



HAL
open science

Metric gradient flows with state dependent functionals: the Nash-MFG equilibrium flows and their numerical schemes

Gabriel Turinici

► **To cite this version:**

Gabriel Turinici. Metric gradient flows with state dependent functionals: the Nash-MFG equilibrium flows and their numerical schemes. 2017. hal-01528480v1

HAL Id: hal-01528480

<https://hal.science/hal-01528480v1>

Preprint submitted on 29 May 2017 (v1), last revised 7 Oct 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Metric gradient flows with state dependent functionals: the Nash-MFG equilibrium flows and their numerical schemes

Gabriel Turinici*

May 29, 2017

Abstract

We investigate the convergence of a relaxed version of the *best reply* numerical schemes (also known as *best response* or *fictitious play*) used to find Nash-mean field games equilibriums. This leads us to consider evolution equations in metric spaces similar to gradient flows except that the functional to be differentiated depends on the current point; these are called equilibrium flows. We give two definitions of solutions and prove that as the time step tends to zero the interpolated (à la de Giorgi) numerical curves converge to equilibrium flows. As a by-product we obtain a sufficient condition for the uniqueness of a mean field games equilibrium. We close with applications to congestion and vaccination mean field games.

Keywords: gradient flows; mean field games; vaccination games

1 Introduction

Let \mathcal{X} be a Polish geodesic metric space (see [11] for an introduction to metric spaces) and $\mathcal{C}(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ a functional. We investigate in this work the equation:

$$\partial_t x_t + \nabla_1 \mathcal{C}(x_t, x_t) = 0, \quad x_0 = \bar{x}. \quad (1)$$

Such an equation is called an *equilibrium flow* or *partial flow* for reasons that will be made clear in the sequel.

*Université Paris-Dauphine, PSL Research University, CNRS UMR 7534, CEREMADE, 75016 PARIS, FRANCE & Institut Universitaire de France, gabriel.turinici@dauphine.fr

20 Discrete (numerically computable) versions of this evolution equation are
 21 the numerical schemes defined by the recurrence:

$$x_0^\tau = \bar{x}, \quad x_{k+1}^\tau \in \operatorname{argmin}_{y \in \mathcal{X}} \frac{d(y, x_k^\tau)^2}{2\tau} + \mathcal{C}(y, x_k^\tau), \quad k \geq 0. \quad (2)$$

22 These numerical schemes are relaxed versions of the *best reply / best response*
 23 / *fictitious play* algorithms (see [7]); the original schemes take sometimes
 24 $\tau = \infty$ i.e., omit the first term in (2).

25 Our goal is to give a rigorous definition of the concept of solution of
 26 (1) and show that the numerical schemes (2) converges, when $\tau \rightarrow 0$, to a
 27 solution of (1). Finally we give examples that show that the equilibrium
 28 flows can be successfully used in the study of mean field games equilibriums.

29 1.1 Motivation and literature review

30 The equilibrium in non-cooperative multi-player games are often formulated
 31 as mixed strategy Nash equilibriums (see [29]). The computation of such
 32 equilibriums and the procedure for players to reach them has been the ob-
 33 ject of many contributions and give rise to several proposals e.g., replicator
 34 dynamics and fictitious play / best reply / best response dynamics, see [14]
 35 for details.

36 The relatively recent introduction of the mean field games (abbreviated
 37 from now on as MFG) by Lasry and Lions [25, 24, 26] and conjointly by
 38 Huang, Malhamé and Caines [17, 16] (see also [6] for entry points to the
 39 litterature) allow to extend this concept to games with an infinite number of
 40 players. In this context the players are considered similar (or decomposed in
 41 several classes, each with an infinite number of individuals) and an equilib-
 42 rium is attained when all agents in a class use same mixed strategy, which
 43 is optimal in the Nash sense. Mixed strategies are probability measures over
 44 the state of pure strategies and thus form a metric space (we will come back
 45 later to the topological description of that space); in order to gain in gen-
 46 erality we will suppose from now on that the space of all mixed strategies
 47 is metrizable and will be denoted \mathcal{X} . The cost of the individual strategy x
 48 depends on the choice of everybody else's strategy $y \in \mathcal{X}$ and is encoded
 49 through the cost function $\mathcal{C}(x, y)$. A MFG equilibrium is thus a point $x \in \mathcal{X}$
 50 such that

$$\mathcal{C}(x, x) \leq \mathcal{C}(z, x), \quad \forall z \in \mathcal{X}. \quad (3)$$

51 In this context, the relaxed best reply algorithm, which corresponds to
 52 (2) has been proposed and tested (see e.g., [7, 8]) with successful results.
 53 However only very few works concern the behavior of solutions for $\tau \rightarrow 0$ in

54 the general framework of metric spaces or the meaning to be given to the
 55 limit equation (1).

