

OPTIMAL TRAJECTORIES IN L^1 AND UNDER L^1 PENALIZATIONS

ANNETTE DUMAS, FILIPPO SANTAMBROGIO

ABSTRACT. Motivated by a MFG model where the trajectories of the agents are piecewise constants and agents pay for the number of jumps, we study a variational problem for curves of measures where the cost includes the length of the curve measures with the L^1 distance, as well as other, non-autonomous, cost terms arising from congestion effects. We prove several regularity results (first in time, then in space) on the solution, based on suitable approximation and maximum principle techniques. We then use modern algorithms in non-smooth convex optimization in order to obtain a numerical method to simulate such solutions.

1. INTRODUCTION

This work studies a class of variational problems which arises in a variant of one of the most classical Mean Field Game models. This variant could be used as a brick for describing the evolution of the real estate market but the mathematical questions which appear are of independent interest and the present paper only concentrates on the variational problem without addressing the full model.

The theory of Mean Field Games was introduced around 2006 at the same time by Lasry and Lions, [9, 10, 11, 12], and by Caines, Huang and Malhamé, [8], in order to describe the evolution of a population of rational agents, each one choosing (or controlling) a path in a state space, according to some preferences which are affected by the presence of other agents nearby in a way physicists call mean-field effect. The evolution is described through a Nash equilibrium in a game with a continuum of players.

In the simplest MFG models we look at a population of agents moving inside a domain Ω and we suppose that every agent chooses his own trajectory solving a minimization problem of the form

$$\min \int_0^T \left(\frac{|x'(t)|^2}{2} + h(t, \rho_t, x(t)) \right) dt + \Psi(x(T)),$$

with given initial point $x(0)$. The mean-field effect will be modeled through the fact that the running cost h explicitly depends on the density ρ_t of the agents at time t .

The initial density of the players ρ_0 is given. The goal in MFG is to find an evolution $t \mapsto \rho_t$ such that, when agents consider the above optimization problem for such a given ρ , the trajectories they choose globally reconstruct the same density ρ . This is a non-trivial fixed point condition which can be described via a system of PDEs (we refer to [7], for instance). In some cases, and in particular when $h(t, \rho, x) = V(t, x) + g(\rho(x))$ for an increasing function $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ (where, by abuse of notation, ρ also denotes the density of the measure ρ , which has to be found absolutely continuous), fixed point can be found by an overall minimization problem. The function g models in this case the cost for congestion (i.e. agents have a higher cost when passing through congested regions, where the density is larger).

The corresponding variational formulation is the following: we consider all the possible population evolutions, i.e. curves $t \mapsto \rho_t \in \mathcal{P}(\Omega)$ and we minimize the following energy

$$\int_0^T \left(\frac{1}{2} |\dot{\rho}|_{W_2}^2(t) + \mathcal{F}_{V_t}(\rho_t) \right) dt + \int_{\Omega} \Psi d\rho_T,$$

where $\mathcal{F}_V(\rho) := \int V(x)\rho(x) + f(\rho(x))dx$, where f is the anti-derivative of g , i.e. $f'(s) = g(s)$ for $s \in \mathbb{R}^+$ with $f(0) = 0$. We fix by convention $f(s) = +\infty$ for $\rho < 0$. The functional \mathcal{F}_V is set to $+\infty$ if ρ is not absolutely continuous. Here, at each instant of time, we use the functional associated with the time-dependent potential $V_t := V(t, \cdot)$. The minimization problem above is a variant of the well-known dynamic formulation of optimal transport (see [4]) which includes congestion effects (as in [6]). Note in particular that f is convex, as its derivative is the increasing function g , and so is \mathcal{F}_V . The notation $|\dot{\rho}|_{W_2}(t)$ stands for the metric derivative (see [1]) of the curve $t \mapsto \rho_t$ in the space of probability measures, endowed with the Wasserstein distance W_2 (a distance induced by optimal transport, see [17, 19]). The reason for this distance to appear is related to the presence of the quadratic term $|x'(t)|^2$ in the minimization problem solved by each agent. If this term was replaced by $|x'(t)|^p$ we would have $|\dot{\rho}|_{W_p}^p(t)$ instead of $|\dot{\rho}|_{W_2}^2(t)$.

We are now interested in a different individual cost, and more precisely we replace the term $\int_0^T \frac{|x'(t)|^2}{2} dt$, defined on curves $x \in H^1([0, T])$ with the cost $S(x)$, defined for $x \in BV([0, T])$ in the following way: if x is piecewise constant $S(x)$ equals the number of its jumps, while $S(x) = +\infty$ if x is not piecewise constant. We then face discontinuous trajectories, and the individual optimization problem can be considered as an impulse control problem (for which we refer, for instance, to the classical book [5]). The variational formulation of such a new Mean Field Game is based on the observation (see, for instance, [19]) that the Wasserstein distance associated with a transport cost

$$c(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

is equal to the total variation distance between the two measures times a factor $1/2$. Indeed, when transporting with this cost a measure μ onto a measure ν with the same mass, the mass which needs to move is equal to the part of the mass of μ which is not in common with ν , which means half of the mass $\|\mu - \nu\|$. When μ, ν are absolutely continuous, this gives $\frac{1}{2}\|\mu - \nu\|_{L^1}$.

This motivates to study a problem where the term $\int_0^T \frac{1}{2} |\dot{\rho}|_{W_2}^2(t) dt$ is replaced by the length of the curve $t \mapsto \rho_t$ computed according to the L^1 distance (we can restrict to absolutely continuous measures ρ_t because of the functional \mathcal{G}_V). Hence, formally, we obtain

$$\min \int_0^T \left(\frac{1}{2} \|\dot{\rho}\|_{L^1} + \mathcal{F}_{V_t}(\rho_t) \right) dt + \int_{\Omega} \Psi d\rho_T.$$

The term $\int_0^T \|\dot{\rho}\|_{L^1} dt$ should be intended as the length of the curve ρ (its total variation) w.r.t. the L^1 distance.

Because of the BV behavior of the curves, the Dirichlet condition $\rho(0) = \rho_0$, as well as the final penalization on ρ_T , have to be suitably interpreted. Indeed, it is always possible to jump exactly at time $t = 0$ or $t = T$, so that the Dirichlet condition at $t = 0$ can be replaced by a penalization $\frac{1}{2}\|\rho(0^+) - \rho_0\|_{L^1}$ and for the final penalization, we can replace $\int_{\Omega} \Psi d\rho_T$ with $\inf_{\mu \in \mathcal{P}(\Omega)} \|\mu - \rho_T\| + \int \Psi d\mu$. This last quantity can be computed and equals $\int \tilde{\Psi} d\rho_T$, where $\tilde{\Psi} = \min\{\Psi, \inf \Psi + 1\}$. This is perfectly coherent with the individual optimization problem: if agents are allowed to jump at a cost 1, the final cost Ψ is automatically replaced by $\tilde{\Psi}$ as there is no point in paying $\Psi(x)$ whenever $\Psi(x) > \Psi(x') + 1$ for some point x' . Up to subtracting a constant to the final penalization, we can thus suppose that we have $|\Psi| \leq 1/2$.

More generally, we will study in this note the variational problem

$$\min \int_0^T \int_{\Omega} (\lambda |\dot{\rho}| + f(\rho_t) + V_t \rho_t) dt + \int_{\Omega} \Psi d\rho_T$$

with a final cost Ψ satisfying $|\Psi| \leq \lambda$ and an initial condition $\rho(0) = \rho_0$ which can also be replaced by a penalization $\lambda \|\rho(0^+) - \rho_0\|_{L^1}$. A variant will be the infinite-horizon case with a discount factor $r > 0$, i.e. the variational problem

$$\min \int_0^{\infty} \int_{\Omega} e^{-rt} (\lambda |\dot{\rho}| + f(\rho_t) + V_t \rho_t) dt$$

under the same initial condition.

For simplicity, we will only consider the case where the function f is uniformly convex (think at $f(\rho) := \frac{1}{2}\rho^2$). We still establish regularity results in both time and space for the optimal solution (which is unique because of strict convexity), a result which, besides its mathematical interest, has also at least two applications in the MFG which motivates the problem. First, it proves that, despite individual trajectories being discontinuous, the global behavior of the density of agents $\rho(t, x)$ is smooth, coherently with the experience about the evolution of residential areas, and second it provides the necessary mathematical properties on the individual running cost $g(\rho) + V$ so as to rigorously prove that minimizers of the variational problem are indeed equilibria of the game (see [18]).

Then, we will also provide a numerical method to approximate the solutions using convex optimization tools, able to deal with the non-smooth penalization given by the L^1 term.

Since proving regularity in a problem set on BV curves could be challenging, in order to develop the relevant techniques (which will be based on a suitable use of the maximum principle) we will first start from a simpler, yet not-so-standard, case, where curves, instead of being valued in the functional space L^1 , will be simply valued in the euclidean space \mathbb{R}^d . This will be object of Section 2, where we will prove Lipschitz behavior in the open interval $(0, T)$. Some explicit examples will also be analyzed, in particular for $d = 1$, in order to have some cases which could be used as a test for the numerical methods of Section 6.

The analysis of the infinite-dimensional case, valued in L^1 , will be the object of Section 3, and will be performed by means of many approximations. Some of them are common with the Euclidean case, but there is an extra approximation which comes from discretization: the infinite-dimensional problem is indeed approximated by a sequence of finite-dimensional ones, where the domain Ω is divided into small cells and only piecewise constant densities are considered. Note that in this Eulerian discretization (different from a Lagrangian one where one should follow the jumping trajectories of the particles) the problem becomes very similar to the one studied in Section 2 with the only exception that the norm used on the finite-dimensional space is not the Euclidean one (i.e. ℓ^2) but the ℓ^1 norm. This makes things slightly more involved, but, surprisingly, the final Lipschitz regularity result will be expressed anyway in the L^2 norm.

Section 4 contains the modifications of the strategy proof which are needed to handle the infinite-horizon case, both in the finite and infinite dimensional case. Then, Section 5 addresses the problem of space regularity. We will prove that the solution $\rho(t, x)$ shares the same modulus of continuity in x (uniformly in t) of the Dirichlet data and of the time-dependent potential $V(t, x)$, and in this proof, differently from what done in sections 2, 3 and 4, the Dirichlet data will be attacked by approximation but without replacing them with penalizations (the L^1 cost will be approximated by other superlinear costs on the velocity, as it was already done in the other sections as well). Indeed, for the previous results, it was crucial to use the transversality conditions coming from a suitable approximation of these penalizations, while here, on the contrary, the transversality conditions are harder to consider. This is why the first regularity result is presented in the case where the problem is endowed with Dirichlet conditions at both $t = 0$ and $t = T$, which is not the natural framework we will be interested in. In order to consider some instances of the problem which are more interesting for applications, we will consider the infinite time horizon with Dirichlet data at $t = 0$ (in this case we do not have penalization at the end), where the whole analysis can be performed, as well as the finite-horizon case when the final penalization ψ is piecewise constant: in this last case we can prove continuity of ρ on each piece where ψ is constant, which is enough, for instance, to obtain $\rho \in L^\infty$.

Finally, Section 6 uses proximal methods from non-smooth convex optimization to attack in a numerical way all these problems, and some examples will be considered. A periodic and explicit case presented at the end of Section 2 will be solved numerically as an example in order to validate the method.

2. LIPSCHITZ REGULARITY IN THE EUCLIDEAN SETTING

As a starting point for our analysis, we consider here the following easier problem

$$(1) \quad \min \left\{ TV(\gamma; [0, 1]) + \int_0^1 F(t, \gamma(t)) dt + \psi_0(\gamma(0)) + \psi_1(\gamma(1)) : \gamma \in BV([0, 1]; \mathbb{R}^d) \right\}.$$

Here $TV(\gamma; [0, 1])$ denotes the total variation of γ on $[0, 1]$, i.e.

$$TV(\gamma; [0, 1]) := \sup \left\{ \sum_{k=0}^{N-1} |\gamma(t_k) - \gamma(t_{k+1})| : 0 = t_0 < t_1 < \dots < t_N = 1 \right\},$$

a value which also coincides with the total mass of the vector measure γ' .

The functions $\psi_0, \psi_1 : \mathbb{R}^d \rightarrow [0, +\infty]$ are just supposed to be l.s.c. and bounded from below, and among possible choices we mention those which impose Dirichlet boundary conditions, i.e.

$$t = 0, 1 \quad \psi_t(x) = \begin{cases} 0 & \text{if } x = x_t, \\ +\infty & \text{if not.} \end{cases}$$

We stress the fact that functions in BV spaces are not continuous and can have jumps ; even if we consider that BV functions of one variable are defined pointwisely, it is possible to change very easily their value at a point. In particular, Dirichlet boundary conditions have a very particular meaning: a curve which takes the value x_0 at $t = 0$ but immediately jumps at another point at $t = 0^+$ is considered to satisfy the condition $\gamma(0) = x_0$. In particular, it is possible to freely choose a curve γ on $(0, 1)$ and then add a jump to x_0 or x_1 at the boundary in order to satisfy the corresponding Dirichlet boundary condition, of course adding a price $|\gamma(0^+) - x_0|$ or $|\gamma(1^-) - x_1|$ to the total variation. In this way, we could decide to identify the values of γ at $t = 0$ or $t = 1$ with their right or left limits at these points, respectively, and replace Dirichlet boundary conditions with a boundary penalization. This could also be done for more general penalizations ψ_0, ψ_1 , for which it is useful to define the relaxed functions

$$\tilde{\psi}_t(x) := \inf_y |y - x| + \psi_t(y).$$

It is important to observe that the functions $\tilde{\psi}_i$ are automatically 1-Lipschitz continuous, as an inf of Lip_1 functions of the variable x , indexed with the parameter y .

In this way the problem (1) becomes

$$\min \left\{ TV(\gamma; [0, 1]) + \int_0^1 F(t, \gamma(t)) dt + \tilde{\psi}_0(\gamma(0^+)) + \tilde{\psi}_1(\gamma(1^-)) : \gamma \in BV([0, 1]; \mathbb{R}^d) \right\}$$

or, equivalently, we can replace $\tilde{\psi}_0(\gamma(0^+)) + \tilde{\psi}_1(\gamma(1^-))$ with $\tilde{\psi}_0(\gamma(0)) + \tilde{\psi}_1(\gamma(1))$ and impose continuity of γ at $t = 0, 1$.

Lemma 2.1. *Let $L : \mathbb{R}^d \rightarrow \mathbb{R}$ be a smooth and uniformly convex function which is supposed to be radial: $L(v) := \ell(|v|)$ for a convex and increasing function $\ell : \mathbb{R}^+ \rightarrow \mathbb{R}$. Let $F : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a C^2 time-dependent potential satisfying $D_{xx}^2 F(t, x) \geq c_0 \mathbf{I}$ for a certain constant $c_0 > 0$ and $|\partial_t \nabla_x F(t, x)| \leq C_0$, and $\psi_0, \psi_1 : \mathbb{R}^d \rightarrow \mathbb{R}$ two Lipschitz continuous functions. Consider a solution γ of*

$$\min \left\{ \int_0^1 (L(\gamma'(t)) + F(t, \gamma(t))) dt + \psi_0(\gamma(0)) + \psi_1(\gamma(1)) : \gamma \in H^1([0, 1]) \right\}.$$

Then γ is Lipschitz continuous and satisfies $|\gamma'| \leq C$ where C is defined by

$$C := \max \left\{ \frac{C_0}{c_0}, (\ell')^{-1}(\text{Lip} \psi_0), (\ell')^{-1}(\text{Lip} \psi_1) \right\}.$$

Proof. Let us start from the Euler-Lagrange system of the above optimization problem. We have

$$\begin{cases} (\nabla L(\gamma'))' = \nabla_x F(t, \gamma(t)) \\ \nabla L(\gamma'(0)) = \nabla \psi_0(\gamma(0)) \\ \nabla L(\gamma'(1)) = -\nabla \psi_1(\gamma(1)). \end{cases}$$

First we observe that $\gamma \in C^0$ and $F \in C^1$ imply that the right-hand side in the first equation is a continuous function, so that we have $\nabla L(\gamma') \in C^1$. Inverting the injective function ∇L we obtain $\gamma \in C^2$ and, since $F \in C^2$, we obtain $\gamma \in C^3$.

Then, the transversality conditions show $|\nabla L(\gamma'(t))| \leq \text{Lip} \psi_t$ for $t = 0, 1$. Using $|\nabla L(v)| = \ell'(|v|)$ we see $|\gamma'(t)| \leq C$ for $t = 0, 1$.

Let us now consider the maximal value of $|\gamma'(t)|$. This maximum exists on $[0, 1]$ since $\gamma \in C^1$ and if it is attained on the boundary $t = 0, 1$ the desired Lipschitz bound $|\gamma'| \leq C$ is satisfied. We can now suppose that it is attained in $(0, 1)$. Since ℓ' is increasing and non-negative, the maximal points of $|\gamma'|$ and of $|\nabla L(\gamma')|^2$ are the same. We can then write the optimality condition differentiating once and twice in t : we do have

$$\nabla L(\gamma') \cdot (\nabla L(\gamma'))' = 0; \quad \nabla L(\gamma') \cdot (\nabla L(\gamma'))'' + |(\nabla L(\gamma'))'|^2 \leq 0.$$

In the last condition we can ignore the positive term $|(\nabla L(\gamma'))'|^2$ and observe that, since $\nabla L(\gamma')$ and γ' are vectors with the same orientation (we do have $\nabla L(\gamma') = \frac{\ell'(|\gamma'|)}{|\gamma'|} \gamma'$), we have $\gamma' \cdot (\nabla L(\gamma'))'' \leq 0$.

