(\mu|\,\mathcal{L}^d) \le \mathscr{F}^f(\mu) + C_f(1+\mathsf{m}_2^2(\mu)) \quad \forall\,\mu\in\mathcal{P}_2(\Omega). \tag{6.40}$$

Therefore

$$\begin{aligned}\Big(\mathscr{F}^f(\overline{M}_\tau)_+\Big)^2 &\le \frac{1}{2}\Big(\mathscr{I}_2(\overline{M}_\tau|\,\mathcal{L}^d)\Big)^2 + 4C_f^2\Big(1+\mathsf{m}_2^4(\overline{M}_\tau)\Big)\\ &\overset{(5.69)}{\le} (d+2)\mathscr{K}_{-1}(\overline{M}_\tau|\,\mathcal{L}^d) + 4C_f^2\Big(1+\mathsf{m}_2^4(\overline{M}_\tau)\Big).\end{aligned}$$

Being the map $t\mapsto \mathscr{F}^f(\overline{M}_{\tau,t})$ nonincreasing by (6.26), we deduce

$$\begin{aligned}\bar t_\tau\Big(\mathscr{F}^f(\overline{M}_\tau)_+\Big)^2 &\le \int_0^{\bar t_\tau}\Big(\mathscr{F}^f(\overline{M}_{\tau,t})_+\Big)^2\,\mathrm{d}t\\ &\overset{(6.34)}{\le} 4(d+2)C(M^0_\tau,t) + 4C_f^2\bar t_\tau\Big(1+4C^2(M^0_\tau,t)\Big).\end{aligned}$$

so that

$$\mathscr{F}^f(\overline{M}_\tau)_+ \le \frac{2\sqrt{d+2}}{\sqrt{\bar t_\tau}}C^{1/2}(M^0_\tau,t) + 2C_f\big(1+2C(M^0_\tau,t)\big). \tag{6.41}$$

A further application of (6.40) and (6.34) yields (6.36). □

We can now conclude the proof of Theorem 4: the crucial point here is that we know an a priori bound of the Relative Fisher information (and of the quadratic moment) of the discrete family $\overline{M}_\tau$ on each interval $(\varepsilon,+\infty)$, $\varepsilon>0$, given by (6.36) (and by (6.34)). Therefore, choosing a decreasing vanishing sequence of elapsed initial times ε_h, $h\in\mathbb{N}$, we can apply Theorem 1 in each interval $(\varepsilon_h,+\infty)$, starting from the approximating family $\overline{M}_{\tau,\varepsilon_h}$. By a standard diagonal argument, we can then extract a convergent subsequence in $(0,+\infty)$. Passing to the limit in (6.34) and (6.36) along a suitable subsequence τ_k as $k\to\infty$ we obtain (1.93) and (1.94) respectively.

All the "integral type" properties which involve integrals in $(0,T)$ (except for (1.80)), i.e. (1.77) and (1.78), follow from (6.34) and (1.93).

6.5. **Proof of Theorem 5 (Asymptotic decay when $f\equiv 0$).** When $f\equiv 0$, setting $\gamma = Z^{-1}\mathcal{L}^d$ so that

$$\mathscr{H}(\mu|\,\gamma) \overset{(2.22)}{=} \mathscr{H}(\mu|\,\mathcal{L}^d) + \log Z,\quad \mathscr{I}_2(\mu|\,\gamma) = \mathscr{I}_2(\mu|\,\mathcal{L}^d),\quad \mathscr{K}_{-1}(\mu|\,\gamma) = \mathscr{K}_{-1}(\mu|\,\mathcal{L}^d),$$

(1.87) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}(\mu_t|\,\gamma) \le -\mathscr{K}_{-1}(\mu_t|\,\gamma) \overset{(1.82)}{\le} -\frac{12}{2+d}\int_\Omega \big|\mathrm{D}^2 r_t\big|^2\,\mathrm{d}x, \tag{6.42}$$

where $\mu_t = r_t^2\mathcal{L}^d$. On the other hand, being Ω bounded and $\partial_{\boldsymbol{n}} r_t\equiv 0$ on $\partial\Omega$, (1.95) yields

$$\int_\Omega \big|\mathrm{D}^2 r_t\big|^2\,\mathrm{d}x \ge \alpha_\Omega\int_\Omega \big|\mathrm{D}r_t\big|^2\,\mathrm{d}x = \frac{\alpha_\Omega}{4}\mathscr{I}_2(\mu_t|\,\gamma).$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{H}(\mu_t|\,\gamma) \le -\frac{3\,\alpha_\Omega}{2+d}\mathscr{I}_2(\mu_t|\,\gamma) \overset{(1.96)}{\le} -\frac{3\,\alpha_\Omega\,\beta_\Omega}{2+d}\mathscr{H}(\mu_t|\,\gamma), \tag{6.43}$$

so that

$$\mathscr{H}(\mu_t|\,\gamma) \le \mathscr{H}(\mu_0|\,\gamma)e^{-\beta\,t},\quad \beta := \frac{3\,\alpha_\Omega\,\beta_\Omega}{2+d}.$$