56 Note that when \mathcal{C} is independent of the second argument, i.e.,

$$\mathcal{C}(x, y) = E(x), \quad (4)$$

57 the relation (2) becomes the celebrated implicit Euler-type scheme of Jordan,
 58 Kinderlehrer and Otto [19] for the definition of gradient flows in metric spaces

$$\partial_t y_t + \nabla E(y_t) = 0, \quad y_0 = \bar{y}, \quad (5)$$

59 and received considerable attention (see [32, 31, 2] for instance). However,
 60 the situation when E has dependence on other variables has not been treated
 61 to the same extent and the related contributions involve gradient flows of
 62 time dependent functionals $\mathcal{E}(t, u)$ with a known dependence on time (see
 63 [20, 12, 30, 27, 28]). Of course, formally one can set

$$\mathcal{E}(t, u) = \mathcal{C}(u, x_t). \quad (6)$$

64 However the cited papers use implicit (in time) optimization

$$y_0^\tau = \bar{x}, \quad y_{k+1}^\tau \in \operatorname{argmin}_{y \in \mathcal{X}} \frac{d(y, y_k^\tau)^2}{2\tau} + \mathcal{E}((k+1)\tau, y_k^\tau), \quad k \geq 0, \quad (7)$$

65 which are not equivalent to the numerical scheme (2) because (7) is fully
 66 implicit while the (relaxed) best-reply scheme is semi-explicit. Further-
 67 more, technical assumptions invoked in previous works assume that the time-
 68 dependence of $\mathcal{E}(t, u)$ is smooth (e.g., in [30, assumption (2.19c) page 109])
 69 or that its differentiability points have additional properties (for instance
 70 independent of x in [12, assumption E3.1], [20, assumption A4]) while here
 71 the mere absolute continuity of $t \mapsto x_t$ does not allow to fulfill *a priori* such
 72 assumptions. Nevertheless our contribution owes much to all these previous
 73 works which it translates to this specific setting.

74 **2 Theoretical results**

75 **2.1 Basic reminders and motivation**

76 The absence of a vector operations in a metric space does no allow to develop
 77 fully a differential calculus and requires adaptation of notions of derivative.
 78 Accordingly the definition of evolution equations have to use alternative prop-
 79 erties.

80 We recall below the main ideas of such an alternative formulation (see [1])
 81 for the particular case (4)-(5); suppose for a moment that \mathcal{X} is an Euclidian
 82 space and E a smooth (C^1 or above) function; then:

$$\frac{d}{dt}E(x_t) = \langle \nabla E(x_t), x'_t \rangle \geq -|\nabla E(x_t)| \cdot |x'_t| \geq -\frac{1}{2}|x'_t|^2 - \frac{1}{2}|\nabla E|^2(x_t),$$

83 or equivalently,

$$\frac{d}{dt}E(x_t) + \frac{1}{2}|x'_t|^2 + \frac{1}{2}|\nabla E|^2(x_t) \geq 0 \quad \forall t, \quad (8)$$

84 with equality only if x is solution of (5). Therefore asking that

$$\frac{d}{dt}E(x_t) + \frac{1}{2}|x'_t|^2 + \frac{1}{2}|\nabla E|^2(x_t) \leq 0 \quad \forall t, \quad (9)$$

85 is an equivalent characterization of (5) (more precisely called the EDI for-
 86 mulation). The integral form can also be used:

$$E(x_b) - E(x_a) + \int_a^b \left(\frac{1}{2}|x'_t|^2 + \frac{1}{2}|\nabla E|^2(x_t) \right) dt \leq 0 \quad \forall 0 \leq a \leq b. \quad (10)$$

87 The advantage of formulation (10) is that it only uses quantities that can
 88 be defined in a metric space (see below for definition of $|x'_t|$ and $|\nabla E|$). The
 89 corresponding computation for a bi-variate functional \mathcal{C} is:

$$\int_a^b \left(\frac{d}{dt}\mathcal{C}(x_t, \nu) \Big|_{\nu=x_t} \right) dt + \int_a^b \left(\frac{1}{2}|x'_t|^2 + \frac{1}{2}|\nabla_1 \mathcal{C}|^2(x_t, x_t) \right) dt \leq 0 \quad \forall 0 \leq a \leq b. \quad (11)$$

90 However this formulation poses specific problems (see also section 3.1) as in
 91 general the solution $(x_t)_{t \geq 0}$ is only absolutely continuous (with respect to
 92 time) while the manipulation of the term $\frac{d}{dt}\mathcal{C}(x_t, \nu) \Big|_{\nu=x_t}$ requires additional
 93 assumptions. This is the object of the next section.

94 Before that, let us recall the following definition:

95 **Definition 1** A curve $x : [0, T] \rightarrow (\mathcal{X}, d)$ is called absolutely continuous if
 96 there exists $f \in L^1(0, T)$ such that

$$d(x_{t_1}, x_{t_2}) \leq \int_{t_1}^{t_2} f(t) dt, \quad \forall t_1 < t_2, \quad t_1, t_2 \in [0, T]. \quad (12)$$

97 For an absolutely continuous curve $(x_t)_{t \in [0, T]}$ the metric derivative of x at r
 98 defined by

$$|x'_r| = \lim_{h \rightarrow 0} \frac{d(x_{r+h}, x_r)}{|h|}, \quad (13)$$

99 exists a.e., belongs to $L^1(0, T)$ and is the smallest L^1 function that verifies
 100 (12).

101 **2.2 Definition of EDI / EVI equilibrium flows**

102 Let us denote

$$\mathcal{D}^s(\mathcal{C}) = \{x \in \mathcal{X} | \mathcal{C}(x, x) < \infty\}. \quad (14)$$

103 We suppose from now on that \mathcal{C} satisfies the assumption:

104 **(A₁)** There exists $C_1 < \infty$ such that $\mathcal{C}(y, x) \geq -C_1, \forall x, y \in \mathcal{X}$.