We now differentiate in time the Euler-Lagrange equation and take the scalar product times γ' , and obtain

$$0 \geq \gamma'(t) \cdot (\nabla L(\gamma'(t)))'' = (\nabla_x F(t, \gamma(t)))' \cdot \gamma'(t) = \partial_t \nabla_x F(t, \gamma) \cdot \gamma'(t) + \gamma'(t) \cdot D_{xx}^2 F(t, \gamma(t)) \gamma'(t).$$

We deduce

$$c_0 |\gamma'(t)|^2 \leq |\gamma'(t)| |\partial_t \nabla_x F(t, \gamma)|,$$

which implies $|\gamma'(t)| \leq \frac{C_0}{c_0} \leq C$ and concludes the proof. \square

We now use the above result on an approximation of the original problem in BV.

Proposition 2.2. *Consider*

$$(2) \quad \min \left\{ TV(\gamma; [0, 1]) + \int_0^1 F(t, \gamma(t)) dt + \psi_0(\gamma(0^+)) + \psi_1(\gamma(1^-)) : \gamma \in BV([0, 1]) \right\}$$

where $F : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a C^2 time-dependent potential satisfying $D_{xx}^2 F(t, x) \geq c_0 \mathbf{I}$ for a certain constant $c_0 > 0$ and $|\partial_t \nabla_x F(t, x)| \leq C_0$, and $\psi_0, \psi_1 \in \text{Lip}_1(\mathbb{R}^d)$ are two penalization functions.

Then a minimizer γ for the above problem exists, is unique, and is actually Lipschitz continuous with $|\gamma'| \leq \frac{C_0}{c_0}$.

Note that we directly state the problem using Lip_1 penalizations instead of first fixing ψ_0 and ψ_1 and then passing to $\tilde{\psi}_0$ and $\tilde{\psi}_1$, but we have already explained why we can restrict to this case. Yet, an important remark is needed:

Remark 2.3. *The solutions with penalizations ψ_t and $\tilde{\psi}_t$ ($t = 0, 1$) coincide in $(0, T)$, but the solution with the original (non-Lip $_1$) penalizations could jump at $t = 0$ or $t = T$, and this jump is intrinsically considered in the definition of $\tilde{\psi}_t$.*

Proof. Given $\varepsilon > 0$, we define $\ell_\varepsilon : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ via $\ell_\varepsilon(s) := \sqrt{\varepsilon^2 + s^2} + \varepsilon h(s)$, where $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a smooth, convex, and increasing function, with $\liminf_{s \rightarrow \infty} h''(s) > 0$. We then define $L_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}$ via $L_\varepsilon(v) = \ell_\varepsilon(|v|)$, so that L_ε is smooth, uniformly convex, and radial¹.

We also choose some numbers $\alpha_\varepsilon < 1$ in such a way that $\lim_{\varepsilon \rightarrow 0} \alpha_\varepsilon = 1$ and $\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{1 - \alpha_\varepsilon^2} = 0$ (for instance $\alpha_\varepsilon = \sqrt{1 - \varepsilon}$).

We consider γ_ε the solution of the variational problem

$$\min \left\{ \int_0^1 (L_\varepsilon(\gamma'(t)) + F(t, \gamma(t))) dt + \alpha_\varepsilon(\psi_0(\gamma(0)) + \psi_1(\gamma(1))) : \gamma \in H^1([0, 1]) \right\}.$$

A solution exists by the direct method of the calculus of variations, and it is unique because of the strict convexity of the function F .

We want to apply Lemma 2.1 to this approximated optimization problem. We first compute $\ell'_\varepsilon(s) = \frac{s}{\sqrt{\varepsilon^2 + s^2}} + \varepsilon h'(s) \geq \frac{s}{\sqrt{\varepsilon^2 + s^2}}$ and observe that we have $(\ell'_\varepsilon)^{-1}(r) \leq \frac{r\varepsilon}{\sqrt{1 - r^2}}$. Since $\text{Lip}(\alpha_\varepsilon \psi_i) = \alpha_\varepsilon \text{Lip}(\psi_i) \leq \alpha_\varepsilon$ we obtain from Lemma 2.1

$$|\gamma'_\varepsilon| \leq \max \left\{ \frac{C_0}{c_0}, \frac{\alpha_\varepsilon \varepsilon}{\sqrt{1 - \alpha_\varepsilon^2}} \right\},$$

and we observe that our choice of α_ε implies that the second term in the max above tends to 0 as $\varepsilon \rightarrow 0$. This means that the Lipschitz constant of γ_ε is at most $\frac{C_0}{c_0}$ if ε is small enough.

By comparing γ_ε with the constant curve $\gamma = 0$ we obtain

$$\int_0^1 F(t, \gamma_\varepsilon(t)) dt + \alpha_\varepsilon(\psi_0(\gamma_\varepsilon(0)) + \psi_1(\gamma_\varepsilon(1))) \leq \int_0^1 F(t, 0) dt + \alpha_\varepsilon(\psi_0(0) + \psi_1(0)) \leq C.$$

This estimate includes an L^2 estimate on γ_ε and, because of the uniform Lipschitz condition on γ_ε , it also implies that the curves γ_ε are equibounded. We can then apply Ascoli Arzelà's theorem to obtain a limit curve $\gamma_\varepsilon \rightarrow \gamma_0$. This curve γ_0 is of course $\frac{C_0}{c_0}$ -Lipschitz continuous, and we can prove that it solves Problem (2).

Indeed, the optimality of γ_ε , together with the inequality $L_\varepsilon(v) \geq |v|$, shows that we have

$$\begin{aligned} TV(\gamma_\varepsilon; [0, 1]) + \int_0^1 F(t, \gamma_\varepsilon(t)) dt + \alpha_\varepsilon(\psi_0(\gamma_\varepsilon(0)) + \psi_1(\gamma_\varepsilon(1))) \\ \leq \int_0^1 L_\varepsilon(\gamma') dt + \int_0^1 F(t, \gamma(t)) dt + \alpha_\varepsilon(\psi_0(\gamma(0)) + \psi_1(\gamma(1))) \end{aligned}$$

for every $\gamma \in H^1$. If we send $\varepsilon \rightarrow 0$ and use the lower semicontinuity of TV for the uniform convergence, we obtain

$$\begin{aligned} TV(\gamma_0; [0, 1]) + \int_0^1 F(t, \gamma_0(t)) dt + \psi_0(\gamma_0(0)) + \psi_1(\gamma_0(1)) \\ \leq \int_0^1 |\gamma'| dt + \int_0^1 F(t, \gamma(t)) dt + \psi_0(\gamma(0)) + \psi_1(\gamma(1)), \end{aligned}$$

where we used the dominated convergence $L_\varepsilon(\gamma') \rightarrow |\gamma'|$ as $\varepsilon \rightarrow 0$, and $\alpha_\varepsilon \rightarrow 1$.

This shows the optimality of γ_0 compared to any H^1 curve. It is now enough to approximate any BV curve with H^1 curves. We take $\gamma \in BV([0, 1])$, we define it as equal to $\gamma(0^+)$ on $[-1, 0]$ and to $\gamma(1^-)$ on $[1, 2]$ and we convolve it with a smooth compactly supported kernel η_δ tending to the identity so as to smooth it, thus obtaining a sequence of curves $\gamma * \eta_\delta$ such that $TV(\gamma * \eta_\delta; [-1, 2]) = \int_{-1}^2 |(\gamma * \eta_\delta)'(t)| dt \leq TV(\gamma, (0, 1))$; moreover, $\gamma * \eta_\delta$ is uniformly bounded and converges to γ at all continuity points of γ , which means that the

¹Note that the easiest choice for h is $h(s) = \frac{1}{2}s^2$, but other choices are possible and reasonable, and the only role of h is to guarantee a lower bound on the Hessian of L_ε (and in particular, to provide a quadratic behavior to ℓ_ε so that the problem is well-posed in H^1) Later one we will see the interest for other choices of h

convergence holds a.e. and at the boundary point. This proves

$$\begin{aligned} \limsup_{\delta \rightarrow 0} \int_0^1 |(\gamma * \eta_\delta)'| dt + \int_0^1 F(t, \gamma * \eta_\delta(t)) dt + \psi_0(\gamma * \eta_\delta(0)) + \psi_1(\gamma * \eta_\delta(1)) \\ \leq TV(\gamma, (0, 1)) + \int_0^1 F(t, \gamma(t)) dt + \psi_0(\gamma(0)) + \psi_1(\gamma(1)) \end{aligned}$$

and concludes the proof of the optimality of γ_0 .

What we proved implies the existence of an optimal curve for (2), and its uniqueness comes, again, from the strict convexity of the term with F . \square

In the case where the target set is \mathbb{R} , we can obtain a very interesting behavior.

Proposition 2.4. *When $d = 1$, i.e. the target space of the curves in Problem (2) is one-dimensional, the minimizer γ satisfies $|\gamma'(t)| |\nabla_x F(t, \gamma(t))| = 0$ a.e., i.e. at each instant of time either γ does not move or it is already located at the optimal point for $F(t, \cdot)$.*

Proof. We consider the same approximation as in Proposition 2.2, using the function $h(s) = (s - M)_+^2$ for a very large M . The uniform Lipschitz bound proven in Lemma 2.1 and Proposition 2.2 makes it irrelevant the choice of $\ell_\varepsilon(s)$ for large values of s , so that we can write the Euler-Lagrange equation for the minimizer γ_ε in the form

$$\frac{\varepsilon^2}{(\varepsilon^2 + |\gamma'_\varepsilon|^2)^{3/2}} \gamma''_\varepsilon = (L'_\varepsilon(\gamma'_\varepsilon))' = \nabla_x F(t, \gamma_\varepsilon),$$

where we explicitly computed² the second derivative of L_ε ignoring the term in h .

We write this as $\varepsilon^2 \gamma''_\varepsilon = (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{3/2} \nabla_x F(t, \gamma_\varepsilon)$. First, note that this implies a uniform bound $\varepsilon^2 |\gamma''_\varepsilon| \leq C$. Then, we differentiate it in time, thus obtaining

$$\varepsilon^2 \gamma'''_\varepsilon = 3(\varepsilon^2 + |\gamma'_\varepsilon|^2)^{1/2} \gamma'_\varepsilon \cdot \gamma''_\varepsilon \nabla_x F(t, \gamma_\varepsilon) + (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{3/2} (\nabla_x F(t, \gamma_\varepsilon))'.$$

Re-using the Euler-Lagrange equation we have

$$\varepsilon^2 \gamma'''_\varepsilon = 3\varepsilon^2 (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{-1} (\gamma'_\varepsilon \cdot \gamma''_\varepsilon) \gamma''_\varepsilon + (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{3/2} (\nabla_x F(t, \gamma_\varepsilon))'.$$

We observe that the last term $(\varepsilon^2 + |\gamma'_\varepsilon|^2)^{3/2} (\nabla_x F(t, \gamma_\varepsilon))'$ is bounded thanks to the Lipschitz bound on γ_ε and the regularity of F . We multiply times γ'_ε and obtain

$$(3) \quad |\varepsilon^2 \gamma'''_\varepsilon \cdot \gamma'_\varepsilon - 3\varepsilon^2 (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{-1} (\gamma'_\varepsilon \cdot \gamma''_\varepsilon)^2| \leq C.$$

We then compute

$$\varepsilon^2 \int_0^1 |\gamma''_\varepsilon|^2 = -\varepsilon^2 \int_0^1 \gamma'''_\varepsilon \cdot \gamma'_\varepsilon + [\varepsilon^2 \gamma'_\varepsilon \cdot \gamma''_\varepsilon]_0^1 \leq C.$$

the last inequality is justified by (3), by the sign of the term $\varepsilon^2 (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{-1} (\gamma'_\varepsilon \cdot \gamma''_\varepsilon)^2$, and by the bound on the boundary term, which is the product of two bounded quantities: γ'_ε and $\varepsilon^2 \gamma''_\varepsilon$.

Coming back to the equality $\varepsilon^2 \gamma''_\varepsilon = (\varepsilon^2 + |\gamma'_\varepsilon|^2)^{3/2} \nabla_x F(t, \gamma_\varepsilon)$ we take the L^2 norms of both sides, thus obtaining

$$\int_0^1 |\gamma'_\varepsilon|^6 |\nabla_x F(t, \gamma_\varepsilon)|^2 \leq \int_0^1 (\varepsilon^2 + |\gamma'_\varepsilon|^2)^3 |\nabla_x F(t, \gamma_\varepsilon)|^2 = \int_0^1 \varepsilon^4 |\gamma''_\varepsilon|^2 \leq C\varepsilon^2.$$

We deduce $\int_0^1 |\gamma'_\varepsilon|^6 |\nabla_x F(t, \gamma_\varepsilon)|^2 dt \rightarrow 0$ and, at the limit, by lower semicontinuity, we have the claim. \square

We will now analyze in details a simple example, both for future use and for better understanding the properties of the minimizers.

We consider the periodic problem

$$(4) \quad \min \left\{ J(\gamma) := \lambda TV(\gamma; \mathbb{S}^1) + \int_{\mathbb{S}^1} \frac{1}{2} |\gamma(t) - \omega(t)|^2 dt : \gamma \in BV(\mathbb{S}^1; \mathbb{R}) \right\},$$

²Note that this computation is based on a 1D cancellation effect, since in higher dimension we have

$$D^2 L_\varepsilon(v) = \frac{(\varepsilon^2 + |v|^2)I - v \otimes v}{(\varepsilon^2 + |v|^2)^{3/2}}$$

and the matrices $|v|^2 I$ and $v \otimes v$ do not cancel out.

where $\omega : \mathbb{S}^1 \rightarrow \mathbb{R}$ is a fixed curve. We suppose that ω is a Lipschitz continuous function such that there exists a finite decomposition of \mathbb{S}^1 into essentially disjoint intervals $I_k = [a_k, b_k]$ ($k = 1, \dots, 4N$ for $N \geq 1$) such that $a_{k+1} = b_k$ and $b_{4N} = a_1$ and satisfying for each $n = 0, \dots, N-1$, the following conditions

- ω is non-decreasing on I_{4n+1} ;
- $\omega(a_{4n+2}) = \omega(b_{4n+2}) := c_{4n+2}$ and $\omega \geq c_{4n+2}$ on I_{4n+2} ;
- ω is non-increasing on I_{4n+3} ;
- $\omega(a_{4n+4}) = \omega(b_{4n+4}) := c_{4n+4}$ and $\omega \leq c_{4n+4}$ on I_{4n+4} ;
- $\int_{I_{4n+2}} \omega - c_{4n+2} = \int_{I_{4n+4}} c_{4n+4} - \omega = 2\lambda$.

We then define a curve γ and a function z via

$$\begin{aligned} \text{on } I_{4n+1} & \quad \gamma = \omega & \quad \text{and} & \quad z = \lambda, \\ \text{on } I_{4n+2} & \quad \gamma = c_{4n+2} & \quad \text{and} & \quad z(t) = \lambda - \int_{a_{4n+2}}^t (\omega - c_{4n+2}), \\ \text{on } I_{4n+3} & \quad \gamma = \omega & \quad \text{and} & \quad z = -\lambda, \\ \text{on } I_{4n+4} & \quad \gamma = c_{4n+4} & \quad \text{and} & \quad z(t) = -\lambda - \int_{a_{4n+2}}^t (\omega - c_{4n+2}). \end{aligned}$$

We can check that γ and z are Lipschitz continuous functions, that we have $|z| \leq \lambda$, $z' = \omega - \gamma$ and $z(t) \in \partial(\lambda|\cdot|)(\gamma'(t))$. This means that we have, for any curve $\tilde{\gamma}$, the inequality

$$(5) \quad \lambda TV(\tilde{\gamma}; \mathbb{S}^1) \geq \lambda TV(\gamma; \mathbb{S}^1) + \int_{\mathbb{S}^1} (\tilde{\gamma}' - \gamma') \cdot z.$$

We claim that γ is a solution of Problem (4). Indeed, for any other competitor $\tilde{\gamma}$, we have

$$J(\tilde{\gamma}) = \lambda TV(\tilde{\gamma}; \mathbb{S}^1) + \int_{\mathbb{S}^1} \frac{1}{2} |\tilde{\gamma}(t) - \omega(t)|^2 dt \geq J(\gamma) + \int_{\mathbb{S}^1} (\tilde{\gamma}' - \gamma') \cdot z + (\tilde{\gamma} - \gamma) \cdot (\gamma - \omega) = J(\gamma),$$

where the last equality is obtained by integrating by parts and using $z' = \omega - \gamma$. The previous inequality comes from the use of (5) and from expanding the square.

This explicit example confirms the behavior predicted in Proposition 2.4: the optimizer in the scalar case either coincides with the minimal point of $F(t, \cdot)$ (here such a point is equal to $\omega(t)$) or it does not move. We will see in Section 6 a numerical confirmation of this solution, which also validates the numerical method.

3. LIPSCHITZ REGULARITY IN THE L^1 SETTING

In this section, we prove the regularity in time of the density ρ which solves the following problem:

$$(6) \quad \min_{\substack{\rho \in E \\ \rho \geq 0 \\ \forall t \in [0, T], \int_{\Omega} \rho(t, x) dx = 1}} \int_0^T \int_{\Omega} (|\dot{\rho}(t, x)| + V(t, x)\rho(t, x) + f(\rho(t, x))) dx dt + \psi_0(\rho(0)) + \psi_T(\rho(T)) := F(\rho),$$

where $E := \text{BV}([0, T], L^1(\Omega)) \cap L^2([0, T] \times \Omega)$ (for one-variable BV functions valued in a Banach space such as L^1 we refer, for instance, to [13, 14]). The function f will be supposed to be uniformly convex (i.e. $f'' \geq c_0 > 0$), and the time-dependent potential V will be supposed to belong to $\text{Lip}([0, T]; L^2(\Omega))$. The domain Ω is a finite measure set in \mathbb{R}^d , and we assume for simplicity that it has unit volume.