On the other hand, (1.87) and (6.43) yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathscr{I}_2(\mu_t|\,\gamma)\Big)^2 \le -4\Big(\frac{3\alpha_\Omega}{d+2}\Big)^2\Big(\mathscr{I}_2(\mu_t|\,\gamma)\Big)^2,$$

and therefore

$$\mathscr{I}_2(\mu_t|\,\gamma) \le \mathscr{I}_2(\mu_0|\,\gamma)e^{-\alpha t} \quad \text{with} \quad \alpha := 2\Big(\frac{3\alpha_\Omega}{d+2}\Big)^2.$$

In order to prove (1.100) when Ω is not bounded, we combine (1.87) (for $\gamma := \mathscr{L}^d$) and (5.69), obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathscr{I}_2(\mu_t|\,\mathscr{L}^d)\Big)^2 \le -4\Big(\mathscr{K}_{-1}(\mu_t|\,\mathscr{L}^d)\Big)^2 \le -(2+d)^{-2}\Big(\mathscr{I}_2(\mu_t|\,\mathscr{L}^d)\Big)^4, \tag{6.44}$$

and therefore

$$\mathscr{I}_2(\mu_t|\,\mathscr{L}^d) \le \mathscr{I}_2(\mu_0|\,\mathscr{L}^d)\frac{1}{\Big(1+(d+2)^{-2}\mathscr{I}_2(\mu_0|\,\mathscr{L}^d)^2\,t\Big)^{1/2}} \le \frac{d+2}{\sqrt{t}}.$$

7. Discrete entropy estimates and proofs of the main results

7.1. **Discrete estimates and entropy inequalities.** In order to prove the last two theorems, we need other *a priori* estimates on the discrete solutions of the Minimizing Movement Scheme. Let us keep the notation of section 2.5 for the discrete solution $\{M_\tau^n\}_{n\in\mathbb{N}}$ of the variational scheme (1.44) related to the Fisher information functional $\phi(\cdot) = \mathscr{G}(\cdot) := \frac{1}{2}\mathscr{I}_2(\cdot|\,\gamma)$. In particular M_τ^n satisfies

$$\frac{W^2(M_\tau^n, M_\tau^{n-1})}{2\tau} + \frac{1}{2}\mathscr{I}_2(M_\tau^n|\,\gamma) \leq \frac{W^2(\mu, M_\tau^{n-1})}{2\tau} + \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma) \quad \forall\, \mu \in \mathcal{P}_2(\Omega), \tag{7.1}$$

and $\boldsymbol{r}_\tau^n, \boldsymbol{v}_\tau^n, v_\tau^n, \boldsymbol{\eta}_\tau^n$ are defined according to

$$\boldsymbol{r}_\tau^n = T_o(M_\tau^n, M_\tau^{n-1}), \quad -\boldsymbol{v}_\tau^n := \frac{\boldsymbol{r}_\tau^n - \boldsymbol{i}}{\tau} \in \partial_s\mathscr{G}(M_\tau^n) \quad v_\tau^n := \frac{dM_\tau^n}{d\gamma}, \quad \boldsymbol{\eta}_\tau^n := \frac{\mathrm{D}v_\tau^n}{v_\tau^n} \tag{7.2}$$

$\overline{M}_\tau, \overline{\boldsymbol{v}}_\tau$ are the corresponding piecewise constant functions

$$\overline{M}_{\tau,t} := M_\tau^n, \quad \overline{\boldsymbol{v}}_\tau := \boldsymbol{v}_\tau^n, \quad \bar{t}_\tau := n\tau \qquad \text{if} \quad (n-1)\tau < t \leq n\tau. \tag{7.3}$$

$\boldsymbol{v}_\tau^n, \boldsymbol{\eta}_\tau^n$ are vector fields in $L^2_{M_\tau^n}(\Omega;\mathbb{R}^d)$ and we also set

$$|||\boldsymbol{v}_\tau^n|||^2 := \int_\Omega |\boldsymbol{v}_\tau^n|^2 \,\mathrm{d}M_\tau^n \overset{(7.2)}{=} \frac{W^2(M_\tau^n, M_\tau^{n-1})}{\tau^2}, \quad |||\boldsymbol{\eta}_\tau^n|||^2 := \int_\Omega |\boldsymbol{\eta}_\tau^n|^2 \,\mathrm{d}M_\tau^n = \mathscr{I}_2(M_\tau^n|\,\gamma). \tag{7.4}$$

Proposition 7.1 (**Discrete estimates**). *For each couple of integer* $0 \leq m < k$ *we have*

$$\mathscr{I}_2(M_\tau^k|\,\gamma) + \tau \sum_{n=m+1}^{k} |||\boldsymbol{v}_\tau^n|||^2 \leq \inf_{\mu\in\mathcal{P}_2^r(\Omega)} \frac{1}{\tau} W^2(\mu, M_\tau^m) + \mathscr{I}_2(\mu|\,\gamma) \leq \mathscr{I}_2(M_\tau^m|\,\gamma), \tag{7.5}$$

$$\mathscr{H}(M_\tau^k|\,\gamma) + \tau \sum_{n=m+1}^{k} \Big(\mathscr{P}(M_\tau^n|\,\gamma) + \frac{\lambda\tau}{2}|||\boldsymbol{v}_\tau^n|||^2\Big) \leq \mathscr{H}(M_\tau^m|\,\gamma), \tag{7.6}$$

$$\Big|\mathscr{H}(M_\tau^k|\,\gamma) - \mathscr{H}(M_\tau^m|\,\gamma)\Big| \leq \tau \sum_{n=m+1}^{k} |||\boldsymbol{v}_\tau^n|||\Big(\mathscr{I}_2(M_\tau^{n-1}|\,\gamma)\Big)^{1/2} - \frac{\lambda}{2}\tau^2 \sum_{n=m+1}^{k} |||\boldsymbol{v}_\tau^n|||^2, \tag{7.7}$$

$$\mathsf{m}_2(M_\tau^k) \leq \mathsf{m}_2(M_\tau^m) + \tau \sum_{n=m+1}^{k} |||\boldsymbol{v}_\tau^n|||. \tag{7.8}$$