105 For any $\alpha, \beta \in \mathbb{R}, \alpha \leq \beta$, we denote by $\mathcal{S}(\alpha, \beta)$ the set of divisions of the
 106 interval $[\alpha, \beta]$. Let $x = (x_t)_{t \in [0, T]} \subset \mathcal{D}^s(\mathcal{C})$ be an absolutely continuous curve
 107 in \mathcal{X} ; define for $0 \leq a \leq b \leq T$ and a division $\Delta = \{a = t_0 < t_1 < \dots < t_{N_\Delta} =$
 108 $b\} \in \mathcal{S}(a, b)$:

$$\Upsilon(\Delta; x, a, b) = \sum_k \mathcal{C}(x_{k+1}, x_k) - \mathcal{C}(x_k, x_k). \quad (15)$$

109

$$\Upsilon(x, a, b) = \liminf_{\Delta \in \mathcal{S}(a, b), |\Delta| \rightarrow 0} \Upsilon(\Delta; x, a, b). \quad (16)$$

110 **Remark 1** When \mathcal{X} is e.g., Euclidian and under regularity assumptions on
 111 \mathcal{C} it is easy to check that $\Upsilon(x, a, b) = \int_a^b \frac{d}{dt} \mathcal{C}(x_t, \nu) \Big|_{\nu=x_t} dt$.

112 **Definition 2 (EDI equilibrium flow)** An absolutely continuous curve $(x_t)_{t \in [0, T]}$
 113 is called an EDI-equilibrium flow starting from \bar{x} if $\lim_{t \rightarrow 0} x_t = \bar{x}$ and:

$$\forall s \geq 0, \quad \Upsilon(x, 0, s) + \frac{1}{2} \int_0^s |x'_r|^2 dr + \frac{1}{2} \int_0^s |\nabla_1 \mathcal{C}|^2(x_r, x_r) dr \leq 0, \quad (17)$$

$$\text{a.e. } t > 0, \quad \forall s \geq t, \quad \Upsilon(x, t, s) + \frac{1}{2} \int_t^s |x'_r|^2 dr + \frac{1}{2} \int_t^s |\nabla_1 \mathcal{C}|^2(x_r, x_r) dr \leq 0, \quad (18)$$

where, for any $y \in \mathcal{D}^s(\mathcal{C})$ and any point (x, y) with $\mathcal{C}(x, y) < \infty$, the slope $|\nabla_1 \mathcal{C}|(x, y)$ of $\mathcal{C}(\cdot, \cdot)$ with respect to the first argument evaluated at (x, y) is:

$$|\nabla_1 \mathcal{C}|(x, y) = \limsup_{z \rightarrow x} \frac{(\mathcal{C}(x, y) - \mathcal{C}(z, y))^+}{d(x, z)} = \max \left\{ \limsup_{z \rightarrow x} \frac{\mathcal{C}(x, y) - \mathcal{C}(z, y)}{d(x, z)}, 0 \right\}.$$

114 **Remark 2** For the particular case (4) the definition above coincides with
 115 the definition of a EDI gradient flow, see [1]. Moreover an equilibrium flow
 116 for $\mathcal{C}(x, y)$ is also an equilibrium flow for $\mathcal{C}(x, y) + G(y)$ for any function G .

117 **Remark 3** A natural question is if there exist functions \mathcal{C} that satisfy the
 118 above assumptions and that cannot be treated with developments in previous
 119 works [20, 12]; the answer is positive, see for instance [20, Example 3 page

120 11]) for $\mathcal{X} = \mathbb{R}$, $d(x, y) = |x - y|$, $\mathcal{C}(x, y) = d(x, y)$ (and also $d(x, y) + \mathcal{F}(x) +$
121 $G(y)$ with $\mathcal{F}(x)$ smooth). One cannot use previous theories because setting
122 $\mathcal{E}(t, u) = \mathcal{C}(u, x_t) = |u - x_t|$ the derivative with respect to t exists except when
123 $u = x_t$, or it is exactly there that it should be used (see [12, equation 2.3] and
124 the curve $x_t = t$). See also section 3.1 below.

125 As the previous works concerned a different setting we do not claim that
126 this contribution is a generalization but rather an extension of these results
127 for our situation.

128 When \mathcal{C} has further convexity properties one can adapt the EVI formu-
129 lation as in [20, 21]. We introduce the following assumption (corresponding
130 to [1, Assumption 4.24 page 77]):

131 **(A₂)** Suppose \mathcal{C} is a lower semicontinuous functional with respect to the first
132 argument and there exists $\lambda \in \mathbb{R}$ such that for any $x_0, x_1, z, v \in \mathcal{X}$ there
133 exists a curve γ connecting x_0 and x_1 such that for all $s \in [0, 1]$:

$$\mathcal{C}(\gamma(s), z) \leq (1 - s)\mathcal{C}(x_0, z) + s\mathcal{C}(x_1, z) - \lambda \frac{(1 - s)s}{2} d^2(x_0, x_1). \quad (19)$$

134

$$d^2(\gamma(s), v) \leq (1 - s)d^2(x_0, v) + sd^2(x_1, v) - (1 - s)s d^2(x_0, x_1). \quad (20)$$

135 In general a function (of one or several variables) that satisfies (19)-(20)
136 with respect to one of its variables is called λ -convex in that variable.

137 **Definition 3 (EVI equilibrium flow)** An absolutely continuous curve $(x_t)_{t \in [0, T]}$
138 is called an EVI-equilibrium flow starting from \bar{x} if $\lim_{t \rightarrow 0} x_t = \bar{x}$ and :

$$\mathcal{C}(x_t, x_t) + \frac{1}{2} \frac{d}{dt} d^2(x_t, y) + \frac{\lambda}{2} d^2(x_t, y) \leq \mathcal{C}(y, x_t), \forall y, \text{ a.e. } t \geq 0. \quad (21)$$

139 Note that the definition is valid because if $(x_t)_{t \in [0, T]}$ is absolutely contin-
140 uous then $t \mapsto d^2(x_t, y)$ is also absolutely continuous thus differentiable a.e.
141 with respect to t .