A few words on the penalizations ψ_0 and ψ_T :

We aim to apply the results when ψ_T is of the following form

$$\begin{aligned} \psi_T : L^1(\Omega) & \rightarrow \mathbb{R} \\ \rho & \mapsto \int_{\Omega} \varphi_T(x) \rho(x) dx, \end{aligned}$$

where $\varphi_T : \Omega \rightarrow \mathbb{R}$ is L^∞ and $\|\varphi_T\|_\infty \leq 1$. In this case, ψ_T is *1-Lipschitz* for the norm $\|\cdot\|_{L^1(\Omega)}$, and it is also weakly continuous in $L^1(\Omega)$ (and hence in $L^2(\Omega)$).

When it comes to ψ_0 , we aim to consider the following case:

$$\begin{aligned} \psi_0 : L^1(\Omega) & \rightarrow \mathbb{R} \\ \rho & \mapsto \|\rho - m_0\|_{L^1(\Omega)}, \end{aligned}$$

where $m_0 \in L^1(\Omega)$. In this case as well, ψ_0 is *1-Lipschitz* for the norm $\|\cdot\|_{L^1(\Omega)}$. It is also weakly lower-semicontinuous in $L^1(\Omega)$ and $L^2(\Omega)$.

More generally, our results apply when the penalizations ψ_0 and ψ_1 are of the following form:

$$\psi_t(\rho) := \int_{\Omega} a_t(x, \rho(x)) dx,$$

for two functions a_t which are 1-Lipschitz continuous and convex in the second variable. In this way the functionals ψ_t are both continuous for the strong L^1 convergence (actually, Lip_1) and lower semicontinuous for the weak L^1 convergence. This general framework includes the two examples above.

Finally, if ψ_0 and ψ_T are defined as previously, we will define $\psi_{0,\alpha} := \alpha\psi_0$ and $\psi_{T,\alpha} := \alpha\psi_T$.

Approximations:

As in the previous section, the absolute value $|\dot{\rho}|$ will be approximated by a smoother function L_ϵ which will be specified later. The mass constraint $\int \rho_t = 1$ will be imposed via a penalization method, adding $\frac{(\int \rho - 1)^2}{2\delta}$. The penalizations on the boundary on $[0, T]$, ψ_0 and ψ_T , will also be approximated by multiplying them times $\alpha < 1$. Finally, the positivity constraint will be handled by approximating f with a sequence $f_n : \mathbb{R} \rightarrow \mathbb{R}$ obtained via $f_n(\rho) := f(\rho) + n(\rho_-)^2$ so that negative values are penalized but the uniform convexity of f is preserved since we have $f_n'' \geq c_0$ for the same $c_0 > 0$.

Taking together these approximations, we obtain the following problem:

$$\min_{\rho \in E} \int_0^T \left(\int_{\Omega} L_\epsilon(\dot{\rho}(t, x) + f_n(\rho(t, x)) + V(t, x)\rho(t, x)) dx + \frac{(\int \rho(t, y) dy - 1)^2}{2\delta} \right) dt + \psi_{0,\alpha}(\rho(0)) + \psi_{T,\alpha}(\rho(T)) := F_n(\rho).$$

Note that the approximated functional is written as F_n , meaning that we have fixed a suitable sequence of valued $(\epsilon, \delta, \alpha)$ such that $\epsilon_n \rightarrow 0$, $\delta_n \rightarrow 0$ and $\alpha_n \rightarrow 1$.

In order not to have difficulties with the infinite-dimensional space $L^1(\Omega)$, we will also use a finite-dimensional discretization. This consists in imposing that the functions $\rho(t)$ belong to a finite-dimensional subspace. More precisely, we divide the space Ω into n small areas of volume $\frac{1}{n}$ called A_i^n (we need their diameter to tend to 0) and we take $\rho_n(t) : \Omega \rightarrow \mathbb{R}$ such that it is constant on each area. This means that $\rho_n(t)$ takes at most n different values and its mass is constant equal to $\rho_i(t)$ on each region A_i^n (and its density equals $n\rho_i(t)$). The problem can be considered as a restriction of the previous one to the subset of E composed of densities which are piecewise constant functions (constant on each A_i^n) for every t , or it can be rewritten as follows:

$$\min_{\rho \in BV([0, T]; \mathbb{R}^n)} \int_0^T \left(\left(\sum_{i=1}^n \frac{1}{n} (L_\epsilon(n\rho_i(t)) + f_n(n\rho_i(t))) + \rho_i(t) \int_{A_i^n} V(t, x) dx \right) + \frac{(\sum_j \rho_j(t) - 1)^2}{2\delta} \right) dt + \psi_{0,\alpha}(\rho(0)) + \psi_{T,\alpha}(\rho(T)),$$

where the functionals $\psi_{0,\alpha}$ and $\psi_{T,\alpha}$ can also be written in terms of values $\rho_i(0)$ and $\rho_i(T)$: they are of the form

$$\psi_{0,\alpha}(\rho(0)) := \alpha \sum_i a_{i,0}^n(\rho_i(0)) \quad \text{and} \quad \psi_{T,\alpha}(\rho(T)) := \alpha \sum_i a_{i,T}^n(\rho_i(T))$$

for some Lip_1 functions $a_{i,t}^n$, which are precisely given by

$$a_{i,t}^n(u) := \frac{1}{n} \int_{A_i^n} a_t(x, nu) dx.$$

In the set E , we say that ρ_n converges to ρ in E in the sense of (7), if

$$(7) \quad \exists C \text{ s.t. } \|\rho_n(t)\|_{L^2(\Omega)} \leq C \text{ for every } n \text{ and every } t \text{ and } \rho_n(t) \rightharpoonup \rho(t) \text{ uniformly in } t$$

where the uniform L^2 bound allows to metrize the weak L^2 convergence and the uniform convergence is defined accordingly.

We observe that the convergence in the sense of (7) implies the weak convergence $\rho_n \rightharpoonup \rho$ weakly in $L^2([0, T] \times \Omega)$.

Lemma 3.1. *Let $(\rho_n)_n$ be a sequence converging to $\rho \in E$ in the sense of (7) such that $F_n(\rho_n)$ is bounded. Then, we have $\rho \geq 0$, for a.e. $t \in [0, T]$ we have $\int_{\Omega} \rho(t, x) dx = 1$, and moreover*

$$(8) \quad F(\rho) \leq \liminf_n F_n(\rho_n).$$

Proof. First, we have

$$\int_0^T \int_{\Omega} |\dot{\rho}_n(t, x)| dx dt \leq \int_0^T \int_{\Omega} L_\epsilon(\dot{\rho}_n(t, x)) dx dt \leq F_n(\rho_n) \leq C,$$

so that $\|\dot{\rho}_n\|_{L^1([0,T]\times\Omega)}$ is bounded. By embedding $L^1([0,T]\times\Omega)$ into $M([0,T]\times\Omega)$, the space of Radon measures, there exists a subsequence of $(\dot{\rho}_n)_n$ which converges weakly in $M([0,T]\times\Omega)$ towards a measure which can only be $\dot{\rho}$. The semicontinuity of the mass for the weak convergence provides

$$\int_0^T \int_{\Omega} |\dot{\rho}| \leq \liminf_{n\rightarrow\infty} \|\dot{\rho}_n\|_{L^1([0,T]\times\Omega)} \leq \int_0^T \int_{\Omega} L_{\epsilon}(\dot{\rho}_n(t,x)) dx dt$$

(where, actually, the first integral is to be intended as a mass in the sense of measures, or as the total variation of the curve $t \mapsto \rho(t)$ in $L^1(\Omega)$). Then, using again $\rho_n \xrightarrow{L^2} \rho$, we obtain

$$\int_0^T \int_{\Omega} V(t,x)\rho(t,x) dx dt = \lim_{n\rightarrow\infty} \int_0^T \int_{\Omega} V(t,x)\rho_n(t,x) dx dt.$$

Using $f \leq f_n$ and the convexity of f , which implies the weak lower semicontinuity of $\rho \mapsto \int \int f(\rho(t,x)) dx dt$, we have

$$\int_0^T \int_{\Omega} f(\rho(t,x)) dx dt \leq \liminf_{n\rightarrow\infty} \int_0^T \int_{\Omega} f_n(\rho_n(t,x)) dx dt.$$

Moreover, by the definition of f_n , we also obtain

$$\int_0^T \int_{\Omega} (\rho_-)^2 dx dt \leq \liminf_n \int_0^T \int_{\Omega} ((\rho_n)_-)^2 dx dt = 0$$

since $\int_0^T \int_{\Omega} ((\rho_n)_-)^2 dx dt \leq \frac{C}{n}$. This shows $\rho \geq 0$.

Since for all $t \in [0, T]$, $\rho_n(t) \rightharpoonup \rho(t)$ weakly in $L^2(\Omega)$, we have in particular $\rho_n(0) \rightharpoonup \rho(0)$ and $\rho_n(T) \rightharpoonup \rho(T)$ weakly in $L^2(\Omega)$ and consequently weakly in $L^1(\Omega)$, so by the lower semi-continuity of ψ_0 and ψ_T , we have

$$\psi_0(\rho(0)) \leq \liminf_n \psi_0(\rho_n(0)) \text{ and } \psi_T(\rho(T)) \leq \liminf_n \psi_T(\rho_n(T)).$$

Since $(\|\rho_n(0)\|_{L^2(\Omega)})_n$ and $(\|\rho_n(T)\|_{L^2(\Omega)})_n$ are bounded we also have L^1 bounds and hence

$$\lim_n \alpha_n \psi_0(\rho_n(0)) = \lim_n \psi_0(\rho_n(0)) \text{ and } \lim_n \alpha_n \psi_T(\rho_n(T)) = \lim_n \psi_T(\rho_n(T))$$

Finally, we use the positivity of the term

$$\int_0^T \frac{(\int \rho_n(t,y) dy - 1)^2}{2\delta} dt$$

to obtain

$$F(\rho) \leq \liminf_n F_n(\rho_n),$$

and its boundedness to obtain $\int \rho(t) = \lim \int \rho_n(t) = 1$ for a.e. t . \square

Lemma 3.2. *Suppose that m_0 is such that $\int f(m_0) < +\infty$ (in particular $m_0 \in L^2(\Omega)$). For all $\rho \in E$ there exists a sequence $(\rho_n)_n \subset D = H^1([0, T]; L^2(\Omega))$ which converges to ρ strongly in $L^2_{t,x}$ and which satisfies*

$$(9) \quad \rho_n(0) = m_0;$$

$$(10) \quad \begin{aligned} \|\dot{\rho}_n\|_{L^1([0,T]\times\Omega)} &\leq \|\dot{\rho}\|_{L^1([0,T]\times\Omega)} + \|\rho(0^+) - m_0\|_{L^1(\Omega)} \\ &= \|\dot{\rho}\|_{L^1([0,T]\times\Omega)} + \psi_0(\rho(0^+)) \end{aligned}$$

$$(11) \quad \limsup_n \int_0^T \int_{\Omega} f(\rho_n(t,x)) dx dt \leq \int_0^T \int_{\Omega} f(\rho(t,x)) dx dt$$

$$(12) \quad \lim_{n\rightarrow\infty} \int_0^T \int_{\Omega} \rho_n(t,x)V(t,x) dx dt = \int_0^T \int_{\Omega} \rho(t,x)V(t,x) dx dt$$

$$(13) \quad \lim_n \psi_T(\rho_n(T)) = \psi_T(\rho(T^-))$$

In addition, if for all $t \in [0, T]$, $\int_{\Omega} \rho(t,x) dx = 1$, then for all $n \in \mathbb{N}$, $\int_{\Omega} \rho_n(t,x) dx = 1$.

Proof. Let $\rho \in E$. We denote by η_n a sequence of mollifiers in the t variable (i.e. η_n is a sequence of smooth probability measures on \mathbb{R} such that $\eta_n \rightharpoonup \delta_0$), and we suppose $\text{spt} \eta_n \subset [0, \frac{1}{n}]$.

The function ρ is only defined on the time interval $[0, T]$, but we can extend ρ to a function $\tilde{\rho} : [-1, T] \times \Omega \rightarrow \mathbb{R}$ setting $\tilde{\rho}(t) = m_0$ for all $t < 0$. The total variation in L^1 of this extension is equal to the sum of that of ρ and the L^1 distance between m_0 and $\rho(0^+)$.

We then define ρ_n as the convolution of $\tilde{\rho}$ with the mollifiers η_n in time:

$$\rho_n(t, x) := \int \eta_n(t-s) \tilde{\rho}(s, x) ds$$

With the assumption on the support of η_n this convolution is well-defined on $[0, T]$ even if $\tilde{\rho}$ has not been extended on $]T, +\infty[$. Moreover, we have $\rho_n(0) = m_0$.

Note that we have

$$|\dot{\rho}_n(t, x)| = \left| \int \eta'_n(t-s) \tilde{\rho}(s, x) ds \right| \leq C(n) \int_{t-1/n}^t \tilde{\rho}(s, x) ds$$

so that we have

$$\|\dot{\rho}_n(t)\|_{L^2(\Omega)} \leq C(n) \int_{t-1/n}^t \|\tilde{\rho}(s)\|_{L^2(\Omega)} ds \leq C(n)$$

(we use here the assumption $m_0 \in L^2(\Omega)$). This proves $\rho_n \in H^1([0, T]; L^2(\Omega))$.

The convexity of the total variation easily implies (10). Again by convexity, for any convex function g we have

$$\int_0^T \int_{\Omega} g(\rho_n) dx dt \leq \int_{-1/n}^T \int_{\Omega} g(\tilde{\rho}) \rightarrow \int_0^T \int_{\Omega} g(\rho),$$

where the last limit is valid whenever $\int_{\Omega} g(m_0) < +\infty$. Applying this to $g(\rho) = \rho^2$ proves that ρ_n strongly convergence in $L^2_{t,x}$ to ρ (it provides an L^2 bound, hence a weak limit up to subsequences; this weak limit can be identified as ρ by testing against continuous functions; the limit is actually strong because the L^2 norm converges to that of the limit). This implies, in particular, (12). As for (11), it is enough to use $g = f$.

To prove (13), we use the property of bounded variation functions: ρ admits a left limit at T in L^1 , i.e for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|\rho(t) - \rho(T^-)\|_{L^1} \leq \varepsilon$ for every $t \in [T - \delta, T]$. By convexity this implies, as soon as $\frac{1}{n} < \delta$, $\|\rho_n(T) - \rho(T^-)\| \leq \varepsilon$ and shows $\rho_n(T) \rightarrow \rho(T^-)$ strongly in L^1 . This implies (13).

If we suppose in addition that for all $t \in [0, T]$, $\int_{\Omega} \rho(t, x) dx = 1$ the same will be true for ρ_n by convexity, which concludes the proof of the statement. \square

Theorem 3.3. *Suppose that $\psi_0: L^1(\Omega) \rightarrow \mathbb{R}$ and $\psi_T: L^1(\Omega) \rightarrow \mathbb{R}$ are 1-Lipschitz and weakly lower semicontinuous on $L^1(\Omega)$ and that $V: [0, T] \times \Omega \rightarrow \mathbb{R}$ belongs to $\text{Lip}([0, T]; L^2(\Omega))$. Suppose also that $f: \mathbb{R} \rightarrow \mathbb{R}$ is c_0 -convex, i.e $f'' \geq c_0$ on \mathbb{R} .*

Then, there exists a unique minimizer ρ to the problem

$$(14) \quad \min_{\substack{\rho \in E \\ \forall t \in [0, T], \int_{\Omega} \rho(t, x) dx = 1}} F(\rho).$$

This solution belongs to $\text{Lip}([0, T]; L^2(\Omega))$ and it satisfies

$$(15) \quad \sup_{t \in [0, T]} \int_{\Omega} |\dot{\rho}(t, x)|^2 dx \leq C.$$

where C only depends on V and on c_0 .

Proof. To obtain the regularity (15), we will approximate the problem (14) as already described at the beginning of this section; i.e. solving

$$(16) \quad \min_{\rho \in E_n} \int_0^T \left(\left(\sum_{i=1}^n \frac{1}{n} (L_{\varepsilon}(n\dot{\rho}_i(t)) + f_n(n\rho_i(t))) + n\rho_i(t) \int_{A_i^n} V(t, x) dx \right) + \frac{(\sum_j \rho_j(t) - 1)^2}{2\delta} \right) dt \\ + \psi_{0,\alpha}(\rho(0)) + \psi_{T,\alpha}(\rho(T)),$$

where we denote by E_n the set of piecewise constant (in space) functions $\rho \in E$ such that

$$\forall t \in [0, T], \forall x \in \Omega, \rho(t, x) = \sum_{i=1}^n n\rho_i(t) \mathbb{1}_{A_i^n}(x),$$

where $\rho_i: [0, T] \rightarrow \mathbb{R}$ is a real-valued function. With the definition, the mass of ρ on each A_i^n equals $\rho_i(t)$.

In particular, if we consider the curves $\rho_i(t)$, for each i the curve ρ_i solves

$$(17) \quad \min_{\rho_i \in H^1([0,T];\mathbb{R})} \int_0^T \left(L_\epsilon(n\dot{\rho}_i(t)) + f_n(n\rho_i(t)) + n\rho_i(t)V_i^n(t) + \frac{(\sum_j \rho_j(t) - 1)^2}{2\delta} \right) dt \\ + \alpha a_{i,0}^n(\rho_i(0)) + \alpha a_{i,T}(\rho_i(T))$$

where $V_i^n(t) := \frac{1}{|A_i^n|} \int_{A_i^n} V(t, x) dx = n \int_{A_i^n} V(t, x) dx$. The function $V^n(t)$ defined to be equal to $V_i^n(t)$ in A_i^n satisfies

$$\|V^n(t)\|_{L^2(\Omega)} \leq \|V(t)\|_{L^2(\Omega)}, \quad \|V^n\|_{L^2([0,T] \times \Omega)} \leq \|V\|_{L^2([0,T] \times \Omega)}, \quad \|\partial_t V^n(t)\|_{L^2(\Omega)} \leq \|\partial_t V(t)\|_{L^2(\Omega)},$$

all these inequalities being a consequence of Jensen's inequality.