Proof. The first inequality of (7.5) follows simply by choosing $\mu := M_\tau^{n-1}$ in (7.1) and summing up from $n = m+2$ to $n = k$, owing (7.4).

In order to get (7.6), we apply the convexity inequality (2.113) with $\mu := M_\tau^n$, $\nu := M_\tau^{n-1}$, obtaining

$$\mathscr{H}(M_\tau^n|\,\gamma) - \mathscr{H}(M_\tau^{n-1}|\,\gamma) \leq \tau \int_\Omega \Big(\boldsymbol{v}_\tau^n \cdot \boldsymbol{\eta}_\tau^n - \frac{\lambda}{2\tau}|x - \boldsymbol{r}_\tau^n|^2\Big) \,\mathrm{d}M_\tau^n. \tag{7.9}$$

Since, by (7.2), $-\boldsymbol{v}_\tau^n = \partial_s\mathscr{G}(M_\tau^n)$, applying (5.38) of Theorem 5.5 we get

$$\int_\Omega \boldsymbol{v}_\tau^n \cdot \boldsymbol{\eta}_\tau^n \,\mathrm{d}M_\tau^n \leq -\mathscr{P}(M_\tau^n|\,\gamma). \tag{7.10}$$

Inserting (7.10) in (7.9) and summing it from $n := m+1$ to k we obtain (7.6). (7.7) follows similarly: first we apply Schwarz inequality to (7.9) getting (recall the notation (7.4))

$$\mathscr{H}(M_\tau^n|\,\gamma) - \mathscr{H}(M_\tau^{n-1}|\,\gamma) \leq \tau|||\boldsymbol{v}_\tau^n|||\;|||\boldsymbol{\eta}_\tau^n||| - \frac{\lambda\tau^2}{2}|||\boldsymbol{v}_\tau^n|||^2. \tag{7.11}$$

By the same argument, the convexity inequality (2.113) with $\mu := M_\tau^{n-1}$ and $\nu := M_\tau^n$ also yields

$$\mathscr{H}(M_\tau^n|\,\gamma) - \mathscr{H}(M_\tau^{n-1}|\,\gamma) \geq -\tau|||\boldsymbol{v}_\tau^n||||||\boldsymbol{\eta}_\tau^{n-1}||| + \frac{\lambda\tau^2}{2}|||\boldsymbol{v}_\tau^n|||^2. \tag{7.12}$$

Recalling that $|||\boldsymbol{\eta}_\tau^n|||^2 = \mathscr{I}_2(M_\tau^n|\,\gamma) \le \mathscr{I}_2(M_\tau^{n-1}|\,\gamma) = |||\boldsymbol{\eta}_\tau^{n-1}|||^2$, (7.11) and (7.12) yield

$$\Big|\mathscr{H}(M_\tau^n|\,\gamma) - \mathscr{H}(M_\tau^{n-1}|\,\gamma)\Big| \le \tau|||\boldsymbol{v}_\tau^n|||\,|||\boldsymbol{\eta}_\tau^{n-1}||| - \frac{\lambda}{2}\tau^2|||\boldsymbol{v}_\tau^n|||^2. \tag{7.13}$$

Summing up form $n = m+1$ to $n = k$ we obtain (7.7).

(7.8) follows easily by summing form $n = m+1$ to $n = k$ the triangular inequality

$$\mathsf{m}_2(M_\tau^n) \le \mathsf{m}_2(M_\tau^{n-1}) + W(M_\tau^n, M_\tau^{n-1}) \overset{(7.4)}{=} \mathsf{m}_2(M_\tau^{n-1}) + \tau|||\boldsymbol{v}_\tau^n|||. \qquad \square$$

It is useful to rewrite the statements of the previous proposition in terms of the piecewise constant interpolants (7.3):

Corollary 7.2. *For every* $0 < s \le t$ *the piecewise constant interpolants satisfy*

$$\mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma) + \int_{\overline{s}_\tau}^{\overline{t}_\tau} |||\overline{\boldsymbol{v}}_{\tau,r}|||^2\,\mathrm{d}r \le \mathscr{I}_2(\overline{M}_{\tau,s}|\,\gamma) \le \mathscr{I}_2(M_\tau^0|\,\gamma), \tag{7.14}$$

$$\mathscr{H}(\overline{M}_{\tau,t}|\,\gamma) + \int_{\overline{s}_\tau}^{\overline{t}_\tau} \mathscr{P}(\overline{M}_{\tau,r}|\,\gamma)\,\mathrm{d}r + \frac{\lambda\tau}{2}\int_{\overline{s}_\tau}^{\overline{t}_\tau} |||\overline{\boldsymbol{v}}_{\tau,r}|||^2\,\mathrm{d}r \le \mathscr{H}(\overline{M}_{\tau,s}|\,\gamma), \tag{7.15}$$