142 2.3 Convergence of numerical schemes: general situa- 143 tion

144 Let us denote

$$\mathcal{M}(x, \tau) = \operatorname{argmin}_{y \in \mathcal{X}} \frac{d(y, x)^2}{2\tau} + \mathcal{C}(y, x). \quad (22)$$

145 With this definition the relaxed *best reply* (or *fictitious play*) (see e.g., [7, 8])
146 numerical scheme in equation (2) can be written as $x_{k+1}^\tau \in \mathcal{M}(x_k^\tau, \tau)$.

147 We investigate in this section whether when $\tau \rightarrow 0$ the set $\{x_k^\tau, k \geq 1\}$
 148 converges to a solution of (1) as defined in (17)-(18) or (21).

149 In order to work with meaningful objects, we introduce the following
 150 assumption which is the analogue of [1, Assumption 4.8 page 67]:

151 **(A₃)** There exists $\bar{\tau} > 0$ such that for any $\tau \leq \bar{\tau}$ and $\bar{x} \in \mathcal{D}^s(\mathcal{C})$:

$$\mathcal{M}(\bar{x}, \tau) \neq \emptyset. \quad (23)$$

152

153 Assuming that assumption **(A₃)** is satisfied, we can define the interpola-
 154 tion á la de Giorgi (see [1]) which is a curve $t \in [0, T] \mapsto x_t^\tau$ such that $x_0^\tau = \bar{x}$
 155 and

$$x_t^\tau \in \mathcal{M}(x_{k\tau}^\tau, t - k\tau), \forall t \in]k\tau, (k+1)\tau], k \geq 1. \quad (24)$$

156 For such a map, we define the discrete speed $\text{Dsp}^\tau : [0, +\infty) \rightarrow [0, +\infty)$ by

$$\text{Dsp}_t^\tau = \frac{d(x_{n\tau}^\tau, x_{(n+1)\tau}^\tau)}{\tau} \text{ for } t \text{ in } (n\tau, (n+1)\tau),$$

157 and the discrete slope $\text{Dsl}^\tau : [0, +\infty) \rightarrow [0, +\infty)$ by

$$\text{Dsl}_t^\tau = \frac{d(x_{n\tau}^\tau, x_t^\tau)}{t - n\tau} \text{ for } t \text{ in } (n\tau, (n+1)\tau).$$

158 We will need some additional hypothesis:

159 **(A₄)** For any $c \in \mathbb{R}$, $r > 0$ and $x \in \mathcal{X}$ the set $\{y \in \mathcal{X} | \mathcal{C}(y, y) \leq c, d(y, x) \leq r\}$
 160 is compact.

161 **(A₅)** The slope $|\nabla_1 \mathcal{C}|$ is lower semicontinuous.

162 **(A₆)** Lipschitz property: there exists $L > 0$ such that:

$$\mathcal{C}(x, y) \leq Ld(y, z) + \mathcal{C}(x, z), \quad \forall x, y, z \in \mathcal{X}. \quad (25)$$

163 **(A₇)** For any absolutely continuous curve $(x_t)_{t \in [a, b]}$:

$$\Upsilon(x, \mathcal{C}, a, b) \leq \liminf_{|\Delta_n| \rightarrow 0, y_n \rightarrow x, \sup_n \int_a^b |\dot{y}_n(t)| < \infty} \Upsilon(\Delta_n; y_n, \mathcal{C}, a, b), \quad (26)$$

164 where the convergence of the curve y_n to x is in the uniform (on com-
 165 pacts) norm.

166 **Remark 4** *The assumption (\mathbf{A}_7) is satisfied if*

167 (\mathbf{A}_8) *There exists $C_L < \infty$ such that for any $x, y, z \in \mathcal{D}^s(\mathcal{C})$:*

$$|\mathcal{C}(x, y) + \mathcal{C}(y, z) - \mathcal{C}(x, z) - \mathcal{C}(y, y)| \leq C_L d(x, y) d(y, z). \quad (27)$$

168 *In this case*

$$\Upsilon(x, \mathcal{C}, a, b) = \lim_{|\Delta| \rightarrow 0} \Upsilon(\Delta; x, \mathcal{C}, a, b) = \lim_{|\Delta_n| \rightarrow 0, y_n \rightarrow x, \sup_n \int_a^b |\dot{y}_n(t)| < \infty} \Upsilon(\Delta_n; y_n, \mathcal{C}, a, b). \quad (28)$$

169 *In particular note that $\mathcal{C}(x, y) = d(x, y)$ satisfies (\mathbf{A}_7) but does **not** sat-*
 170 *isfy (28) (thus neither (\mathbf{A}_8)) while $\mathcal{C}(x, y) = d^2(x, y)$ satisfies (\mathbf{A}_8) with*
 171 *$C_L = 2$. On the other hand for $\mathcal{X} = \mathbb{R}$ the mapping $(x, y) \mapsto \mathcal{C}(x, y) =$*
 172 *$\int_0^x \int_0^y f(s, t) ds dt$, with f a bounded function will also satisfy assumption*
 173 *(\mathbf{A}_8) .*

174 Assumptions $(\mathbf{A}_1), (\mathbf{A}_4), (\mathbf{A}_5)$ are classical (see for instance [1, Assump-

175 tion 4.13 page 69]) and (\mathbf{A}_7) belongs to the same class; on the other hand,

176 the assumption (\mathbf{A}_6) is the analogue of [20, assumption A4 page 11] through

177 the correspondence in equation (6).

178 The properties of the curves obtained by the numerical scheme (2) are

179 detailed in the following result.