The proof is divided in several steps.

STEP 1 The minimizers of (16), which is a finite-dimensional variational problem in H^1 , exist by the direct method.

STEP 2 Let ρ_n be a minimizer of F_n for all n . In this step, we bound $\|\dot{\rho}_n(t)\|_{L^2(\Omega)}$ independently of n so that we will be able to pass to the limit $n \rightarrow \infty$. We remind that ρ_n is piecewise constant in space: $\rho_n(t, x) := \sum_{i=1}^n n\rho_{n,i}(t)\mathbb{1}_{A_i^n}(x)$. In the following, we fix n and write $\rho_i(t)$ instead of $\rho_{n,i}(t)$ to enlighten the notation.

The Euler-Lagrange equation of (17) is

$$(18) \quad (L'_\epsilon(n\dot{\rho}_i(t)))' = V_i^n(t) + f'_n(n\rho_i(t)) + \frac{(\sum_j \rho_j(t) - 1)}{n\delta}.$$

Differentiating the equation (18) yields

$$(L'_\epsilon(n\dot{\rho}_i(t)))'' = (V_i^n)'(t) + n\dot{\rho}_i(t)f''_n(n\rho_i(t)) + \frac{\sum_j \dot{\rho}_j(t)}{n\delta}.$$

Multiplying by $\dot{\rho}_i(t)$ and summing over i gives

$$\sum_{i=1}^n \dot{\rho}_i(t)(L'_\epsilon(n\dot{\rho}_i(t)))'' = \sum_{i=1}^n (V_i^n)'(t)\dot{\rho}_i(t) + n(\dot{\rho}_i(t))^2 f''_n(n\rho_i(t)) + \frac{(\sum_j \dot{\rho}_j(t))^2}{n\delta}.$$

Since the term $(\sum_j \dot{\rho}_j(t))^2$ is positive, using $f''_n \geq c_0$ as well, we obtain the inequality

$$(19) \quad \sum_{i=1}^n \dot{\rho}_i(t)(L'_\epsilon(n\dot{\rho}_i(t)))'' \geq \sum_{i=1}^n (V_i^n)'(t)\dot{\rho}_i(t) + c_0 n(\dot{\rho}_i(t))^2.$$

Now, we need to estimate the left-hand side of (19). By expanding the second derivative, we have

$$(20) \quad \sum_{i=1}^n \dot{\rho}_i(t)(L'_\epsilon(n\dot{\rho}_i(t)))'' = \sum_{i=1}^n \dot{\rho}_i(t)(n\ddot{\rho}_i(t)L''_\epsilon(n\dot{\rho}_i(t)))' \\ = \sum_{i=1}^n \dot{\rho}_i(t)((n\ddot{\rho}_i(t))^2 L'''_\epsilon(n\dot{\rho}_i(t)) + n\ddot{\rho}_i(t)L''_\epsilon(n\dot{\rho}_i(t))).$$

Let us define the function $h: \mathbb{R} \rightarrow \mathbb{R}$ such that $h(s) = sL'_\epsilon(s) - L_\epsilon(s)$. Then, h verifies

$$h'(s) = sL''_\epsilon(s) \quad \text{and} \quad h''(s) = L''_\epsilon(s) + sL'''_\epsilon(s).$$

We now consider

$$(21) \quad \max_{t \in [0,T]} \sum_{i=1}^n h(n\dot{\rho}_i(t)).$$

Two cases will be distinguished:

- the maximum is reached on $(0, T)$,
- the maximum is reached on $\{0, T\}$.

(i) Suppose there exists $t_0 \in]0, T[$ such that $\sum_{i=1}^n h(n\dot{\rho}_i(t_0)) = \max_{t \in [0,T]} \sum_{i=1}^n h(n\dot{\rho}_i(t))$.

In particular, we have

$$\begin{aligned} \sum_{i=1}^n n \ddot{\rho}_i(t_0) h'(n \dot{\rho}_i(t_0)) &= \sum_{i=1}^n n^2 \ddot{\rho}_i(t_0) \dot{\rho}_i(t_0) L_\epsilon''(n \dot{\rho}_i(t_0)) = 0 \\ \text{and } \sum_{i=1}^n n \ddot{\rho}_i(t_0) h'(n \dot{\rho}_i(t_0)) + (n \ddot{\rho}_i(t_0))^2 h''(n \dot{\rho}_i(t_0)) &\leq 0, \\ \text{i.e. } \sum_{i=1}^n n^2 \ddot{\rho}_i(t_0) \dot{\rho}_i(t_0) L_\epsilon''(n \dot{\rho}_i(t_0)) + (n \ddot{\rho}_i(t_0))^2 (L_\epsilon''(n \dot{\rho}_i(t_0)) + n \dot{\rho}_i(t_0) L_\epsilon'''(n \dot{\rho}_i(t_0))) &\leq 0. \end{aligned}$$

Inserting (19) and (20) in t_0 , the last inequality becomes

$$(22) \quad \sum_{i=1}^n ((V_i^n)'(t_0) \dot{\rho}_i(t_0) + c_0 n \dot{\rho}_i(t_0)^2) \leq - \sum_{i=1}^n (n \ddot{\rho}_i(t_0))^2 L_\epsilon''(n \dot{\rho}_i(t_0)) \leq 0.$$

Let us precise the expression of h :

$$\begin{aligned} h(s) &= s L_\epsilon'(s) - L_\epsilon(s) \\ &= \frac{s^2}{\sqrt{s^2 + \epsilon^2}} + 2\epsilon s^2 - \sqrt{s^2 + \epsilon^2} - \epsilon s^2 = \frac{s^2 - s^2 - \epsilon^2}{\sqrt{s^2 + \epsilon^2}} + \epsilon s^2 \\ &= -\frac{\epsilon^2}{\sqrt{s^2 + \epsilon^2}} + \epsilon s^2 < \epsilon s^2. \end{aligned}$$

Since t_0 is a maximizer of $\sum_{i=1}^n h(n \dot{\rho}_i)$, for all $t \in [0, T]$, we obtain

$$(23) \quad \begin{aligned} \sum_{i=1}^n h(n \dot{\rho}_i(t)) &= \sum_{i=1}^n -\frac{\epsilon^2}{\sqrt{(n \dot{\rho}_i(t))^2 + \epsilon^2}} + \epsilon (n \dot{\rho}_i(t))^2 \\ &\leq \sum_{i=1}^n h(n \dot{\rho}_i(t_0)) < \epsilon \sum_{i=1}^n (n \dot{\rho}_i(t_0))^2. \end{aligned}$$

In particular, we have

$$\sum_{i=1}^n -\frac{\epsilon}{\sqrt{(n \dot{\rho}_i(t))^2 + \epsilon^2}} + (n \dot{\rho}_i(t))^2 < \sum_{i=1}^n (n \dot{\rho}_i(t_0))^2.$$

Since $\frac{\epsilon}{\sqrt{(n \dot{\rho}_i(t))^2 + \epsilon^2}} \leq 1$, for all $t \in [0, T]$,

$$(24) \quad \sum_{i=1}^n n \dot{\rho}_i(t)^2 < 1 + \sum_{i=1}^n n \dot{\rho}_i(t_0)^2.$$

Besides, thanks to Cauchy-Schwarz inequality, the inequality (22) becomes

$$(25) \quad \begin{aligned} c_0 \sum_{i=1}^n n \dot{\rho}_i(t_0)^2 &\leq - \sum_{i=1}^n (V_i^n)'(t_0) \dot{\rho}_i(t_0) \\ &\leq \sqrt{\sum_{i=1}^n \frac{1}{n} ((V_i^n)'(t_0))^2} \sqrt{\sum_{i=1}^n n \dot{\rho}_i(t_0)^2} \\ &\leq \sqrt{\sup_{t \in [0, T]} \sum_{i=1}^n \int_{A_i^n} (V^n)'(t)^2 dx} \sqrt{\sum_{i=1}^n n \dot{\rho}_i(t_0)^2} \\ &\leq \sup_{t \in [0, T]} \|(V^n)'(t, \cdot)\|_{L^2(\Omega)} \sqrt{\sum_{i=1}^n n \dot{\rho}_i(t_0)^2} \\ &\leq \sup_{t \in [0, T]} \|V'(t, \cdot)\|_{L^2(\Omega)} \sqrt{\sum_{i=1}^n n \dot{\rho}_i(t_0)^2}. \end{aligned}$$

Finally, (25) gives

$$(26) \quad c_0^2 \sum_{i=1}^n n \dot{\rho}_i(t_0)^2 \leq \sup_{t \in [0, T]} \|V'(t, \cdot)\|_{L^2(\Omega)}^2.$$

By gathering the inequalities (24) and (26), there exists a constant $C_0^2 := \sup_{t \in [0, T]} \|V'(t, \cdot)\|_{L^2(\Omega)}^2 \geq 0$ independent from n such that for all $t \in [0, T]$,

$$(27) \quad \sum_{i=1}^n n \dot{\rho}_i(t)^2 \leq 1 + \sum_{i=1}^n n \dot{\rho}_i(t_0)^2 \leq 1 + \frac{C_0^2}{c_0^2}.$$

Yet, for all $t \in [0, T]$, the L^2 -norm of $\dot{\rho}_n(t)$ is

$$(28) \quad \|\dot{\rho}_n(t)\|_{L^2(\Omega)}^2 = \sum_{i=1}^n \int_{A_i^n} (n \dot{\rho}_i(t))^2 = \sum_{i=1}^n n \dot{\rho}_i(t)^2,$$

so, for all $n \in \mathbb{N}$, a minimizer of (17) verifies

$$(29) \quad \sup_{t \in [0, T]} \|\dot{\rho}_n(t)\|_{L^2(\Omega)}^2 \leq 1 + \frac{C_0^2}{c_0^2}.$$

(ii) Suppose that the maximum of (21) is reached at 0, i.e $\sum_{i=1}^n h(n \dot{\rho}_i(0)) = \max_{t \in [0, T]} \sum_{i=1}^n h(n \dot{\rho}_i(t))$. The transversality condition yields

$$L'_\epsilon(n \dot{\rho}_i(0)) = \alpha a_{i,0}^n (\rho_i(0)),$$

and we know that $\alpha a_{i,0}^n$ is Lip_α . Using $L'_\epsilon(s) = \frac{s}{\sqrt{s^2 + \epsilon^2}} + 2\epsilon s$, we get

$$\left| \frac{n \dot{\rho}_i(0)}{\sqrt{(n \dot{\rho}_i(0))^2 + \epsilon^2}} + 2\epsilon n \dot{\rho}_i(0) \right| \leq \alpha.$$

The same computations as in Section 2 lead to

$$|n \dot{\rho}_i(0)| \leq \frac{\alpha \epsilon}{1 - \alpha}.$$

Thanks to this inequality, we can conclude similarly to (27) that

$$\begin{aligned} \sum_{i=1}^n n \dot{\rho}_i(t)^2 &< 1 + \sum_{i=1}^n n \dot{\rho}_i(0)^2, \\ &\leq 1 + \frac{n^2 \alpha^2 \epsilon^2}{n^2(1 - \alpha^2)} = 1 + \frac{\alpha^2 \epsilon^2}{1 - \alpha^2}, \end{aligned}$$

so, if $\alpha \rightarrow 1$ and $\epsilon \rightarrow 0$ in a way that $\frac{\alpha^2 \epsilon^2}{1 - \alpha^2}$ remains below 1, we have

$$(30) \quad \sup_{t \in [0, T]} \|\dot{\rho}_n(t)\|_{L^2(\Omega)}^2 = \sup_{t \in [0, T]} \sum_{i=1}^n n \dot{\rho}_i(t)^2 \leq 2.$$

Remark 3.4. *The upper bounds in (29) and (30) can be improved to quantities which are arbitrarily close to C_0^2/c_0^2 and 0, respectively.*

For the first one, it suffices to define a slightly different approximation of $|\cdot|$. If one take $L_{\epsilon, A}(s) = \sqrt{s^2 + \epsilon^2} + A\epsilon s^2$ instead of L_ϵ , the calculus remains the same, except for (23), which becomes

$$\begin{aligned} \sum_{i=1}^n h(n \dot{\rho}_i(t)) &= \sum_{i=1}^n -\frac{\epsilon^2}{\sqrt{(n \dot{\rho}_i(t))^2 + \epsilon^2}} + A\epsilon (n \dot{\rho}_i(t))^2 \\ &\leq \sum_{i=1}^n h(n \dot{\rho}_i(t_0)) < A\epsilon \sum_{i=1}^n (n \dot{\rho}_i(t_0))^2. \end{aligned}$$

and yields the inequality

$$\begin{aligned} A \sum_{i=1}^n n \dot{\rho}_i(t)^2 &< 1 + A \sum_{i=1}^n n \dot{\rho}_i(t_0)^2, \\ \text{i.e } \sum_{i=1}^n n \dot{\rho}_i(t)^2 &< \frac{1}{A} + \sum_{i=1}^n n \dot{\rho}_i(t_0)^2. \end{aligned}$$

Since the constant A can be chosen as large as we want, the bound can be C_0^2/c_0^2 .

For the second estimate, it is enough to choose $\alpha \rightarrow 1$ such that $\frac{\alpha^2 \epsilon^2}{1-\alpha^2} \rightarrow 0$ which allows to replace the second term in (30) by almost 0 (and the first one can be taken small as well, as we have just explained).

The same computations lead to the same conclusion if the maximum were reached on T .

To conclude, by (29) and (30), we have shown that

$$(31) \quad \forall n, \sup_{t \in [0, T]} \|\dot{\rho}_n(t)\|_{L^2(\Omega)}^2 \leq \max \left\{ \frac{C_0^2}{c_0^2} + 1, 2 \right\} := C.$$

STEP 3 In this step, we prove that there exists a subsequence of $(\rho_n)_n$ which converges to a function $\rho \in \mathcal{C}([0, T], L^2(\Omega))$.

The sequence $(\rho_n(t))_n$ is equi-Lipschitz in $L^2(\Omega)$ but the norm $L^2_{t,x}$ is also bounded, which provides a uniform bound $\|\rho_n(t)\|_{L^2(\Omega)} \leq C$. Bounded sets in $L^2(\Omega)$ are not compact for the strong convergence, but they are compact, by Banach-Alaoglu's theorem, for the weak convergence. The equicontinuity that we have is in the strong sense, so it also holds in the weak sense, and we can then apply the Arzelà-Ascoli's theorem to extract a subsequence of $(\rho_n)_n$ which converges to ρ in the sense of (7).

Additionally, we obtain by the lower-semicontinuity property of the L^2 -norm that the limit curve ρ is also Lipschitz in time for the strong L^2 norm:

$$\forall (t, s) \in [0, T]^2, \|\rho(t) - \rho(s)\|_{L^2(\Omega)} \leq C|t - s|.$$

STEP 4 The goal of this step is to prove that the limit ρ found previously is actually a minimizer of (14). Comparing to a suitable competitor (for instance a discretization of m_0 , constant in time), we can see that there exists $C > 0$ such that $F_n(\rho_n) \leq C$ for all n .

First, let us notice that $\int_{\Omega} \rho(t, x) dx = 1$, for all $t \in [0, T]$.

Indeed, we have

$$\int_0^T \left(\int_{\Omega} \rho(t, x) dx - 1 \right)^2 dt \leq \liminf_n \int_0^T \left(\int_{\Omega} \rho_n(t, x) dx - 1 \right)^2 dt = 0,$$

Second, by Lemma 3.1, we have that $F(\rho) \leq \liminf_n F_n(\rho_n)$.

Third, let $m \in D = H^1([0, T]; L^2(\Omega))$ be such that $\int_{\Omega} m(t, y) dy = 1$. In what follows, we prove that there exists a sequence $(m_n)_n$ such that $\limsup_n F_n(m_n) \leq F(m)$.

We define the sequence $(m_n)_n$ by choosing piecewise constant functions such that

$$(32) \quad \forall t \in [0, T], \forall x \in A_i^n, m_n(t, x) = n \int_{A_i^n} m(t, y) dy.$$

which verifies $\int_{\Omega} m_n(t, y) dy = \int_{\Omega} m(t, y) dy = 1$. This lets the term penalizing the mass of m_n disappear in $F_n(m_n)$ and, using $m \geq 0$ and $m_n \geq 0$, the penalization of the negative part disappears as well:

$$F_n(m_n) = \int_0^T \int_{\Omega} L_{\epsilon}(\dot{m}_n(t, x)) + f(m_n(t, x)) + m_n(t, x) V^n(t, x) dx dt + \psi_{0, \alpha}(m_n(0)) + \psi_{T, \alpha}(m_n(T)).$$

We observe that, by Jensen's inequality, exactly as it happens for V^n , the sequence m_n is bounded in L^2 (the norm is bounded by that of m) and clearly weakly converges to m (it is enough to test against continuous test functions). So, it strongly converges to m in L^2 . This is true both in $L^2_{t,x}$ and in L^2_x for every t . In particular, the boundary terms and the linear term $\int m_n V^n$ converge to the corresponding terms with m . As for the term $\int f(m_n)$, it can be bounded (again thanks to Jensen's inequality) by $\int f(m)$. We are left to bound the term involving the time-derivative.

Using again the Jensen inequality we have $\int L_{\epsilon}(\dot{m}_n) \leq \int L_{\epsilon}(\dot{m})$ and using $m \in D$ (i.e. $\dot{m} \in L^2_{t,x}$) this quantity tends to $\|\dot{m}\|_{L^1([0, T] \times \Omega)}$.