$$\big|\mathscr{H}(\overline{M}_{\tau,t}|\,\gamma) - \mathscr{H}(\overline{M}_{\tau,s}|\,\gamma)\big| \le \int_{\overline{s}_\tau}^{\overline{t}_\tau} |||\overline{\boldsymbol{v}}_{\tau,r}|||\Big(\mathscr{I}_2(\overline{M}_{\tau,r-\tau}|\,\gamma)\Big)^{1/2}\,\mathrm{d}r - \frac{\lambda\tau}{2}\int_{\overline{s}_\tau}^{\overline{t}_\tau} |||\overline{\boldsymbol{v}}_{\tau,r}|||^2\,\mathrm{d}r. \tag{7.16}$$

$$\mathsf{m}_2(\overline{M}_{\tau,t}) \le \mathsf{m}_2(\overline{M}_{\tau,s}) + \int_{\overline{s}_\tau}^{\overline{t}_\tau} |||\overline{\boldsymbol{v}}_{\tau,r}|||\,\mathrm{d}r. \tag{7.17}$$

From the previous relations we easily derive *a priori* bounds on the discrete solutions which provide useful information on their limit points as $\tau \downarrow 0$. We distinguish two cases: in the first one (which corresponds to Theorem 6) we are assuming an upper bound on the initial Fisher information, whereas in the second one (corresponding to Theorem 7) we are only assuming an upper bound on the initial logarithmic entropy *and the strong uniform convexity of* V.

Corollary 7.3 (**A priori estimates:** $\mathscr{I}_2(M_\tau^0|\,\gamma)$ **is bounded.**)**.** *Let us suppose that, as in* (1.102)

$$\sup_{0<\tau<\tau_0} \mathscr{I}_2(M_\tau^0|\,\gamma) = \mathscr{J}_0 < +\infty, \qquad \sup_{0<\tau<\tau_0} \mathsf{m}_2(M_\tau^0) = \mathsf{m}_{2,0} < +\infty, \tag{7.18}$$

and let us set, as in (2.24)*,*

$$c_\vartheta := \log\Big(\int_\Omega e^{-\frac{\vartheta}{2}|x|^2}\,\mathrm{d}\gamma(x)\Big), \qquad \gamma_\vartheta := e^{-\frac{\vartheta}{2}|x|^2 - c_\vartheta}\cdot\gamma \in \mathcal{P}_2^r(\Omega) \quad \forall\,\vartheta > 0, \tag{7.19}$$

where we also allow $\vartheta = 0$ *if* $\gamma \in \mathcal{P}_2(\Omega)$ *(so that* $\gamma_0 = \gamma,\ c_0 = 0$*). Then for every* $\tau \in (0, \tau_0)$ *and* $\vartheta \ge 0$ *we have*

$$\mathscr{H}_0 := \sup_{0<\tau<\tau_0} \mathscr{H}(M_\tau^0|\,\gamma) \le \frac{1}{2}\mathscr{J}_0 + (1-\lambda)\mathsf{m}_{2,0}^2 + C_{\lambda,\vartheta}, \quad C_{\lambda,\vartheta} := (1-\lambda-\vartheta/2)\mathsf{m}_2^2(\gamma_\vartheta) - c_\vartheta, \tag{7.20}$$

$$\sup_{t\ge0} \mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma) \le \mathscr{J}_0, \qquad \int_0^{+\infty} |||\overline{\boldsymbol{v}}_{\tau,t}|||^2\,\mathrm{d}t \le \mathscr{J}_0, \tag{7.21}$$

$$\big|\mathscr{H}(\overline{M}_{\tau,t}|\,\gamma) - \mathscr{H}(\overline{M}_{\tau,s}|\,\gamma)\big| \le \big(\mathscr{I}_0\big)^{1/2}\int_{\overline{s}_\tau}^{\overline{t}_\tau} |||\overline{\boldsymbol{v}}_{\tau,r}|||\,\mathrm{d}r + \frac{1}{2}\lambda^-\tau\mathscr{J}_0, \tag{7.22}$$

$$\mathsf{m}_2(\overline{M}_{\tau,t}) \le \mathsf{m}_{2,0} + \big(\overline{t}_\tau\mathscr{J}_0\big)^{1/2}, \qquad \mathscr{H}(\overline{M}_{\tau,t}|\,\gamma) \le \mathscr{H}_0 + \lambda^-\mathscr{J}_0\big(t+2\tau\big), \tag{7.23}$$

$$\int_0^{\overline{t}_\tau} \Big(\mathscr{K}_{-1}(\overline{M}_{\tau,r}|\,\gamma) + \lambda^+\mathscr{I}_2(\overline{M}_{\tau,r}|\,\gamma)\Big)\,\mathrm{d}r \le \mathscr{H}_0 + \lambda^-\mathscr{J}_0\big(t+2\tau\big) + \vartheta\big(\mathsf{m}_{2,0}^2 + \overline{t}_\tau\mathscr{J}_0\big) + c_\vartheta. \tag{7.24}$$