180 **Theorem 1** *Let \mathcal{C} satisfying assumptions $(\mathbf{A}_1), (\mathbf{A}_3), (\mathbf{A}_4), (\mathbf{A}_5), (\mathbf{A}_6)$*
 181 *and (\mathbf{A}_7) . Take $x_0 \in \mathcal{D}^s(\mathcal{C})$. Then the set of curves $\{(x_t^\tau)_{t \in [0, T]}; \tau \leq \bar{\tau}\}$*
 182 *defined in (24) is relatively compact in the set of curves in \mathcal{X} with local*
 183 *uniform convergence and any limit curve is an EDI equilibrium flow in the*
 184 *sense of Definition 2.*

185 **Proof** We follow, when possible, the proof of [1, Theorem 4.14 page 69]; first

186 we proceed as in [1, Theorem 4.9 page 67] and note that for fixed $x \in \mathcal{X}$,

187 $\tau \leq \bar{\tau}$ and $x_\tau \in \mathcal{M}(x, \tau)$, the map $\tau \mapsto \frac{d(x_\tau, x)^2}{2\tau} + \mathcal{C}(x_\tau, x)$, is locally Lipschitz

188 and

$$\frac{d}{d\tau} \left(\frac{d(x_\tau, x)^2}{2\tau} + \mathcal{C}(x_\tau, x) \right) = -\frac{d^2(x, x_\tau)}{2\tau^2}. \quad (29)$$

189 The proof of this identity follow precisely the reference and the details are

190 left to the reader.

191 The minimal property of $x_{(k+1)\tau}^\tau$ implies that for any k : $\mathcal{C}(x_{(k+1)\tau}^\tau, x_{k\tau}^\tau) -$

192 $\mathcal{C}(x_{k\tau}^\tau, x_{k\tau}^\tau) \leq 0$; combined with (25) this shows that $\mathcal{C}(x_{(k+1)\tau}^\tau, x_{(k+1)\tau}^\tau) < \infty$.

193 On the other hand, as in [1, Lemma 4.10 page 67] we obtain (again we
 194 skip the details) that $\tau \mapsto d(x, x_\tau)$ is non decreasing and

$$|\nabla_1 \mathcal{C}|(x_\tau, x) \leq \frac{d(x, x_\tau)}{2\tau}. \quad (30)$$

195 Integrating (29) we obtain the following identity:

$$\sum_{k=n}^{m-1} \mathcal{C}(x_{(k+1)\tau}^\tau, x_{k\tau}^\tau) - \mathcal{C}(x_{k\tau}^\tau, x_{k\tau}^\tau) + \frac{1}{2} \int_{n\tau}^{m\tau} (\text{Dsp}_r^\tau)^2 dr + \frac{1}{2} \int_{n\tau}^{m\tau} (\text{Dsl}_r^\tau)^2 dr = 0, \quad (31)$$

196 or equivalently, denoting Δ^τ the division of $[0, T]$ containing the points $k\tau$
 197 ($1 \leq k \leq T/\tau$):

$$\Upsilon(\Delta^\tau, x^\tau, n\tau, m\tau) + \frac{1}{2} \int_{n\tau}^{m\tau} (\text{Dsp}_r^\tau)^2 dr + \frac{1}{2} \int_{n\tau}^{m\tau} (\text{Dsl}_r^\tau)^2 dr = 0. \quad (32)$$

198 In particular for any $k \in \mathbb{N}^*$:

$$\begin{aligned} & \left(\sum_{\ell=0}^k d(x_{\ell\tau}^\tau, x_{\ell+1\tau}^\tau) \right)^2 \leq \left(\int_0^{(k+1)\tau} |\text{Dsp}_r^\tau| dr \right)^2 \leq (k+1)\tau \int_0^{(k+1)\tau} |\text{Dsp}_r^\tau|^2 dr \\ & \leq -2(k+1)\tau \Upsilon(\Delta^\tau; x^\tau, 0, (k+1)\tau) \\ & \leq -2(k+1)\tau \left(\mathcal{C}(\bar{x}, \bar{x}) - \mathcal{C}(x_{(k+1)\tau}^\tau, x_{(k+1)\tau}^\tau) + L \sum_{\ell=0}^k d(x_{\ell\tau}^\tau, x_{\ell+1\tau}^\tau) \right). \end{aligned} \quad (33)$$

199 This shows that in particular $\sum_{\ell=0}^k d(x_{\ell\tau}^\tau, x_{\ell+1\tau}^\tau)$ is bounded by a constant
 200 depending on $T, L, \mathcal{C}(\bar{x}, \bar{x})$ and C_1 but independent on τ . On the other hand
 201 for $t \in]k\tau, (k+1)\tau]$, $t \leq T$:

$$d^2(x_t^\tau, \bar{x}) \leq \left(d(x_t^\tau, x_{k\tau}^\tau) + \sum_{\ell=0}^{k-1} d(x_{\ell\tau}^\tau, x_{\ell+1\tau}^\tau) \right)^2 \leq \left(\sum_{\ell=0}^k d(x_{\ell\tau}^\tau, x_{\ell+1\tau}^\tau) \right)^2,$$

202 where we used that $\tau \mapsto d(x_t^\tau, x_{k\tau}^\tau)$ is non decreasing. Thus the set of curves
 203 $\{(x_t^\tau)_{t \in [0, T]}; \tau \leq \bar{\tau}\}$ is uniformly bounded with respect to τ . As a by-product
 204 $\Upsilon(\Delta^\tau; x^\tau, n\tau, m\tau)$ is bounded uniformly with respect to τ , and $n, m \leq T/\tau$.