Summing up, we have proved that for all $m \in D$ such that $\int_{\Omega} m(t, x) dx = 1$, there exists a sequence $(m_n)_n$ which converges to m strongly in $L^2([0, T] \times \Omega)$ and which verifies

$$(33) \quad \limsup_n F_n(m_n) \leq F(m).$$

Now, we can prove that ρ minimizes F . Let $m \in D$ be a competitor such that $\int_{\Omega} m(t, x) dx = 1$ and $(m_n)_n$ the sequence defined in (32). Since for all $n \in \mathbb{N}$, ρ_n minimizes F_n , we have

$$\forall n \in \mathbb{N}, F_n(\rho_n) \leq F_n(m_n).$$

By Lemma 3.1 and the property (33), we obtain

$$(34) \quad \forall m \in D, F(\rho) \leq F(m).$$

The inequality is true for all $m \in D$, there remains to show that it is true for all $m \in E$.

Let $m \in E$ be such that $\int_{\Omega} m(t, x) dx = 1$. By lemma 3.2, there exists a sequence $(m_n)_n \subset D$ which converges to m strongly in $L^2_{t,x}$ and such that $\int_{\mathbb{R}^N} m_n(t, x) dx = 1$ for all $n \in \mathbb{N}$. This sequence may be different from (32). By Inequality (34), we have

$$\forall n \in \mathbb{N}, F(\rho) \leq F(m_n).$$

By using the properties (10), (11) and (12) and the fact that m is supported in $[0, T] \times \Omega$, we obtain

$$\forall m \in E, F(\rho) \leq F(m),$$

which shows that the limit ρ is a minimizer of F .

To conclude, the function ρ is a solution to the problem (14) and verifies the property (15).

Moreover, the solution ρ is unique by the strict convexity of F , given by f . \square

4. INFINITE HORIZONS PROBLEMS

Similar results can be proven in the infinite horizon case, with an exponential discount factor:

$$\min_{\substack{\rho \in E_{\infty}; \\ \forall t \in [0, +\infty[, \int_{\Omega} \rho(t, x) dx = 1}} \int_0^{+\infty} e^{-rt} \int_{\Omega} (|\dot{\rho}| + V(t, x)\rho + f(\rho)) dx dt + \psi_0(\rho(0)) := F_{\infty}(\rho)$$

where $E_{\infty} := \text{BV}_{loc}([0, +\infty[, L^1(\Omega)) \cap L^2_{loc}([0, +\infty[\times \Omega)$. This kind of infinite horizon problem is very classical in economical models and we will not discuss any more its economical motivations. From the mathematical point of view, it is interesting as we get rid of some difficulties (which will appear in the next section) related to the transversality condition at $t = T$ but the computations which provide the time regularity can be re-done up to minor modifications.

The proof is similar to Theorem 3.3. We will need to use the following approximated functional

$$F_T(\rho) := \int_0^T e^{-rt} \int_{\Omega} (|\dot{\rho}| + V(t, x)\rho + f(\rho)) dx dt + \psi_0(\rho(0))$$

and its approximation

$$F_T^n(\rho) := \int_0^T e^{-rt} \left(\sum_{i=1}^n \frac{1}{n} \left(L_{\epsilon}(n\rho_i(t)) + f_n(n\rho_i(t)) + \frac{(\sum_j \rho_j(t) - 1)^2}{2\delta} \right) + n\rho_i(t) \int_{A_i^n} V(t, x) dx \right) dt + \psi_{0,\alpha}(\rho(0)),$$

where L_{ϵ} and the partition $\Omega = \cup_{i=1}^n A_i^n$ are the same as in the proof of Theorem 3.3.

Theorem 4.1. *Suppose that $\psi_0: L^1(\Omega) \rightarrow \mathbb{R}$ is 1-Lipschitz and weakly lower semicontinuous on $L^1(\Omega)$ and that $V: [0, T] \times \Omega \rightarrow \mathbb{R}$ belongs to $\text{Lip}([0, T]; L^2(\Omega))$. Suppose also that $f: \mathbb{R} \rightarrow \mathbb{R}$ is c_0 -convex, i.e $f'' \geq c_0$.*

Then, there exists a unique solution ρ to the problem

$$(35) \quad \min_{\substack{\rho \in E_{\infty}; \\ \forall t \in [0, +\infty[, \int_{\Omega} \rho(t, x) dx = 1}} F_{\infty}(\rho),$$

and it satisfies

$$(36) \quad \sup_{t \in [0, +\infty[} \int_{\Omega} |\dot{\rho}(t, x)|^2 dx \leq C_{\infty},$$

where $C_{\infty} > 0$ only depends on V (through its Lipschitz constant) and on c_0 .

Proof. The proof will be very similar to the case discussed in Section 3, and we will only highlight the differences. First, there is an extra approximation due to the infinite horizon: we fix T (and, later, we will consider $T \rightarrow \infty$) and we solve

$$(37) \quad \min_{\rho \in E_n} F_n^T(\rho).$$

We adapt the proof of theorem 3.3 and we mention here the main differences.

STEP 1 Using the fact that $e^{-rT} \leq e^{-rt}$, we can conclude similarly to Step 1 in theorem 3.3 that there exists a solution to problem (37).

STEP 2 Let ρ_n (denoted ρ in this step for simplicity) be a minimizer of F_n . We exploit the Euler-Lagrange equation for each component $\rho_i(t)$ in order to obtain bounds. The equation is slightly different now, due to the coefficient e^{-rt} . We have

$$(38) \quad (e^{-rt} L'_\epsilon(\dot{\rho}_i))' = e^{-rt} \left(V_i^n(t, x_i) + f'_n(n\rho_i) + \frac{\sum_j \rho_j - 1}{\delta} \right)$$

$$\text{i.e. } -rL'_\epsilon(\dot{\rho}_i) + (L'_\epsilon(\dot{\rho}_i))' = V_i^n(t, x_i) + f'_n(n\rho_i) + \frac{\sum_j \rho_j - 1}{\delta n}.$$

Similarly to the proof of Theorem 3.3, we differentiate Eq. (38), multiply it by $\dot{\rho}_i$ and sum over i :

$$\sum_i (-r\dot{\rho}_i \dot{\rho}_i L''_\epsilon(\dot{\rho}_i) + \dot{\rho}_i (L'_\epsilon(\dot{\rho}_i))') = \sum_i (\dot{\rho}_i (V_i^n)'(t, x_i) + n\dot{\rho}_i^2 f''_n(n\rho_i)) + \frac{(\sum_i \dot{\rho}_i)^2}{\delta n}.$$

Compared to Section 3, there is an extra term equal to $\sum_i (-r\dot{\rho}_i \dot{\rho}_i L''_\epsilon(\dot{\rho}_i))$. We now take again h the function such that

$$h(s) = sL'_\epsilon(s) - L_\epsilon(s)$$

and look at $\max_{t \in [0, T]} \sum_{i=1}^n h(n\dot{\rho}_i(t))$.

The extra term in the Euler-Lagrange equation appears in considering the maximum in $(0, T)$, but in this case we also have

$$\sum_{i=1}^n h'(n\dot{\rho}_i(t)) \dot{\rho}_i(t) = 0.$$

Using $h'(s) = sL''_\epsilon(s)$ we see that this lets the extra term vanish, and the computations are then unchanged.

As for the case where the maximum is reached at $t = 0$, the transversality condition is exactly the same as in Section 3 because of $e^{-rt} = 1$. When the maximum is reached at $t = T$, the transversality condition now gives $e^{-rT} L'_\epsilon(n\dot{\rho}_i(T)) = 0$ for every i . This allows to obtain $\dot{\rho}_i(T) = 0$ and the maximum cannot be reached at $t = T$.

This allows to bound $\sum_i n\dot{\rho}_i(t)^2$ by a constant which only depends on c_0 and $\sup_{t \in [0, T_n]} \|V'(t, \cdot)\|_{L^2(\Omega)}^2 \leq \sup_{t \in [0, \infty)} \|V'(t, \cdot)\|_{L^2(\Omega)}^2$.

STEP 3 Now, we pass to the limit in n for fixed T . This follows the very same procedure based on the Arzelà-Ascoli's theorem, as in Section 3. In this way, we obtain a family $(\rho_T)_T$ of equilipschitz functions in t valued in $L^2(\Omega)$, satisfying the very same uniform bound on the time derivative, and each ρ_T minimizes F_T by proceeding the same way as in Step 4 of Theorem 3.3.

STEP 4 We pass to the limit $T \rightarrow \infty$. The previous step provides a family $(\rho_T)_{T \in \mathbb{N}}$ of equilipschitz minimizers of F_T . By choosing a density m such that $\int f(m) < +\infty$ and comparing ρ_T to the constant curve equal to m , using the integrability of the exponential discount coefficient e^{-rt} , we obtain a uniform bound $F_T(\rho_T) \leq C$ (independent of T). Let us fix $T_0 < \infty$. Since ρ_T is a minimizer of F_T , we have $F_{T_0}(\rho_T) \leq F_T(\rho_T) \leq C$. Thus we have a bound (depending on T_0 , because of the coefficient e^{-rt})

$$\sup_{t \in [0, T_0]} \|\rho_T(t)\|_{L^2(\Omega)} \leq C_{T_0}.$$

This allows to apply Arzelà-Ascoli's theorem for the convergence in the sense of (7), and extracting a diagonal subsequence we obtain a subsequence of $(\rho_T)_T$ which converges towards $\rho \in \mathcal{C}([0, +\infty[, L^2(\Omega))$ in the sense of (7), on each interval $[0, T_0]$.

STEP 5 In this step, we show that ρ defined previously is a minimizer of F . Let us denote by $(\rho_{T_k})_k$, the subsequence which converges to ρ . Let $m \in E_\infty$ be a competitor and T a fixed value. For all $T_k \geq T$, we have

$$F_T(\rho_{T_k}) \leq F_T(m) \leq F_\infty(m).$$

By the lower semi-continuity of F_T we obtain $F_T(\rho) \leq F_\infty(m)$ and, taking the limit $T \rightarrow \infty$, we see $F_\infty(\rho) \leq F_\infty(m)$.

This shows that ρ is a minimizer of F . Additionally, it verifies

$$\forall t, s \in [0, +\infty[, \|\rho(t) - \rho(s)\|_{L^2(\Omega)} \leq C_\infty |t - s|,$$

hence (36). □

5. REGULARITY IN SPACE

While the regularity shown in the previous sections is in time, we can show regularity in space under additional conditions on V and on the boundary conditions. We will start from the problem with Dirichlet boundary conditions:

$$(39) \quad \min_{\substack{\rho \in E \\ \rho(0, \cdot) = m_0, \rho(T, \cdot) = m_T \\ \forall t \in [0, T], \int_{\Omega} \rho(t, x) dx = 1}} \int_0^T \int_{\Omega} (|\dot{\rho}(t, x)| + V(t, x)\rho(t, x) + f(\rho(t, x))) dx dt := F(\rho),$$

where we remind that $E := \text{BV}([0, T], L^1(\Omega)) \cap L^2([0, T] \times \Omega)$. Similarly to Section 3 and 4, we approximate the problem (48) by

$$(40) \quad \min_{\substack{\rho \in E_n \\ \rho(0, \cdot) = m_0^n, \\ \rho(T, \cdot) = m_T^n}} \int_0^T \left(\left(\sum_{i=1}^n \frac{1}{n} L_{\epsilon}(n\dot{\rho}_i(t)) + n\rho_i(t) \int_{A_i^n} V(t, x) dx + \frac{1}{n} f(n\rho_i(t)) \right) + \frac{(\sum_j \rho_j(t) - 1)^2}{2\delta} \right) dt := F_n(\rho),$$

with the usual choices for $L_{\epsilon}(s) = \sqrt{s^2 + \epsilon^2} + \epsilon s^2$ and E_n , the set of piecewise constant functions $\rho \in E$ such that

$$\forall t \in [0, T], \forall x \in \Omega, \rho(t, x) = \sum_{i=1}^n n\rho_i(t) \mathbb{1}_{A_i^n}(x),$$

where $\rho_i: [0, T] \rightarrow \mathbb{R}$ is a real-valued function. Note that here we prefer not to replace the Dirichlet boundary conditions with L^1 penalizations, so that we need to discretize the initial and final data as well. We then define m_i^n to be piecewise constant approximations of m_t (for $t = 0, T$), taking in particular $m_{i,t}^n := \frac{1}{|A_i^n|} \int_{A_i^n} m_t(x) dx$. We also set, as usual, $V_i^n(t) := \frac{1}{|A_i^n|} \int_{A_i^n} V(t, x) dx$

We will consider the problem solved by $\rho_i(t)$ for all $i \in \{1, \dots, n\}$:

$$(41) \quad \min_{n\rho_i(0) = m_{i,0}^n, n\rho_i(T) = m_{i,T}^n} \int_0^T \left(L_{\epsilon}(n\dot{\rho}_i(t)) + f(n\rho_i(t)) + n\rho_i(t)V_i^n + \frac{(\sum_j \rho_j(t) - 1)^2}{2\delta} \right) dt.$$

The Euler-Lagrange system of (41) is

$$(42) \quad \begin{cases} (L'_{\epsilon}(n\dot{\rho}_i(t)))' &= V_i^n(t, x_i) + f'(n\rho_i(t)) + c(t), \\ n\rho_i(0) &= m_{i,0}^n, \\ n\rho_i(T) &= m_{i,T}^n, \end{cases}$$

where $c(t) := \frac{\int \rho(t, y) dy - 1}{\delta n}$.

Lemma 5.1. *If ρ is a minimizer of (40), then for every i, j we have the following inequality:*

$$(43) \quad n \sup_{t \in [0, T]} \rho_i(t) - \rho_j(t) \leq \max \left(\sup_{t \in [0, T]} V_j^n(t) - V_i^n(t), m_{i,0}^n - m_{j,0}^n, m_{i,T}^n - m_{j,T}^n \right).$$

Proof. We consider

$$\max_{t \in [0, T]} \rho_i(t) - \rho_j(t).$$

We distinguish three cases:

- the maximum is reached on $(0, T)$,
- the maximum is reached at 0
- the maximum is reached at T .

(i) If the maximum is attained at $t_0 \in (0, T)$, we have $\dot{\rho}_i(t_0) = \dot{\rho}_j(t_0)$ as well as $\ddot{\rho}_i(t_0) \leq \ddot{\rho}_j(t_0)$. This implies $\ddot{\rho}_i(t_0)L'_{\epsilon}(n\rho_i(t_0)) \leq \ddot{\rho}_j(t_0)L'_{\epsilon}(n\rho_j(t_0))$. By using (42) and subtracting the equation for i and

that for j , we have

$$\begin{aligned} n\rho_i(t_0) - n\rho_j(t_0) &\leq \frac{1}{c_0}(f'_n(n\rho_i(t_0)) - f'_n(n\rho_j(t_0))) \\ &= \ddot{\rho}_i(t_0)L''_\epsilon(n\dot{\rho}_i(t_0)) - \ddot{\rho}_j(t_0)L''_\epsilon(n\dot{\rho}_j(t_0)) + V_j^n(t_0) - V_i^n(t_0) \\ &\leq V_j^n(t_0) - V_i^n(t_0) \leq \sup_{t \in [0, T]} V_j^n(t) - V_i^n(t), \end{aligned}$$

Consequently, $n(\rho_i(t_0) - \rho_j(t_0))$ is bounded by $\sup_{t \in [0, T]} V_j^n(t) - V_i^n(t)$.

(ii) If the maximum is reached at $t = 0$ we have

$$n\rho_i(t) - n\rho_j(t) \leq na_i^n - a_j^n$$

using the Dirichlet condition.

(iii) If the maximum is reached at $t = T$ we have

$$n\rho_i(t) - n\rho_j(t) \leq nb_i^n - b_j^n$$

using the other Dirichlet condition.

The claim follows by putting together the three cases □

Theorem 5.2. *Let ρ be a minimizer of (48) and ω a function such that*

$$\frac{1}{c_0}(V(t, x) - V(t, x')), m_t(x) - m_t(x') \leq \omega(x - x')$$

for all t, x, x' . Suppose either that ω is a constant or that $\lim_{z \rightarrow 0} \omega(z) = 0$ (i.e., ω is a modulus of continuity).

Then we have for a.e. t, x, x'

$$\rho(t, x) - \rho(t, x') \leq \omega(x - x').$$

Proof. We first consider the curves ρ_n minimizing (40). We write the estimate from Lemma 5.1 in terms of the densities $\rho_n(t, x) = n\rho_i(t)$. We obtain

$$\rho_n(t, x) - \rho_n(t, x') \leq \omega(x - x') + \varepsilon_n.$$

Indeed, if the function ω is a constant M , then the oscillations of $\frac{V}{c_0}, a$ and b are bounded by M and so $n\rho_i - n\rho_j \leq M$ as well. Otherwise, if ω is a modulus of continuity, then the functions $\frac{V}{c_0}, a$ and b are continuous and their averages on small pieces A_i^n can be replaced by the values at the center of these pieces up to a small error ε_n (which also takes into account that x and x' can differ from the centers of the corresponding pieces).

The family $(\rho_n)_n$ is bounded in $L^2([0, T] \times \Omega)$. Differently from the case of Section 3, we do not have Lipschitz estimates in time (note that this is not the same approximation as in Section 3, since we impose the Dirichlet boundary conditions instead of penalizing them), but luckily we will not need them. Indeed, the equicontinuity in time was essential to obtain uniform and pointwise bounds and deal with the boundary terms. Here we just use weak convergence in $L^2([-1, T+1] \times \Omega)$ after extending ρ_n to a on $[-1, 0]$ and to b on $[T, T+1]$. It is easy to see that ρ_n admits a weakly converging subsequence and that the limit solves (48). Note that the extension before $t = 0$ and after $t = T$ is needed to include the possible jump at those instants of time in the total variation and hence play the role of the Dirichlet boundary condition.