Proof. (7.20) is a direct consequence of (2.116) which yields

$$\begin{aligned}\mathscr{H}(M_\tau^0|\,\gamma) &\le \frac{1}{2}\mathscr{I}_2(M_\tau^0|\,\gamma) + \frac{1-\lambda}{2}W^2(M_\tau^0,\gamma_\vartheta) + \mathscr{H}(\gamma_\vartheta|\,\gamma)\\ &\le \frac{1}{2}\mathscr{J}_0 + (1-\lambda)\mathsf{m}_2^2(M_\tau^0) + (1-\lambda)\mathsf{m}_2^2(\gamma_\vartheta) - \frac{\vartheta}{2}\mathsf{m}_2^2(\gamma_\vartheta) - c_\vartheta \le \frac{1}{2}\mathscr{J}_0 + (1-\lambda)\mathsf{m}_{2,0}^2 + C_{\lambda,\vartheta}.\end{aligned}$$

(7.21) follows immediately from (7.18) and (7.14). (7.22) is a consequence of Hölder inequality, (7.16), and (7.21).

Concerning the upper bound on the second order moments (7.23), it follows from (7.17) and (7.21). The entropy bound of (7.23) follows from (7.15), since

$$\mathscr{P}(\mu|\,\gamma) = \mathscr{K}_{-1}(\mu|\,\gamma) + \mathscr{I}_2^V(\mu|\,\gamma) \ge \mathscr{K}_{-1}(\mu|\,\gamma) + \lambda\mathscr{I}_2(\mu|\,\gamma) \quad \forall\,\mu\in\mathcal{P}_2^r(\Omega), \tag{7.25}$$

and therefore

$$\mathscr{P}(\overline{M}_{\tau,t}|\,\gamma) \ge -\lambda^-\mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma) \overset{(7.21)}{\ge} -\lambda^-\mathscr{J}_0.$$

(7.24) still follows from (7.15) and (7.25), since

$$\mathscr{H}(\overline{M}_{\tau,t}|\,\gamma) \overset{(2.25)}{\ge} -\frac{\vartheta}{2}\mathsf{m}_2^2(\overline{M}_{\tau,t}) - c_\vartheta \overset{(7.23)}{\ge} -\vartheta\big(\mathsf{m}_{2,0}^2 + \bar{t}_\tau\mathscr{J}_0\big) - c_\vartheta. \qquad \square$$

Corollary 7.4 (A priori estimates: $\mathscr{H}(M_\tau^0|\,\gamma)$ is bounded). *Let us suppose that V is λ-convex for $\lambda > 0$, $\gamma(\Omega) = 1$, and $\sup_{0<\tau<\tau_0}\mathscr{H}(M_\tau^0|\,\gamma) = \mathscr{H}_0 < +\infty$. Then for every time step $\tau\in(0,\tau_0)$*

$$\lambda\bar{t}_\tau\cdot\mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma) \le \mathscr{H}_0 \quad \forall\, t>0, \tag{7.26}$$

and

$$\mathscr{H}(\overline{M}_{\tau,t}|\,\gamma) + \int_0^{\bar{t}_\tau}\mathscr{K}_{-1}(\overline{M}_{\tau,r}|\,\gamma)\,\mathrm{d}r \le \mathscr{H}_0, \quad \frac{\lambda}{4}\mathsf{m}_2^2(\overline{M}_{\tau,t}) \le \mathscr{H}_0 + \frac{\lambda}{2}\mathsf{m}_2^2(\gamma) \quad \forall\, t>0. \tag{7.27}$$

In particular for every $\varepsilon > 0$ we have

$$\sup_{0<\tau<\tau_0,\, t\ge\varepsilon}\mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma) < +\infty, \qquad \sup_{0<\tau<\tau_0,\, t\ge 0}\mathsf{m}_2(\overline{M}_{\tau,t}) < +\infty. \tag{7.28}$$

Proof. Since the map $t\mapsto\mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma)$ is non increasing, (7.26) follows immediately from (7.24) with $\vartheta = 0$.

(7.27) still follows from (7.24); finally, the uniform bound of the second order moments of $\overline{M}_{\tau,t}$ is a consequence of Talagrand's inequality (2.115). $\square$

7.2. **Proof of Theorem 6 (the case $\mathscr{I}_2(\mu_0|\,\gamma) < +\infty$).** Since the relative Fisher information $\mathscr{I}_2(\mu_0|\,\gamma)$ of the initial datum μ_0 is finite, the main points of Theorem 6 are a direct consequence of the abstract result Theorem 2.13, while the statements about the Entropy follows from the discrete estimates of the previous section 7.1. Let us quickly check the various claims.

- **i)** follows from the properties **a)** and **b)** of Theorem 2.13, since $\mathscr{G}(\mu) := \frac{1}{2}\mathscr{I}_2(\mu|\,\gamma)$ satisfies assumptions (2.66a,b).

  Concerning the convergence of the Entropy, it follows from the estimates (7.21) and (7.22) and a modified version of Ascoli-Arzelà Theorem [AGS05, Prop. 3.3.1].