205 A similar estimation starting from:

$$d^2(x_{n\tau}^\tau, x_{n\tau}^\tau) \leq \left(\int_{n\tau}^{m\tau} |\text{Dsp}_r^\tau| dr \right)^2 \leq (n\tau - m\tau) \int_{n\tau}^{m\tau} |\text{Dsp}_r^\tau|^2 dr,$$

206 allows to see that the set of curves is also equicontinuous. By Arzela-Ascoli
 207 one obtains the relative compactness.

208 Let now $\tau_n \downarrow 0$ and $(x_t)_{t \in [0, T]}$ a limit curve of $\{(x_t^{\tau_n})_{t \in [0, T]}; n \geq 1\}$. From
 209 (34) one obtains that $(x_t)_{t \in [0, T]}$ is absolutely continuous and for $t \leq s \leq T$:

$$\int_t^s |\dot{x}_r|^2 dr \leq \liminf_{n \rightarrow \infty} \int_t^s |\text{Dsp}_r^{\tau_n}|^2 dr. \quad (34)$$

210 On the other hand, from the lower semicontinuity of $|\nabla_1 \mathcal{C}|$ and (30) we obtain

$$|\nabla_1 \mathcal{C}|(x_t, x_t) \leq \liminf_{n \rightarrow \infty} |\nabla_1 \mathcal{C}|(x_t^{\tau_n}, x_{k_n \tau_n}^{\tau_n}) \leq \liminf_{n \rightarrow \infty} \text{Dsl}_t^{\tau_n}. \quad (35)$$

211 where k_n is such that $t \in]k_n \tau_n, (k_n + 1) \tau_n]$.

212 Using **(A₇)** and similar arguments as in the end of the proof of [1, The-
 213 orem 4.14 page 69] (details are left to the reader) one can pass to the limit
 214 in (32) and obtain relations (17)-(18). ■

215 2.4 Convergence of numerical schemes: convex case

216 For the convex case more precise information can be obtained and is gathered
 217 in the following theorem.

218 **Theorem 2** *Let \mathcal{C} satisfying assumptions **(A₁)**, **(A₂)**, **(A₄)** and **(A₆)**.*

- 219 1. *For every $x \in \mathcal{D}^s(\mathcal{C})$, the set $\mathcal{M}(x, \tau)$ contains exactly one element for*
 220 *any $\tau \leq 1/\lambda^-$. Moreover, for $x_0 \in \mathcal{D}^s(\mathcal{C})$ the set of curves $\{(x_t^\tau)_{t \in [0, T]}; \tau \leq$
 221 $\bar{\tau}\}$ defined in (24) is relatively compact and any limit curve $(x_t)_{t \in [0, T]}$
 222 *is an EVI equilibrium flow in the sense of Definition 3.**
- 223 2. *Suppose in addition that for any $x \in \mathcal{D}^s(\mathcal{C})$ and $u, v \in \mathcal{D}^s(\mathcal{C})$ in a*
 224 *neighborhood of x :*

$$|C(u, v) + C(v, u) - C(u, u) - C(v, v)| = o(d(u, v)). \quad (36)$$

225 *Then the (EVI) equilibrium flow is unique.*

- 226 3. *If the stronger assumption **(A₈)** (instead of (36)) is satisfied, then for*
 227 *any two EVI equilibrium flows $(x_t)_{t \in [0, T]}$ and $(y_t)_{t \in [0, T]}$ starting from*
 228 *$\bar{x} \in \mathcal{D}^s(\mathcal{C})$ and $\bar{y} \in \mathcal{D}^s(\mathcal{C})$ respectively:*

$$d(x_t, y_t) \leq e^{-(\lambda - C_L)t} d(\bar{x}, \bar{y}), \text{ a.e., } t \geq 0. \quad (37)$$

- 229 4. *Under assumption **(A₈)** if $\lambda > C_L$ then the mean field game with cost*
 230 *functional $\mathcal{C}(\cdot, \cdot)$ has an unique Nash-MFG equilibrium $x^{\mathcal{C}}$ and any equi-*
 231 *librium flow $(x_t)_{t \geq 0}$ starting from some $x(0) = \bar{x} \in \mathcal{D}^s(\mathcal{C})$ converges*
 232 *exponentially fast to $x^{\mathcal{C}}$.*

233 **Proof**

234 **Proof of item 1:** Following the proof of [1, Theorem 4.25 page 77] it is
 235 possible to prove the existence and uniqueness of the minimizers in $\mathcal{M}(x, \tau)$
 236 using the same arguments i.e. taking a minimizing sequence $(x_n)_{n \geq 1}$ and
 237 using assumption **(A₂)** for $s = 1/2$, $x_0 = x_n$, $x_1 = x_m$, $y = x$, $v = x$.
 238 Moreover, the following estimation (replacing [1, equation (69) page 78]) is
 239 obtained using arguments in the same proof as soon as $x^\tau \in \mathcal{M}(x, \tau)$:

$$\frac{d^2(x^\tau, y) - d^2(x, y)}{2\tau} + \frac{\lambda}{2} d^2(x^\tau, y) \leq \mathcal{C}(y, x) - \mathcal{C}(x^\tau, x), \quad \forall y \in \mathcal{X}. \quad (38)$$

240 Summing up such estimations it follows that for $t = n\tau < m\tau = s$:

$$\begin{aligned} & \frac{d^2(x_t^\tau, y) - d^2(x_s^\tau, y)}{2(s-t)} + \frac{\lambda\tau}{2(s-t)} \sum_{\ell=n}^{m-1} d^2(x_{(\ell+1)\tau}^\tau, y) \\ & \leq \frac{\tau}{s-t} \left(\sum_{\ell=n}^{m-1} \mathcal{C}(y, x_{\ell\tau}^\tau) - \mathcal{C}(x_{(\ell+1)\tau}^\tau, x_{\ell\tau}^\tau) \right), \quad \forall y \in \mathcal{X}. \end{aligned} \quad (39)$$