The conclusion comes from the following lemma 5.3. □

Lemma 5.3. *Let ρ_n be a sequence weakly converging to ρ in $L^2([0, T] \times \Omega)$ and suppose that there exists a function ω such that*

$$(44) \quad \forall n \in \mathbb{N}^*, \forall t \in [0, T], \forall x, x' \in \Omega, \rho_n(t, x) - \rho_n(t, x') \leq \omega(x - x') + \varepsilon_n$$

for a sequence $\varepsilon_n \rightarrow 0$.

Then we have

$$\rho(t, x) - \rho(t, x') \leq \omega(x - x')$$

whenever (t, x) and (t, x') are Lebesgue points of ρ .

Proof. Let us take (t_0, x_0) and (t_0, y_0) in $[0, T] \times \Omega$, two Lebesgue points of ρ . Let $r > 0$ be such that $]t_0 - r, t_0 + r[\subset [0, T]$, $B(x_0, r) \subset \Omega$ and $B(y_0, r) \subset \Omega$, we have in particular

$$(45) \quad \int_{t_0-r}^{t_0+r} \int_{B(x_0, r)} \rho_n(t, x) dx dt \xrightarrow{n \rightarrow \infty} \int_{t_0-r}^{t_0+r} \int_{B(x_0, r)} \rho(t, x) dx dt,$$

$$(46) \quad \text{and } \int_{t_0-r}^{t_0+r} \int_{B(y_0, r)} \rho_n(t, x) dx dt \xrightarrow{n \rightarrow \infty} \int_{t_0-r}^{t_0+r} \int_{B(y_0, r)} \rho(t, x) dx dt.$$

By assumption we have the inequality

$$\begin{aligned} \int_{t_0-r}^{t_0+r} \left(\int_{B(x_0, r)} \rho_n(t, x) dx - \int_{B(y_0, r)} \rho_n(t, x') dx' \right) dt &= \int_{t_0-r}^{t_0+r} \int_{B(x_0, r)} (\rho_n(t, x) - \rho_n(t, x - x_0 + y_0)) dx dt \\ &\leq \omega(x_0 - y_0) + \varepsilon_n \end{aligned}$$

By taking the limit $n \rightarrow \infty$ and applying (45) and (46) to the left-hand side of the inequality, we get

$$\int_{t_0-r}^{t_0+r} \left(\int_{B(x_0, r)} \rho(t, x) dx - \int_{B(y_0, r)} \rho(t, x') dx' \right) dt \leq \omega(x_0 - y_0).$$

Since (t_0, x_0) and (t_0, y_0) are Lebesgue points, we can pass to the limit $r \rightarrow 0$ and obtain

$$\rho(t_0, x_0) - \rho(t_0, y_0) \leq \omega(x_0 - y_0). \quad \square$$

The case where the Dirichlet conditions are replaced by penalizations are harder to deal with. The only case that is easy to consider requires that the transversality condition is the same for ρ_i and ρ_j . We can obtain the following results for which we just sketch the modifications to the previous proofs.

Theorem 5.4. *Let ρ be a minimizer of*

$$(47) \quad \min_{\substack{\rho \in E \\ \rho(0, \cdot) = m_0, \\ \forall t \in [0, T], \int_{\Omega} \rho(t, x) dx = 1}} \int_0^T \int_{\Omega} (|\dot{\rho}(t, x)| + V(t, x)\rho(t, x) + f(\rho(t, x))) dx dt + \int \Psi_T \rho(T, x).$$

Let $A \subset \Omega$ be a set where Ψ_T is constant. Suppose that ω is a function (either constant or a modulus of continuity) such that $\frac{1}{c_0}(V(t, x) - V(t, x')), m_0(x) - m_0(x') \leq \omega(x - x')$ for all t, x, x' . Then we have for a.e. $x, x' \in A$

$$\rho(t, x) - \rho(t, x') \leq \omega(x - x').$$

Proof. The approximation will be the same as before, but the Dirichlet boundary condition at $t = T$ is replaced by a transversality condition. Choosing a decomposition into pieces A_i^n such that x, x' belong to two pieces contained in A , this transversality condition will be the same for the two curves $\rho_i(t)$ and $\rho_j(t)$ that we need to consider to estimate $\rho(t, x) - \rho(t, x')$. In particular, we will have $\dot{\rho}_i(T) = \dot{\rho}_j(T)$. Hence, when considering $\max_i \rho_i(t) - \rho_j(t)$, the maximum could be attained on $t_0 = T$ but in this case the first-order optimality condition will be satisfied, and this allows to obtain the second-order one, which is the main tool to estimate $\rho_i(t_0) - \rho_j(t_0)$. The rest of the analysis goes as in the rest of the Section. \square

Theorem 5.5. *Let ρ be a minimizer of*

$$(48) \quad \min_{\substack{\rho \in E \\ \rho(0, \cdot) = m_0, \\ \forall t \in [0, T], \int_{\Omega} \rho(t, x) dx = 1}} \int_0^T e^{-rt} \int_{\Omega} (|\dot{\rho}(t, x)| + V(t, x)\rho(t, x) + f(\rho(t, x))) dx dt := F(\rho).$$

Suppose that ω is a function (either constant or a modulus of continuity) such that $\frac{1}{c_0}(V(t, x) - V(t, x')), m_0(x) - m_0(x') \leq \omega(x - x')$ for all t, x, x' . Then we have for a.e. x, x'

$$\rho(t, x) - \rho(t, x') \leq \omega(x - x').$$

Proof. We first replace F with F_T as in Section 4. Then, the claim is the same as in Theorem 5.4 with $\phi_T = 0$ (hence we can take $A = \Omega$) with the only difference that have an extra coefficient e^{-rt} . This lets an extra term $-rL'_\varepsilon(n\dot{\rho}_i)$ appear in the Euler-Lagrangian equation, but this term cancels when taking the difference between i and j because of the first-order optimality condition for t_0 . The rest of the analysis goes as in Theorem 5.4 and in the rest of the Section. This provides a uniform estimate on ρ_T , independent of T , and we can then take the limit $T \rightarrow \infty$. \square

We conclude summarizing the result that we can obtain in terms of spatial regularity:

- For the problem with two Dirichlet boundary conditions, if V, m_0 , and m_T are continuous, then the solution ρ shares the same modulus of continuity of $\frac{V}{c_0}, m_0$, and m_T . Combining this with the L^2 Lipschitz regularity in time obtained in Section 3 this gives a uniform continuity result in (t, x) .
- In the same problem, if V, m_0 , and m_T are only bounded, then ρ is bounded, since its oscillation is bounded by that of $\frac{V}{c_0}, m_0$, and m_T and its L^2 norm is also bounded.
- For the problem with a Dirichlet boundary condition at $t = 0$ and a penalization Ψ_T at $t = T$, if Ψ_T is piecewise constant and V and m_0 are continuous, then ρ is piecewise continuous, and hence bounded. The solution ρ is also bounded if we only assume V and m_0 to be bounded, and Ψ_T to be piecewise constant.
- In the infinite-horizon problem with Dirichlet boundary condition at $t = 0$, the solution shares the same modulus of continuity of $\frac{V}{c_0}$ and m_0 and is uniformly continuous in (t, x) . It is uniformly bounded if V and m_0 are bounded.
- In the periodic case (which we briefly presented in Section 2 but did not develop here) the solution ρ shares the same continuity or boundedness of $\frac{V}{c_0}$.

6. NUMERICAL APPROXIMATION

In this section, numerical simulations are carried out on the following problem

$$(49) \quad \min_{\substack{\rho \in E \\ \forall t \in [0, T], \int_{\Omega} \rho(t, x) dx = 1 \\ \rho \geq 0}} \int_0^T \int_{\Omega} \left(\lambda |\dot{\rho}(t, x)| + V(t, x) \rho(t, x) + \frac{\rho(t, x)^2}{2} \right) dx dt + \psi_0(\rho(0)) + \psi_T(\rho(T)) = F(\rho)$$

and on some of its variants (Dirichlet conditions, periodic case...). The parameter $\lambda > 0$ allows us to add more or less importance to the L^1 -norm.

For these numerical examples, we take the domain $\Omega := [0, S]$ to be one-dimensional with $S > 0$. In this section, we study the following cases:

- periodic solutions in 1D (only the time variable, Section 6.2, Figure 1).
- periodic solutions in 2D (periodic in time, Section 6.3, Figures 2, 3).
- non-periodic solutions in 2D, with or without Dirichlet conditions or penalizations at the time boundary (Section 6.4, Figures 4, 5, and 6).

6.1. The numerical method. Let $\{t_0, \dots, t_K\}$ and $\{x_0, \dots, x_N\}$ be the regular subdivisions of respectively $[0, T]$ and $[0, S]$. To approximate the integral in F , the left-rectangle method will be used, hence we will consider solutions $(\rho(t_i, x_j))_{\substack{0 \leq i \leq K-1 \\ 0 \leq j \leq N-1}}$ in the vector space $\mathbb{R}^{K \cdot N}$. With an abuse of notations, we also write $V \in \mathbb{R}^{K \cdot N}$ the vector $(V(t_i, x_j))_{\substack{0 \leq i \leq K-1 \\ 0 \leq j \leq N-1}}$. In the following, the notations $\langle \cdot | \cdot \rangle$ and $\|\cdot\|$ designate respectively the scalar product and the euclidean norm in a vector space of finite dimension.

The problem that will be numerically solved can be written in the form

$$\min_{\rho \in \mathbb{R}^{K \cdot N}} f(\rho) + g(\mathcal{A}\rho),$$

where, of course, f and g have nothing to do with the functions introduced in the previous sections. Here $\rho \in \mathbb{R}^{K \cdot N}$, $f(\rho) = h \langle V | \rho \rangle + h \frac{\|\rho\|_2^2}{2}$ is a proper, closed and h -strongly convex function and $g(\rho_0, \rho_1, \rho_2) = \lambda \|\rho_0\|_1 + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2)$ is a proper, closed and convex function. The notation $\|\cdot\|_1$ refers to the l_1 -norm in a vector space of finite dimension and δ_{C_i} is the indicator function defined by

$$\delta_{C_i}(\rho) = \begin{cases} 0, & \rho \in C_i, \\ +\infty, & \rho \notin C_i, \end{cases}$$

where $C_0 = (\mathbb{R}^+)^{K \cdot N}$ and $C_1 = \{\rho \in \mathbb{R}^{K \cdot N}; \forall i \in \{0, \dots, K-1\}, \sum_{j=0}^{N-1} \rho(t_i, x_j) \cdot l = 1\}$ are the sets of constraints. The linear transformation \mathcal{A} is $\mathcal{A}(\rho) = (A\rho, \rho, \rho)$, where the matrix A will be detailed for each case.

An algorithm to approximate the solution of this problem is already known and is called (Fast) Dual Proximal Gradient Method which can be found in [2]. We will use its primal representation which does not involve the dual representation of the problem.

Since g is a sum of separable functions, its proximal operator is

$$\text{prox}_g(\rho_0, \rho_1, \rho_2) = \left(\text{prox}_{\lambda \|\cdot\|_1}(\rho_0), \text{prox}_{\delta_{C_0}}(\rho_1), \text{prox}_{\delta_{C_1}}(\rho_2) \right).$$

We remind that for all $x \in \mathbb{R}$,

$$\text{prox}_{\lambda|\cdot|}(x) = [|x| - \lambda]_+ \text{sgn}(x) \text{ and } \text{prox}_{\delta_C}(x) = P_C(x)$$

where P_C is the projection on the set C .

The algorithm is described below.

Initialization: $L \geq \frac{\|A\|^2}{h}$, $w^0 = y^0 \in (\mathbb{R}^{K \cdot N})^3$, $t_0 = 1$.

Step for $k \geq 0$:

- $u^k = \text{argmax}_u \{ \langle u, \mathcal{A}^T w^k \rangle - f(u) \} = \frac{\mathcal{A}^T w^k}{h} - V$
 $= \frac{1}{h} (A^T w_0^k + w_1^k + w_2^k) - V$
- $y_0^{k+1} = w_0^k - \frac{1}{L} A u^k + \frac{1}{L} \text{prox}_{L\lambda\|\cdot\|_1} (A u^k - L w_0^k)$
- $y_1^{k+1} = w_1^k - \frac{1}{L} u^k + \frac{1}{L} \text{prox}_{L\delta_{C_0}} (u^k - L w_1^k)$
- $y_2^{k+1} = w_2^k - \frac{1}{L} u^k + \frac{1}{L} \text{prox}_{L\delta_{C_1}} (u^k - L w_2^k)$
- $t^{k+1} = \frac{k+1+a}{a}$
- $w^{k+1} = y^{k+1} + \frac{t^k - 1}{t^{k+1}} (y^{k+1} - y^k)$

The role of the parameter t^k is to accelerate the algorithm, in the spirit of Nesterov's accelerated gradient or (in the proximal case) of the FISTA algorithm (see [15, 3]).

Note that our problem is essentially, up to minor modifications and the presence of extra constraints (unit mass, positivity, Dirichlet conditions), a simplified version of some standard problems in image denoising based on total variation (see, for instance, the classical paper [16]): here, whether the problem is 1D or 2D or higher-dimensional, the main feature is that the total variation is only computed in time.

The following sections describe different examples of solutions to (49) by using this algorithm. For each case, the differences with the description above will be specified.

6.2. 1D-periodic. When we consider the periodic problem (in time, so that the interval $[0, T]$ becomes a circle of length T), and we assume $S = T$ and $V(t, x) = v(t - x)$ for an S -periodic function v , it is possible to reduce the problem to the one dimensional case, namely a problem with one only variable in $[0, T]$ instead of $[0, T] \times [0, S]$. One expects the solution $\rho(t, x) = u(t - x)$ to be transported according to time. The uniqueness of the solution and the symmetry with respect to translations in both time and space (replacing (t, x) with $(t + \delta, x + \delta)$) show that the solution should indeed be of this form. Then, a change of variables $y = t - x$ can be carried out in F as following:

$$\begin{aligned} F(\rho) &= \int_0^T \int_0^S \left(\lambda |\dot{u}(t - x)| + v(t - x) u(t - x) + \frac{u(t - x)^2}{2} \right) dx dt \\ &= \int_0^T \int_t^{t-S} \left(\lambda |\dot{u}(y)| + v(y) u(y) + \frac{u(y)^2}{2} \right) dy dt \\ &= T \int_0^S \left(\lambda |\dot{u}(y)| + v(y) u(y) + \frac{u(y)^2}{2} \right) dy. \end{aligned}$$

The problem reduces hence to the search of an S -periodic solution with one variable. By the way, up to multiplicative and additive constants this problem is equivalent to minimizing

$$\int_0^S \left(\lambda |\dot{u}(y)| + \frac{|u(y) + v(y)|^2}{2} \right) dy,$$

which is exactly the problem described as an example at the end of Section 2, with $\omega = -v$, except for the constraints $\int_0^S u(y) dy = 1$ and $u \geq 0$.

In the following subsections, we will see that the solution obtained through numerical simulations coincide with u computed as above.

For the numerical simulation shown in Figure 1 we take $v(y) = -a_0 \cos(\frac{2\pi}{S} y)$ and the choice of the parameters are shown in Table 1.

Let $\{y_0, \dots, y_K\}$ be a regular subdivision of $[0, S]$ and $h := y_1 - y_0$ the step-size of the subdivision. By the left-rectangle method, the problem is approximated by

$$(50) \quad \sum_{i=0}^{K-1} \left(\lambda |u(y_{i+1}) - u(y_i)| + v(y_i) u(y_i) h + \frac{u(y_i)^2}{2} h \right) + \delta_{C_0}(u) + \delta_{C_1}(u).$$

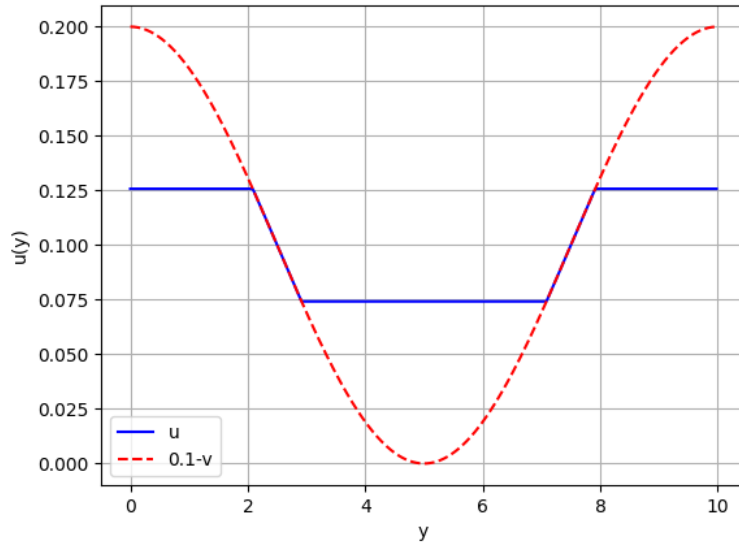


FIGURE 1. The simulation of the solution u to the 1D periodic case with $v(y) = -a_0 \cos(\frac{2\pi}{S}y)$. The parameters are displayed in Table 1. The blue solid line corresponds to the solution u , while the red dashed line is the profile of $c - v$ with $c = 1/S$.

Parameter	Value
S	10
K	500
h	0.02
L	$6/h$
a_0	0.1
λ	0.1

TABLE 1. Parameters for the solution to the problem in Fig. 1.

We observe that with this choice of parameters the solution u is strictly positive and “follows” the profile of $c - v$ for a constant c which appears as a Lagrange multiplier for the mass constraint and allows to obtain unit mass. The function u is also a(n unconstrained) solution of the problem at the end of Section 2 with $\omega = c - v$ and its profile, shown in Figure 1, is consistent with the explicit description of the solution which we gave.