  Up to a further extraction, we can assume that the functions $t\mapsto|\!|\!|\overline{\boldsymbol{v}}_{\tau_k,t}|\!|\!|$ admits a weak limit w in $L^2(0,+\infty)$. Passing to the limit in (7.16) we get

$$|\mathscr{H}_t - \mathscr{H}_s| \le \int_s^t w(r)\big(\mathscr{J}_r\big)^{1/2}\,\mathrm{d}r \tag{7.29}$$

  which shows the absolute continuity of $t\mapsto\mathscr{H}_t$.

  The regularity (1.106) of the square root s of the density of μ follows by the lower semicontinuity of the functional $\mathscr{K}_{-1}(\cdot|\,\gamma)$ of Corollary 3.8, the a priori upper bounds (7.21),(7.24), and the local estimates of Lemma 3.7.

ii) is a consequence of point **c)** of Theorem 2.13 and of Theorem 5.9: we write the distributional formulation of the continuity equation (2.103)

$$\iint_{\Omega_T} \Big(\partial_t \zeta + \boldsymbol{v}\cdot \mathrm{D}\zeta\Big)\,\mathrm{d}\mu_t\,\mathrm{d}t = 0 \quad \forall\,\zeta \in C^\infty_c(\mathbb{R}^d), \quad \partial_{\boldsymbol{n}}\zeta = 0 \text{ on } \partial\Omega, \tag{7.30}$$

and we use the description of $\boldsymbol{v}$ given by (5.74) choosing $\boldsymbol{\zeta} := \mathrm{D}\zeta$.

iii) is simply (2.107)

iv) (1.110) follows immediately by passing to the limit in (7.15) and recalling that

$$\mathscr{P}(\mu|\,\gamma) \ge \mathscr{K}_{-1}(\mu|\,\gamma) + \lambda \mathscr{I}_2(\mu|\,\gamma). \tag{7.31}$$

In order to prove (1.111) we have to use the refined Lyapunov inequality (2.95), obtained by the De Giorgi variational interpolants $\tilde M_{\tau,t}$ (2.91), (2.92), (2.93). We observe that (7.10) and (7.31) yield

$$\lambda \mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma) + \mathscr{K}_{-1}(\overline{M}_{\tau,t}|\,\gamma) \le \Big(\int_\Omega |\overline{\boldsymbol{v}}_{\tau,t}|^2\,\mathrm{d}\overline{M}_{\tau,t}\cdot \mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma)\Big)^{1/2}. \tag{7.32}$$

Analogously, being $-\tilde{\boldsymbol{v}}_{\tau,t} \in \partial_s\mathscr{G}(\tilde M_{\tau,t})$ we obtain

$$\lambda \mathscr{I}_2(\tilde M_{\tau,t}|\,\gamma) + \mathscr{K}_{-1}(\tilde M_{\tau,t}|\,\gamma) \le \mathscr{P}(\tilde M_{\tau,t}|\,\gamma) \le \Big(\int_\Omega |\tilde{\boldsymbol{v}}_{\tau,t}|^2\,\mathrm{d}\tilde M_{\tau,t}\cdot \mathscr{I}_2(\tilde M_{\tau,t}|\,\gamma)\Big)^{1/2}. \tag{7.33}$$

We introduce the quantities

$$\kappa_t := \liminf_{k\uparrow+\infty}\int_\Omega |\overline{\boldsymbol{v}}_{\tau,t}|^2\,\mathrm{d}\overline{M}_{\tau,t}, \qquad \tilde\kappa_t := \liminf_{k\uparrow+\infty}\int_\Omega |\tilde{\boldsymbol{v}}_{\tau,t}|^2\,\mathrm{d}\tilde M_{\tau,t},$$

observing that the lower semicontinuity of $\mathscr{K}_{-1}(\cdot|\,\gamma)$, (7.32), (7.33), (2.100), (2.101) yield

$$\Big(\lambda\,\mathscr{J}_t + \mathscr{K}_{-1}(\mu_t|\,\gamma)\Big)^2 \le \frac12\big(\kappa_t + \tilde\kappa_t\big)\,\mathscr{J}_t, \qquad \text{for a.e. } t > 0. \tag{7.34}$$

On the other hand, passing to the limit in (2.95) and applying Fatou's Lemma we get

$$\frac12\mathscr{J}_t + \frac12\int_s^t \big(\kappa_r + \tilde\kappa_r\big)\,\mathrm{d}r \le \frac12\mathscr{J}_s \quad \forall\, 0 < s < t, \tag{7.35}$$

and therefore

$$-\frac12\frac{d}{dt}\mathscr{J}_t \ge \frac12\big(\kappa_t + \tilde\kappa_t\big) \quad \text{for a.e. } t \in (0,+\infty). \tag{7.36}$$

Combining (7.36) with (7.34) we get (1.111).

7.3. **Proof of Theorem 7 (the case $\mathscr{H}(\mu_0|\,\gamma) < +\infty$).** The proof of Theorem 7 is completely similar to the conclusion of the proof of Theorem 4: thanks to (7.28) of Corollary 7.4, we get an a priori bound of the quadratic moments of of the discrete family $\overline{M}_{\tau,t}$ for ever $t \ge 0$ and of the Relative Fisher information $\mathscr{I}_2(\overline{M}_{\tau,t}|\,\gamma)$ on each interval $(\varepsilon,+\infty)$, $\varepsilon > 0$.