241 We can prove as in theorem 1 that the curves $\{x_t^\tau\}_{\tau \leq 1/\lambda}$ form a relatively
 242 compact set. Taking $\tau_n \downarrow 0$ one obtains for $0 \leq t < s \leq T$:

$$\begin{aligned} & \frac{d^2(x_t, y) - d^2(x_s, y)}{2(s-t)} + \frac{\lambda}{2(s-t)} \int_t^s d^2(x_r, y) dr \\ & \leq \frac{1}{s-t} \left(\int_t^s [\mathcal{C}(y, x_r) - \mathcal{C}(x_r, x_r)] dr \right), \quad \forall y \in \mathcal{X}, \end{aligned} \quad (40)$$

243 which is the integral form of the definition 3 of a EVI equilibrium flow. In the
 244 passage from (39) to (40) the terms $\sum_{\ell=n}^{m-1} d^2(x_{(\ell+1)\tau}^\tau, y)$ and $\sum_{\ell=n}^{m-1} \mathcal{C}(y, x_{\ell\tau}^\tau)$
 245 are, thanks to the Lipschitz property of $\mathcal{C}(y, \cdot)$ and $d^2(\cdot, y)$, arbitrarily close
 246 to Riemann sums of the corresponding integrals. On the other hand for the
 247 term $\sum_{\ell=n}^{m-1} \mathcal{C}(x_{(\ell+1)\tau}^\tau, x_{\ell\tau}^\tau)$ the inequality is obtained from the Fatou lemma
 248 applied to $f_n(r) = \sum_{\ell=n}^{m-1} 1_{[\ell\tau_n, (\ell+1)\tau_n]}(r) \mathcal{C}(x_{(\ell+1)\tau_n}^\tau, x_{\ell\tau_n}^\tau)$ (recall that we have
 249 a lower bound on \mathcal{C} from assumption **(A₁)**) and the lower semi-continuity of
 250 \mathcal{C} .

251 **Proof of item 3** To prove contraction (37) we use the EVI inequality for x_t
 252 and y_t and write:

$$\begin{aligned} & \frac{1}{2} \frac{d}{ds} d^2(x_s, y_t) \Big|_{s=t} + \frac{\lambda}{2} d^2(x_t, y_t) \leq \mathcal{C}(y_t, x_t) - \mathcal{C}(x_t, x_t) \quad a.e. \quad t \geq 0. \\ & \frac{1}{2} \frac{d}{ds} d^2(y_s, x_t) \Big|_{s=t} + \frac{\lambda}{2} d^2(y_t, x_t) \leq \mathcal{C}(x_t, y_t) - \mathcal{C}(y_t, y_t) \quad a.e. \quad t \geq 0. \end{aligned} \quad (41)$$

253 Adding up the two identities and invoking (27) for $x = z = x_t$, $y = y_t$ we
 254 obtain (heuristically for the moment):

$$\frac{1}{2} \frac{d}{dt} d^2(x_t, y_t) + \lambda d^2(x_t, y_t) \leq C_L d^2(x_t, y_t) \text{ a.e. } t \geq 0, \quad (42)$$

255 and (37) follows. In practice, the manipulations above can be made precise
 256 by a doubling of variables argument (see [1, page 80], [22], [9, pages 179,
 257 183]).

258 **Proof of item 2:** the proof is similar to that of item 3 and is left as an
 259 exercise to the reader. The key ingredient is the [2, Lemma 4.1.8] which we
 260 recall below:

261 **Lemma 1** *Let $f : [0, \infty[\rightarrow \mathbb{R}$ be a locally absolutely continuous function,*
 262 *and $a, b \in L^1_{loc}([0, \infty[)$ be given functions satisfying for $\lambda \in \mathbb{R}$:*

$$\frac{d}{dt} f^2(t) + 2\lambda f^2(t) \leq a(t) + 2b(t)f(t), \text{ a.e., } t > 0. \quad (43)$$

263 *Then for any $T > 0$*

$$e^{\lambda T} |f(T)| \leq \sqrt{f^2(0) + \sup_{t \in [0, T]} \int_0^t e^{2\lambda s} a(s) ds} + 2 \int_0^T e^{\lambda t} |b(t)| dt. \quad (44)$$

264 **Proof of item 4:** the proof is a consequence of item 3 as soon as one proves
 265 that the curve $y_t = x^c$ is an EVI equilibrium flow. To prove this, note that
 266 the definition (3) of a Nash-MFG equilibrium implies $\mathcal{M}(x^c, \tau) = \{x^c\}$ for
 267 all $\tau > 0$. Then $x_t^\tau = x^c$ for any $\tau, t \geq 0$ thus the (only) limit curve is $y_t = x^c$
 268 and by item 1 it is an EVI equilibrium flow. ■

269 **Remark 5** *The item 4 is interesting because it gives a different route to*
 270 *ensure uniqueness of mean field games equilibriums, very few such procedures*
 271 *exist in the literature.*

272 3 Applications

273 For the notations used in this section see the appendix A.

274 3.1 The treatment of the term $\int_a^b \frac{d}{dt} \mathcal{C}(x_t, \nu) \Big|_{\nu=x_t} dt$

275 We consider first an example that shows how our proposed methodology
 276 avoids problems with the term $\int_a^b \frac{d}{dt} \mathcal{C}(x_t, \nu) \Big|_{\nu=x_t} dt$ which does not always

277 has a well-defined meaning, even when the dynamics is C^∞ with respect to
 278 time.