Remark 6.1. *If $v(y) = -a_0 \cos(\frac{2\pi}{S}y)$, it is possible to compute the critical λ at which the aspect of the solution u switches from Figure 1 to the constant solution. Assuming that the solution becomes constant when $\epsilon = \frac{S}{4}$ and extending the solution periodically on $[-S, 0]$, we integrate the Euler-Lagrange equation $z' = v + u - c$ over $[-\epsilon, \epsilon]$:*

$$\begin{aligned} -2\lambda &= \int_{-\epsilon}^{\epsilon} z'(y) dy = \int_{-\epsilon}^{\epsilon} (v(y) + u(y) - c) dy = \int_{-\epsilon}^{\epsilon} (v(y) - v(\epsilon)) dy \\ &= \int_{-\epsilon}^{\epsilon} \left(-a_0 \cos\left(\frac{2\pi}{S}y\right) + a_0 \cos\left(\frac{2\pi}{S}\epsilon\right) \right) dy = -a_0 \frac{S}{\pi} \sin\left(\frac{2\pi}{S}\epsilon\right) + 2\epsilon a_0 \cos\left(\frac{2\pi}{S}\epsilon\right). \end{aligned}$$

By taking $\epsilon = \frac{S}{4}$, we obtain that $\lambda = \frac{a_0 S}{2\pi}$.

When $\lambda > \frac{a_0 S}{2\pi}$, the solution u is constant equal to c . While the condition $\lambda < \frac{a_0 S}{2\pi}$ gives the solution u as in Figure 1. This explains the choice of parameter λ .

6.3. 2D periodic in time. We consider now the case where we keep two variables, but we assume the time domain to be periodic. This case is slightly simpler to handle from the point of view of the discretization of the time derivative.

By using the subdivisions described in Section 6.1, we approximate the integral F by the left-rectangle method as following:

$$(51) \quad \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \left(\lambda |\rho(t_{i+1}, x_j) - \rho(t_i, x_j)| l + V(t_i, x_j) \rho(t_i, x_j) h l + \frac{\rho(t_i, x_j)^2}{2} h l \right),$$

where $h := t_1 - t_0$ and $l := x_1 - x_0$ are the step-sizes of each subdivision. We will consider that $\rho(t_0, x_j) = \rho(t_K, x_j)$ for all $j \in \{0, N-1\}$.

With this discretization, we look for a solution in the space $\mathbb{R}^{K \cdot N}$ where ρ is viewed as a vector.

The discretized problem is

$$(52) \quad \min_{\rho \in \mathbb{R}^{K \cdot N}} \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \left(\lambda |\rho(t_{i+1}, x_j) - \rho(t_i, x_j)| l + V(t_i, x_j) \rho(t_i, x_j) h l + \frac{\rho(t_i, x_j)^2}{2} h l \right) + \delta_{C_0}(\rho) + \delta_{C_1}(\rho).$$

To avoid numerical errors due to small values, searching a minimizer of (52) is the same as searching a minimizer to the problem divided by l :

$$\min_{\rho \in \mathbb{R}^{K \cdot N}} \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \left(\lambda |\rho(t_{i+1}, x_j) - \rho(t_i, x_j)| + V(t_i, x_j) \rho(t_i, x_j) h + \frac{\rho(t_i, x_j)^2}{2} h \right) + \delta_{C_0}(\rho) + \delta_{C_1}(\rho).$$

The function $f: \mathbb{R}^{K \cdot N} \rightarrow \mathbb{R}$ defined by

$$(53) \quad f(\rho) = \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \left(V(t_i, x_j) \rho(t_i, x_j) h + \frac{\rho(t_i, x_j)^2}{2} h \right) = h \langle V | \rho \rangle + h \frac{\|\rho\|^2}{2}$$

corresponds to the function f described in Section 6.1.

The linear transformation $\mathcal{A}: \mathbb{R}^{K \cdot N} \rightarrow (\mathbb{R}^{K \cdot N})^3$ is the same as in Section 6.1. The squared matrix A is of order $K \cdot N$ and its coordinates are

$$A_{i,j} = \begin{cases} -1 & \text{if } j = i, \\ 1 & \text{if } j \equiv i + N[K \cdot N], \\ 0 & \text{otherwise,} \end{cases}$$

namely

$$A = \begin{pmatrix} -I_N & I_N & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & -I_N & I_N \\ I_N & & & & & -I_N \end{pmatrix}$$

where I_N is the identity matrix of order N and blank space corresponds to zeros.

The function $g: (\mathbb{R}^{K \cdot N})^3 \rightarrow \mathbb{R}$ is defined as in Section 6.1:

$$(54) \quad \begin{aligned} g(\rho_0, \rho_1, \rho_2) &= \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \lambda |\rho_0(t_i, x_j)| + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2) \\ &= \lambda \|\rho_0\|_1 + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2). \end{aligned}$$

In Figure 2, we consider the same case as in Figure 1, but with a 2D approach. The profile of the solution ρ at time 0 is the same as in Figure 1. Whilst in Figure 1, the problem is viewed on one dimensional space, here we can see that the solution is transported according to time from the left to the right.

Parameter	Value
T	10
S	10
K	100
N	500
h	0.02
L	$6/h$
a_0	0.1
λ	0.1

TABLE 2. Parameters for the numerical simulation of the solution to (49).

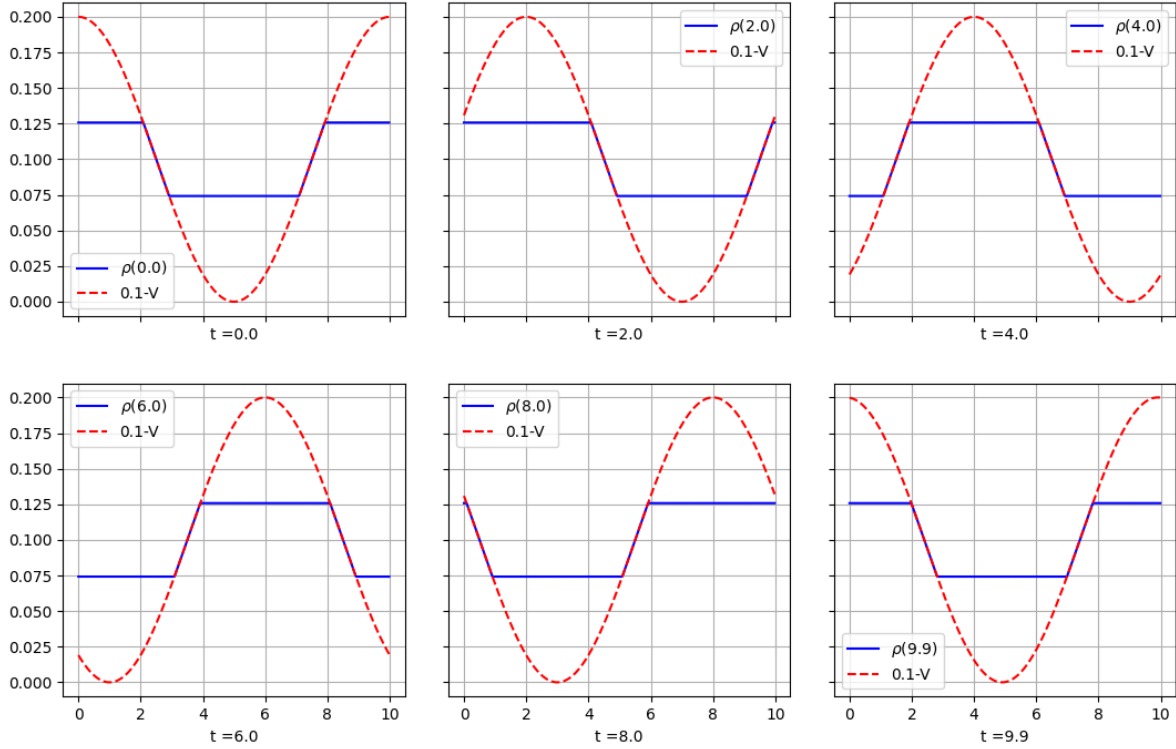


FIGURE 2. Profile of the solution to (49) with $V(t, x) = -a_0 \cos(\frac{2\pi}{S}(t - x))$ at times $t = 0, 2, 4, 6, 8, 9.9$. The blue solid line describes the solution ρ at each specified time and the red dashed line is the profile of $0.1 - V$ at each time. The parameters are displayed in Table 2.

The next case (see Figure 3) we show is different. Here we impose a Dirichlet boundary condition at time 0 which corresponds here to m_0 being the constant density equal to $1/S$. This is the same as studying the non-periodic problem on $[0, T]$ and imposing two Dirichlet boundary conditions, which are (by chance) equal. In this case, keeping the periodic structure allows to use a simpler form of the matrix A .

Let $m_0 \in \mathbb{R}^N$ a vector verifying the constraints, i.e.,

$$m_0 \in (\mathbb{R}^+)^N \text{ and } \sum_{j=0}^{N-1} m_{0,j} l = 1.$$

Let us define the new set $C_2 = \{\rho \in \mathbb{R}^{K \cdot N}; \forall j \in \{0, \dots, N-1\}, \rho(0, x_j) = m_{0,j}\}$ which encodes the constraint $\rho(0) = m_0$.

The function f is the same as in (53). The linear transformation is now $\mathcal{A}\rho = (A\rho, \rho, \rho, \rho)$ and the function $g: (\mathbb{R}^{K \cdot N})^4 \rightarrow \mathbb{R}$ is

$$\begin{aligned} g(\rho_0, \rho_1, \rho_2, \rho_3) &= \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \lambda |\rho_0(t_i, x_j)| + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2) + \delta_{C_2}(\rho_3) \\ &= \lambda \|\rho_0\|_1 + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2) + \delta_{C_2}(\rho_3). \end{aligned}$$

In Figure 3 we see that the Dirichlet condition is indeed verified at $t = 0$ and then, immediately, the solution jumps (coherently with Remark 2.3) from the constant state to $t = 0.1$. One can notice that the profile is different from Figure 2 in a way that on the subintervals where the solution does not follow $0.1 - V$, it is not constant anymore. However, between times $t = 4$ and $t = 6$, the solution comes back to the profile when there is no boundary condition, namely, it is constant when it does not follow $0.1 - V$. When $t \geq 6$, the profile of the solution varies again on the subintervals where it should be constant.

6.4. 2D, non periodic. Unlike Sections 6.2 and 6.3, we do not impose anymore a periodic time behavior, and the value $\rho(t_K, x_j)$ of the solution $\rho \in \mathbb{R}^{(K+1) \cdot N}$ may be different from $\rho(t_0, x_j)$. The integral in F is

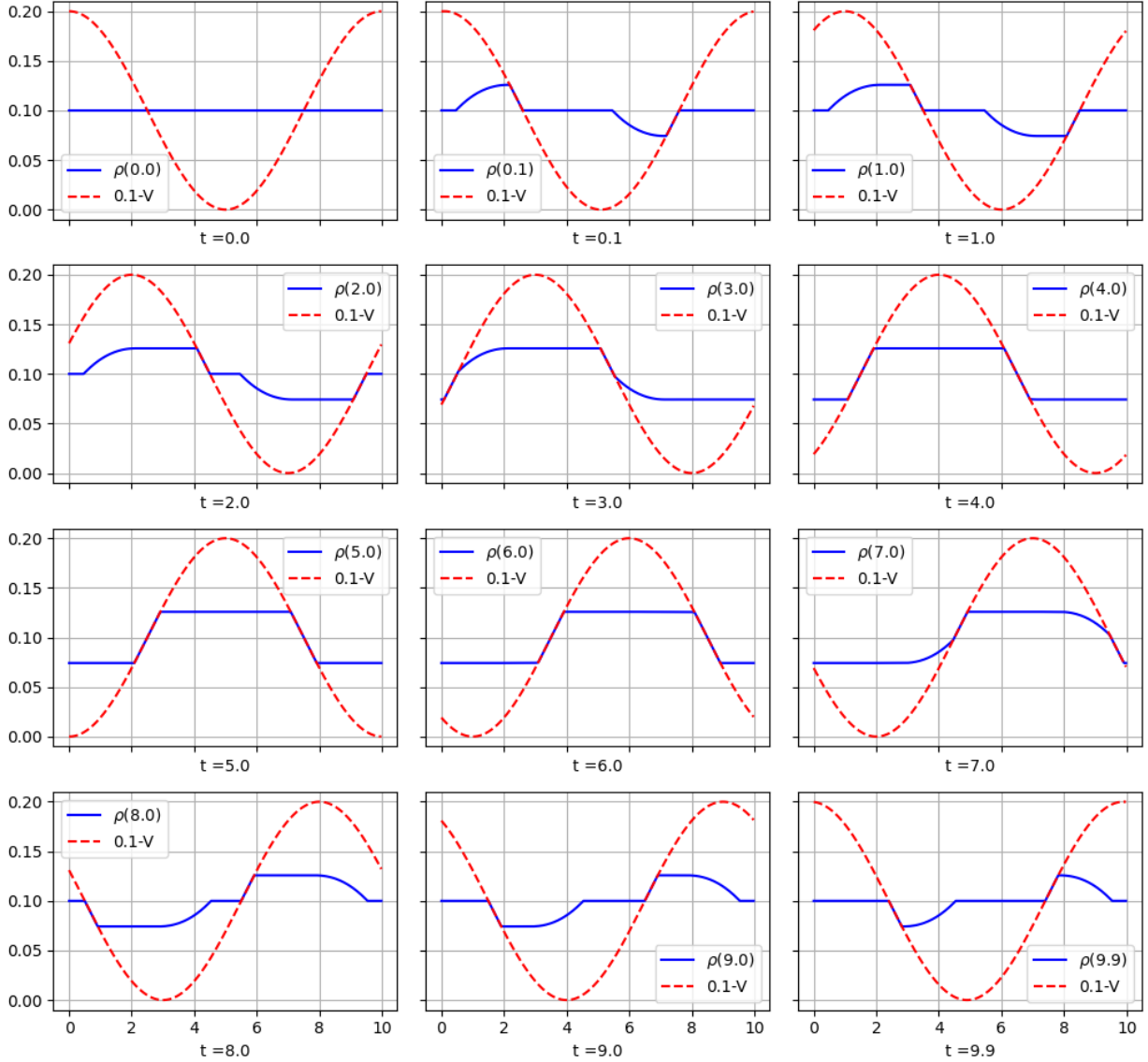


FIGURE 3. Simulation of the solution to (49) at times $t = 0, 0.1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 9.9$ with $V(t, x) = -a_0 \cos(\frac{2\pi}{S}(t - x))$ and $m_0(x) = 1/S$ with parameters from Table 2 except for parameter L where $L = 7/h$. The blue solid line is the profile of the solution ρ and the red dashed line corresponds to $0.1 - V$ taken at each time t .

now approximated by:

$$(55) \quad \sum_{j=0}^{N-1} \left(\sum_{i=0}^{K-1} \lambda |\rho(t_{i+1}, x_j) - \rho(t_i, x_j)| l + \sum_{i=0}^K \left(V(t_i, x_j) \rho(t_i, x_j) h l + \frac{\rho(t_i, x_j)^2}{2} h l \right) \right).$$

Remark 6.2. The formula (55) contains an extra term in t_K which should not be in the left-rectangle method. However, with this choice of discretization, the function f is strongly convex which allows us to apply the algorithm. The error is of order $o(h)$ so it is negligible.