Choosing a decreasing vanishing sequence of elapsed initial times ε_h, $h \in \mathbb{N}$, we can apply Theorem 6 in each interval $(\varepsilon_h,+\infty)$, starting from the approximating family $\overline{M}_{\tau,\varepsilon_h}$. By a standard diagonal argument, we can then extract a convergent subsequence in $(0,+\infty)$.

Passing to the limit in (7.26) along the subsequence τ_k as $k \uparrow +\infty$ we obtain (1.117).

In order to prove (1.119), let us first observe that applying (2.26) for every $\vartheta > 0$ we get

$$\mathscr{H}(\mu|\,\gamma) = \mathscr{H}(\mu|\,\mathcal{L}^d) + \int_\Omega V\,\mathrm{d}\mu \ge \int_\Omega V\,\mathrm{d}\mu - \frac{\vartheta}{2}\mathsf{m}_2^2(\mu) + \frac{d}{2}\log\big(\tfrac{\vartheta}{2\pi}\big) \quad \forall\,\mu \in \mathcal{P}_2^r(\Omega). \tag{7.37}$$

Since we know by the previous Corollary 7.4 that the second moments of $\overline{M}_{\tau,t}$ are uniformly bounded, we obtain choosing $\vartheta = 1$

$$\sup_{\tau,t}\int_\Omega V\,\mathrm{d}\overline{M}_{\tau,t} \le \frac{\lambda+2}{\lambda}\mathscr{H}_0 + \mathsf{m}_2^2(\gamma) + \frac{d}{2}\log(2\pi) < +\infty. \tag{7.38}$$

Since V has a super-quadratic growth, the above uniform bound and the weak convergence of $\overline{M}_{\tau_k,t}$ to μ_t yields convergence w.r.t. the Wasserstein distance in $\mathcal{P}_2(\Omega)$. The uniform upper bound of the Fisher information (7.28) and (2.114) yield (1.119) and conclude the proof.

References

[AFP00] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Clarendon Press, Oxford, 2000.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005. MR MR2129498

[AP93] Antonio Ambrosetti and Giovanni Prodi, *A primer of nonlinear analysis*, Cambridge Studies in Advanced Mathematics, vol. 34, Cambridge University Press, Cambridge, 1993. MR MR1225101 (94f:58016)

[AS06] Luigi Ambrosio and Giuseppe Savaré, *Gradient flows of probability measures*, Handbook of Evolution Equations (III), Elsevier, 2006.

[BB00] Jean-David Benamou and Yann Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393. MR 2000m:65111

[BÉ85] Dominique Bakry and Michel Émery, *Inégalités de Sobolev pour un semi-groupe symétrique*, C. R. Acad. Sci. Paris Sér. I Math. **301** (1985), no. 8, 411–413. MR MR808640 (86k:60141)

[Bla65] Nelson M. Blachman, *The convolution inequality for entropy powers*, IEEE Trans. Information Theory **IT-11** (1965), 267–271. MR MR0188004 (32 #5449)

[BLS94] Pavel M. Bleher, Joel L. Lebowitz, and Eugene R. Speer, *Existence and positivity of solutions of a fourth-order nonlinear PDE describing interface fluctuations*, Comm. Pure Appl. Math. **47** (1994), no. 7, 923–942. MR 95e:35173

[Bog98] Vladimir I. Bogachev, *Gaussian measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998. MR 2000a:60004

[Bre70] Haïm Brezis, *On a characterization of flow-invariant sets*, Comm. Pure Appl. Math. **23** (1970), 261–263. MR MR0257511 (41 #2161)

[Bre83] ———, *Analyse fonctionnelle - Théorie et applications*, Masson, Paris, 1983.

[Bre91] Yann Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417. MR 92d:46088

[But89] Giuseppe Buttazzo, *Semicontinuity, relaxation and integral representation in the calculus of variations*, Pitman Research Notes in Mathematics Series, vol. 207, Longman Scientific & Technical, Harlow, 1989. MR 91c:49002

[Car91] Eric A. Carlen, *Superadditivity of Fisher's information and logarithmic Sobolev inequalities*, J. Funct. Anal. **101** (1991), no. 1, 194–211. MR MR1132315 (92k:94006)

[CCT05] María J. Cáceres, José A. Carrillo, and Giuseppe Toscani, *Long-time behavior for a nonlinear fourth-order parabolic equation*, Trans. Amer. Math. Soc. **357** (2005), no. 3, 1161–1175 (electronic). MR MR2110436 (2005k:35188)

[CL04] Eric A. Carlen and Michael Loss, *Logarithmic sobolev inequalities and spectral gaps*, Recent advances in the theory and applications of mass transport (Providence RI) (Amer. Math. Soc., ed.), Contemp. Math., vol. 353, 2004, pp. 53–60.

[CT02] José A. Carrillo and Giuseppe Toscani, *Long-time asymptotics for strong solutions of the thin film equation*, Comm. Math. Phys. **225** (2002), no. 3, 551–571. MR MR1888873 (2002m:35115)

[De 93] Ennio De Giorgi, *New problems on minimizing movements*, Boundary Value Problems for PDE and Applications (Claudio Baiocchi and Jacques Louis Lions, eds.), Masson, 1993, pp. 81–98.

[DGJ06] Jean Dolbeault, Ivan Gentil, and Ansgar Jüngel, *A nonlinear fourth-order parabolic equation and related logarithmic-sobolev inqualities*, Comm. Math. Sci. (2006).