The new function f is

$$(56) \quad f(\rho): \mathbb{R}^{(K+1) \cdot N} \rightarrow \mathbb{R}$$

$$\rho \mapsto h \langle V | \rho \rangle + h \frac{\|\rho\|^2}{2} = \sum_{i=0}^K \sum_{j=0}^{N-1} \left(V(t_i, x_j) \rho(t_i, x_j) h + \frac{\rho(t_i, x_j)^2}{2} h \right).$$

The function g is defined as in (54) and \mathcal{A} is the same as in Section 6.1 with a different matrix A which is rectangle of dimension $(K \cdot N) \times ((K + 1) \cdot N)$ such that

$$A_{i,j} = \begin{cases} -1 & \text{if } i = j, \\ 1 & \text{if } j = i + N, \\ 0 & \text{otherwise,} \end{cases}$$

namely

$$A = \begin{pmatrix} -I_N & I_N & & \\ & \ddots & \ddots & \\ & & -I_N & I_N \end{pmatrix}.$$

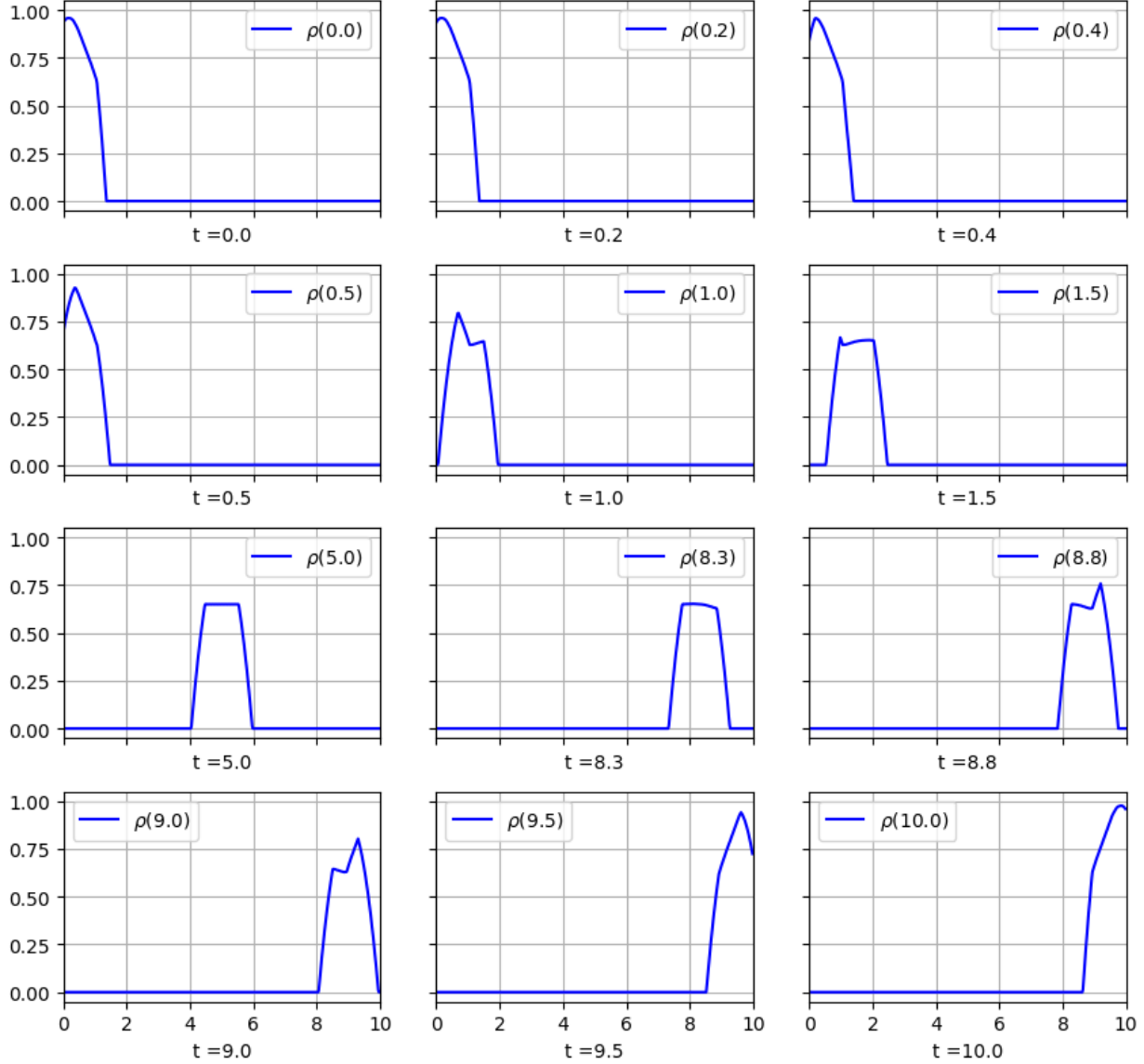


FIGURE 4. The simulation of the solution to (49) at times $t = 0, 0.2, 0.4, 0.5, 1, 1.5, 5, 8.3, 8.8, 9, 9.5, 10$ with $V(t, x) = (t - x)^2$ and parameters from Table 2. The blue solid line corresponds to the solution ρ at different times.

Figure 4 shows the simulation of the solution ρ to the problem (49) with a given V which is not periodic anymore. No Dirichlet boundary conditions nor penalizations ψ_t at $t = 0, T$ are considered. The different times are chosen to show the most significant changes in the profile of ρ . Since $V(t, x)$ is minimal at $x = t$,

the general behavior is that the solution is transported from the left to the right. Between times $t = 1.5$ and $t = 8.3$, the solution either follows $c - V$ (where c is a constant to define) or it is constant. Close to the time boundaries, the behavior is different.

Remark 6.3. *The constant c such that ρ follows $c - V$ depends on λ . The parameter λ can be studied as in remark 6.1. Different values of λ can either imply that ρ follows $c - V$ or that it follows $c - V$ and is constant in the middle.*

The next case that we present (Figure 5) involves Dirichlet conditions in time. Let $m_0 \in \mathbb{R}^N$ and $m_T \in \mathbb{R}^N$ be two vectors verifying the boundary conditions, i.e,

$$m_0 \in (\mathbb{R}^+)^N \text{ and } \sum_{j=0}^{N-1} m_{0,j} l = 1,$$

$$m_T \in (\mathbb{R}^+)^N \text{ and } \sum_{j=0}^{N-1} m_{T,j} l = 1.$$

Let us define the new set of constraints by $C_2 = \{\rho \in \mathbb{R}^{K \cdot N}; \forall j \in \{0, \dots, N-1\}, \rho(0, x_j) = m_{0,j} \text{ and } \rho(T, x_j) = m_{T,j}\}$. The function f is defined as in (56), $\mathcal{A}\rho = (A\rho, \rho, \rho, \rho)$ and the function $g: (\mathbb{R}^{K \cdot N})^4 \rightarrow \mathbb{R}$ is

$$g(\rho_0, \rho_1, \rho_2, \rho_3) = \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \lambda |\rho_0(t_i, x_j)| + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2) + \delta_{C_2}(\rho_3)$$

$$= \lambda \|\rho_0\|_1 + \delta_{C_0}(\rho_1) + \delta_{C_1}(\rho_2) + \delta_{C_2}(\rho_3).$$

Remark 6.4. *Instead of adding the constraint $\rho(0) = m_0$ in the problem, one could have discretized the integral with the right-rectangle method in time and directly used that $\rho(0) = m_0$ in the algorithm.*

The profile of Figure 5 is quite similar to Figure 4 and mainly differs around $t = 0, T$. Again, a jump in time is observed.

The very last example we consider is non-periodic in time and involves both a Dirichlet condition at $t = 0$ and penalization at $t = T$. We consider the penalization

$$\psi_T(\rho(T)) = \int_0^S \Psi_T(x) \rho(T, x) dx.$$

The left-rectangle method applied to ψ_T gives

$$\psi_T(\rho(T)) = \sum_{j=0}^{N-1} \Psi_T(x_j) \rho(T, x_j) l.$$

We use the discretization from (55). The difference is in the definition of V . We shall take \tilde{V} such that

$$\tilde{V} = V + (0, \dots, 0, \frac{\Psi_T}{h}).$$

Again, the profile of Figure 6 is similar to Figures 4 and 5 except for the behavior close to the time boundaries. In Figure 6, we have put penalization at times 0 and T such that it costs less to be close to 0 in space, hence the concentration of the mass at time 0 in $x = 0$ and the jump at time T .

REFERENCES

- [1] L. AMBROSIO AND P. TILLI, *Topics on analysis in metric spaces*. Oxford Lecture Series in Mathematics and its Applications (25). Oxford University Press, Oxford, 2004.
- [2] A. BECK, *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, 2017.
- [3] A. BECK AND M. TEOULLE, A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems, *SIAM Journal on Imaging Sciences* Vol. 2, Iss. 1, 183-202 (2009)
- [4] J.-D. BENAMOU AND Y. BRENIER, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84, 375-393, 2000.
- [5] A. BENSOUSSAN, J.-L. LIONS, *Contrôle impulsionnel et inéquations variationnelles*, Gauthier-Villars, Paris, 1982.
- [6] G. BUTTAZZO, C. JIMENEZ, AND E. OUDET, An Optimization Problem for Mass Transportation with Congested Dynamics *SIAM J. Control Optim.* 48 (2010), 1961-1976.
- [7] P. CARDALIAGUET, Notes on Mean Field Games (from P.-L. Lions' lectures at Collège de France), available at <https://www.ceremade.dauphine.fr/~cardalia/>
- [8] M. HUANG, R.P. MALHAMÉ, P.E. CAINES, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Communication in information and systems*, 6 (2006), No. 3, 221-252.

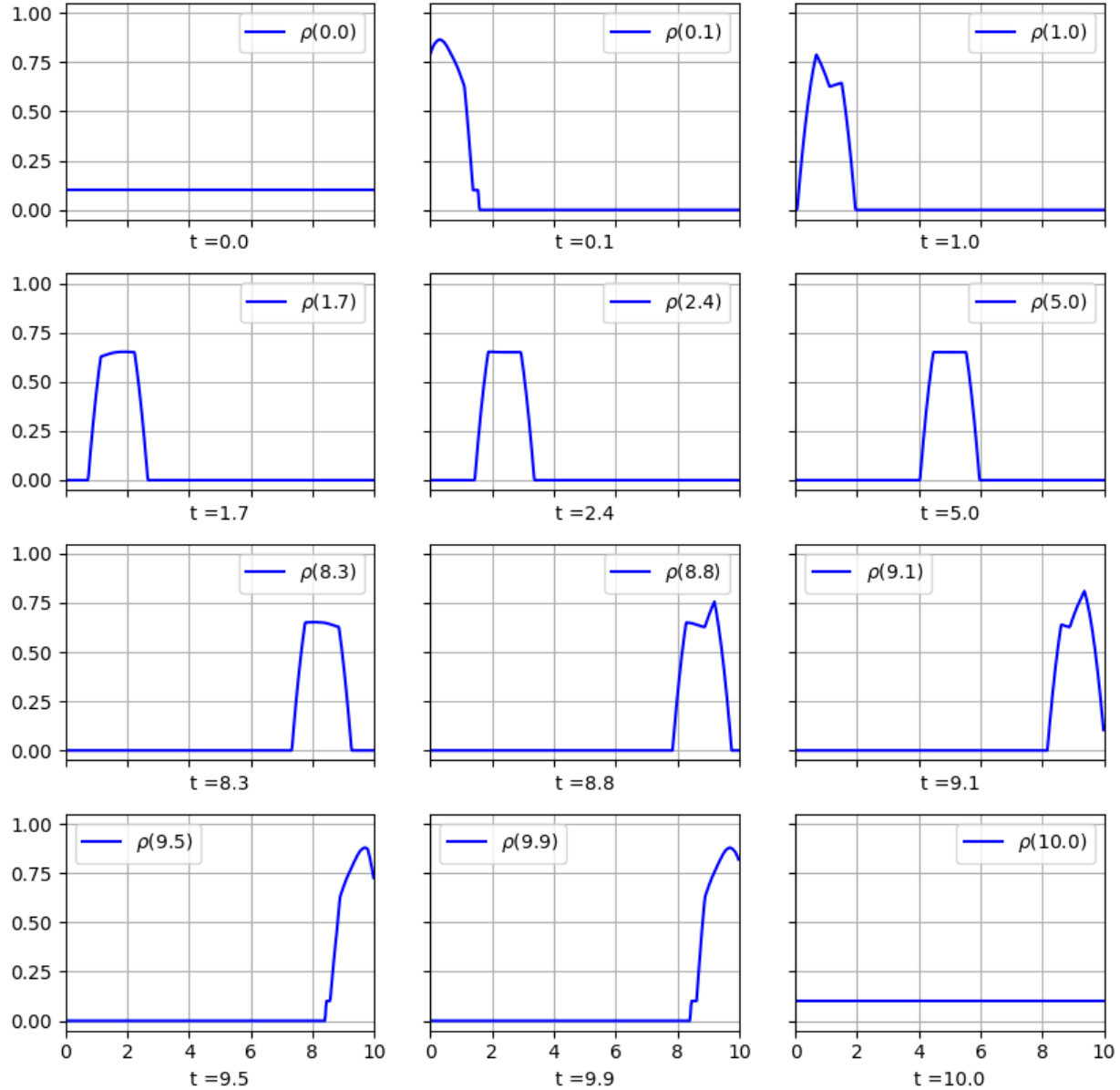


FIGURE 5. The simulation of the solution to (49) at times $t = 0, 0.1, 1, 1.7, 2.4, 5, 8.3, 8.8, 9.1, 9.5, 9.9, 10$ with $V(t, x) = (t - x)^2$, $m_0(x) = m_T(x) = 1/S$ and parameters from Table 2 except from L which is $L = 7/h$. The blue solid line corresponds to the solution ρ at different times.

- [9] J.-M. LASRY, P.-L. LIONS, Jeux à champ moyen. I. Le cas stationnaire, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 9, 619–625.
- [10] J.-M. LASRY, P.-L. LIONS, Jeux à champ moyen. II. Horizon fini et contrôle optimal, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 10, 679–684.
- [11] J.-M. LASRY AND P.-L. LIONS, Mean-Field Games, *Japan. J. Math.* 2, 229–260, 2007.
- [12] P.-L. LIONS, Series of lectures on Mean Filed Games, *Collège de France*, Paris, 2006-2012, video-recorded and available at the web page <http://www.college-de-france.fr/site/audio-video/>
- [13] J.-J. MOREAU Bounded variation in time, in *Topics in Nonsmooth Mechanics*, (Ed. J.J. Moreau, P.D. Panagiotopoulos and G. Strang), Birkhäuser (Basel, Boston, Berlin), 1988, p. 1-74.
- [14] J.-J. MOREAU Jump functions of a real interval to a Banach space *Annales de la Faculté des sciences de Toulouse : Mathématiques, Série 5, Tome S10* (1989), pp. 77-91.
- [15] Y. NESTEROV A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. *Doklady AN USSR*. 269: 543â547, (1983).

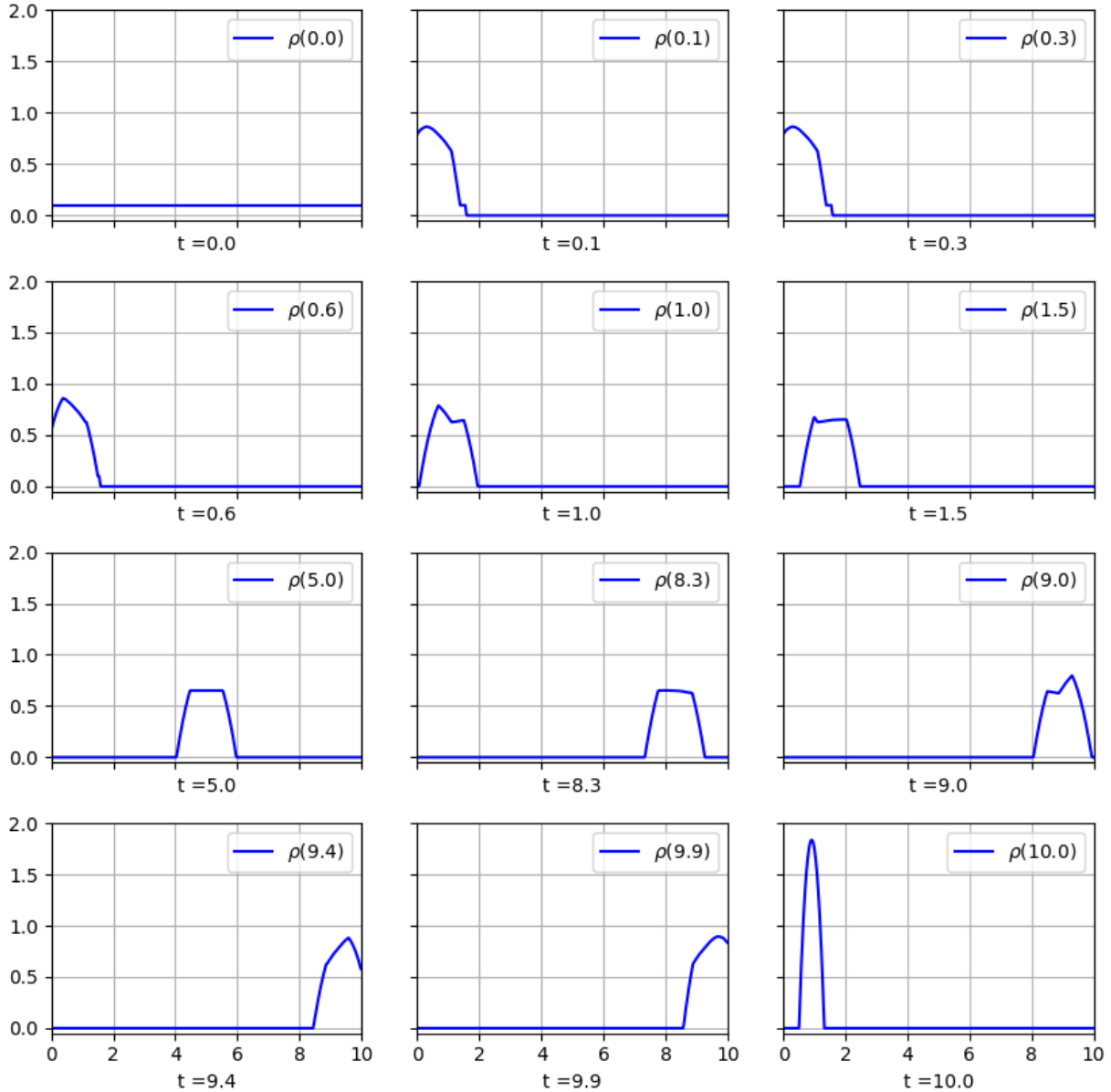


FIGURE 6. The simulation of the solution to (49) at times $t = 0, 0.1, 0.3, 0.6, 1, 1.5, 5, 8.3, 9, 9.4, 9.9, 10$ with $V(t, x) = (t - x)^2$, $m_0(x) = 1/S$, $\varphi_T(x) = x^2$ and parameters from Table 2 except from L which is $7/h$. The blue solid line corresponds to the solution ρ at different times.

- [16] L. I. RUDIN, S. OSHER, AND E. FATEMI Nonlinear total variation based noise removal algorithms. *Physica D.* 60 (1&4): 259&268, (1992).
- [17] F. SANTAMBROGIO *Optimal Transport for Applied Mathematicians*, book, Birkhäuser, Progress in Nonlinear Differential Equations and Their Applications 87, Birkhäuser Basel (2015)
- [18] F. SANTAMBROGIO, Lecture notes on Variational Mean Field Games, *Mean Field Games – Cetraro, Italy, 2019*, Cardaliaguet and Porretta (Eds), Springer, C.I.M.E. Foundation Subseries.
- [19] C. VILLANI *Topics in Optimal Transportation*. Graduate Studies in Mathematics, AMS, 2003.