[DGMT80] Ennio De Giorgi, Antonio Marino, and Mario Tosques, *Problems of evolution in metric spaces and maximal decreasing curve*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **68** (1980), no. 3, 180–187. MR 83m:49052

[DLSS91] Bernard Derrida, Joel L. Lebowitz, Eugene R. Speer, and Herbert Spohn, *Fluctuations of a stationary nonequilibrium interface*, Phys. Rev. Lett. **67** (1991), no. 2, 165–168. MR MR1113639 (92b:82052)

[DPL04a] Giuseppe Da Prato and Alessandra Lunardi, *Elliptic operators with unbounded drift coefficients and Neumann boundary condition*, J. Differential Equations **198** (2004), no. 1, 35–52. MR 2 037 749

[DPL04b] ———, *Elliptic operators with unbounded drift coefficients and Neumann boundary condition*, J. Differential Equations **198** (2004), no. 1, 35–52. MR MR2037749 (2004k:35066)

[GJT06] Maria Pia Gualdani, Ansgar Jüngel, and Giuseppe Toscani, *A nonlinear fourth-order parabolic equation with nonhomogeneous boundary conditions*, SIAM J. Math. Anal. **37** (2006), no. 6, 1761–1779 (electronic). MR MR2213393

[GM96] Wilfrid Gangbo and Robert J. McCann, *The geometry of optimal transportation*, Acta Math. **177** (1996), no. 2, 113–161. MR 98e:49102

[Gri85] Pierre Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, London, 1985.

[Gro76] Leonard Gross, *Logarithmic sobolev inequalities.*, Amer. Jour. Math. **97** (1976), 1061–1073 (English).

[GS64] Casper Goffman and James Serrin, *Sublinear functions of measures and variational integrals*, Duke Math. J. **31** (1964), 159–178. MR 29 #206

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17 (electronic). MR 2000b:35258

[JM06] Ansgar Jüngel and Daniel Matthes, *An algorithmic construction of entropies in higher-order nonlinear PDEs*, Nonlinearity **19** (2006), no. 3, 633–659. MR MR2209292

[JP00] Ansgar Jüngel and René Pinnau, *Global nonnegative solutions of a nonlinear fourth-order parabolic equation for quantum systems*, SIAM J. Math. Anal. **32** (2000), no. 4, 760–777 (electronic). MR 2002j:35153

[JT03] Ansgar Jüngel and Giuseppe Toscani, *Exponential time decay of solutions to a nonlinear fourth-order parabolic equation*, Z. Angew. Math. Phys. **54** (2003), no. 3, 377–386. MR MR2048659 (2005a:35135)

[LT95] Pierre-Louis Lions and Giuseppe Toscani, *A strengthened central limit theorem for smooth densities*, J. Funct. Anal. **129** (1995), no. 1, 148–167. MR 95m:60043

[McC97] Robert J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179. MR 98e:82003

[McK66] H. P. McKean, Jr., *Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas*, Arch. Rational Mech. Anal. **21** (1966), 343–367. MR MR0214112 (35 #4963)

[MV00] Peter A. Markowich and Cédric Villani, *On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis*, Mat. Contemp. **19** (2000), 1–29, VI Workshop on Partial Differential Equations, Part II (Rio de Janeiro, 1999). MR MR1812873 (2002d:82058)

[Nag42] Mitio Nagumo, *Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen*, Proc. Phys.-Math. Soc. Japan (3) **24** (1942), 551–559. MR MR0015180 (7,381e)

[Ott01] Felix Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), no. 1-2, 101–174. MR 2002j:35180

[OV00] Felix Otto and Cédric Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR 2001k:58076

[RR98] Svetlozar T. Rachev and Ludger Rüschendorf, *Mass transportation problems. Vol. I*, Probability and its Applications, Springer-Verlag, New York, 1998, Theory. MR 99k:28006

[SK87] C. S. Smith and M. Knott, *Note on the optimal transportation of distributions*, J. Optim. Theory Appl. **52** (1987), no. 2, 323–329. MR 88d:90076

[Sta59] A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Information and Control **2** (1959), 101–112. MR MR0109101 (21 #7813)

[Tal96] M. Talagrand, *Transportation cost for gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), 587–600.

[Tos97] Giuseppe Toscani, *Sur l'inégalité logarithmique de Sobolev*, C. R. Acad. Sci. Paris Sér. I Math. **324** (1997), no. 6, 689–694. MR MR1447044 (98g:26019)

[Vil03] Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR 2004e:90003

Dipartimento di Matematica "F. Casorati", Università di Pavia. Via Ferrata, 1 – 27100 Pavia, Italy.
E-mail address: `gianazza@imati.cnr.it`
URL: `http://www.imati.cnr.it/~gianazza`

Dipartimento di Matematica "F. Casorati", Università di Pavia. Via Ferrata, 1 – 27100 Pavia, Italy.
E-mail address: `giuseppe.savare@unipv.it`
URL: `http://www.imati.cnr.it/~savare`

Dipartimento di Matematica "F. Casorati", Università di Pavia. Via Ferrata, 1 – 27100 Pavia, Italy.
E-mail address: `giuseppe.toscani@unipv.it`
URL: `http://www-dimat.unipv.it/~toscani`