279 Set $\mathcal{X} = \mathbb{R}$ and $\mathcal{C}(x, y) = \frac{x^2}{2} + \theta|x - y|$ for some fixed $\theta > 0$. After
 280 elementary computations we obtain from the definition (3) that the set of
 281 Nash equilibriums is the entire interval $[-\theta, \theta]$. In particular there is no
 282 uniqueness; we check thus the assumptions of the theorem 1:

- 283 • (\mathbf{A}_1) is valid with $C_1 = 0$;
- 284 • (\mathbf{A}_3) is valid because we minimize a continuous coercive real function;
- 285 • (\mathbf{A}_4) is valid trivially on \mathbb{R} ;
- 286 • (\mathbf{A}_5) results from the computation $|\nabla_1 \mathcal{C}|(x, y) = (x - \theta \cdot 1_{x \leq y} + \theta \cdot 1_{x > y})_+$
- 287 • (\mathbf{A}_6) is valid with $L = \theta$;
- 288 • (\mathbf{A}_7) results from the formula $\Upsilon(x, a, b) = \frac{x_b^2 - x_a^2}{2} + \text{length}(x, a, b)$ (with
 289 $\text{length}(x, a, b)$ being the length of the curve $(x_t)_{t \geq 0}$ between $t = a$ and
 290 $t = b$); note that (\mathbf{A}_8) does not hold in this case.

291 We can apply the theorem 1 to conclude that at least an equilibrium flow
 292 exists. In fact, a more careful inspection reveals that the flow is:

- 293 - when $|x_0| \leq \theta$: the flow is the constant curve $x_t = x_0$;
- 294 - when $x_0 > \theta$: $x_t = \theta + e^{-t}(x_0 - \theta)$;
- 295 - when $x_0 < \theta$: $x_t = -\theta + e^{-t}(x_0 + \theta)$.

296 The dynamics is thus smooth; however the term $\int_a^b \frac{d}{dt} \mathcal{C}(x_t, \nu) \Big|_{\nu=x_t} dt$ can-
 297 not be given any meaning because of the non-differentiability with respect
 298 to t of $|x_t - x_s|$ in $s = t$.

299 **Remark 6** *In this case the assumption (\mathbf{A}_2) is also valid thus one can ap-
 300 ply the first point of the theorem 2 to conclude to the existence of a EVI
 301 - equilibrium flow. However one can propose examples of \mathcal{C} being the sum
 302 of a smooth, coercive but non-convex function of x (as $(x^2 - 1)^2$) and the
 303 term $\theta|x - y|$ that can still check the hypothesis of theorem 1 but not those of
 304 the theorem 2. Here we took the first term to be quadratic in order to have
 305 explicit analytic solutions.*

306 **3.2 Multiple agent types games with potential and in- 307 teraction terms**

308 We follow here the framework in [7] and in particular show how their Corro-
 309 lary 5.2 (corresponding to their example in equation (5.9) Section 5) can be

310 treated with tools provided in this work. In this situation, the players are
 311 assigned to one of several types $x \in X$; once an agent is given a type x it
 312 cannot change it. As a consequence, the probability measure on the set X
 313 is fixed, let us denote it by $\mu \in \mathcal{P}(X)$. Each agent can take actions y in a
 314 set Y . Both X and Y are supposed compact and moreover Y is a convex,
 315 closed subset of an Euclidian space. A strategy is a joint probability measure
 316 $\gamma \in \mathcal{P}(X \times Y)$ with $\pi_{\#}^1 \gamma = \mu$. We assign:

$$\tilde{\mathcal{X}} = \{ \gamma \in \mathcal{P}(X \times Y) | \pi_{\#}^1 \gamma = \mu \}. \quad (45)$$

317 The cost of the action y for the individual of type $x \in X$ when everybody
 318 else is doing γ is $C(x, y, \gamma)$. Denote $\mathcal{C}(\gamma_1, \gamma_2) = \int_{X \times Y} C(x, y, \gamma_2) \gamma_1(dx, dy)$.
 319 The Nash equilibrium (termed Cournot-Nash in [7]) is defined as in equation
 320 (3).

321 Note that by the disintegration theorem [2, Theorem 5.3.1 page 121] for
 322 any $\gamma \in \tilde{\mathcal{X}}$ there exist a a unique (μ - a.e.) set $(\gamma_x)_{x \in X} \subset \mathcal{P}(Y)$ such that
 323 for any Borel function f :

$$\int_{X \times Y} f(x, y) \gamma(dx, dy) = \int_X \left(\int_Y f(x, y) \gamma_x(dy) \right) \mu(dx). \quad (46)$$

324 When there is no ambiguity, we will use the notation $(\gamma_x)_{x \in X} \subset \mathcal{P}(Y)$ to
 325 designate the disintegration of the measure γ along the elements $x \in X$.

326 The measure γ_x can be interpreted as the strategy of the agents of type
 327 x .

328 **Remark 7** In [7] a specific situation is considered when the cost does not
 329 depend on the type of other agents choosing a given action in Y but only
 330 on how many they are i.e., on the measure $\pi_{\#}^2 \gamma$. In this case $C(x, y, \gamma) =$
 331 $C(x, y, \pi_{\#}^2 \gamma)$.

332 For $p \geq 1$ take $\nu \in \mathcal{P}_p(Y)$ (arbitrary but fixed) and define

$$\mathcal{X}_p = \left\{ \gamma \in \tilde{\mathcal{X}} \left| \int_X \mathbb{W}_p(\gamma_x, \nu)^p \mu(dx) < \infty \right. \right\}. \quad (47)$$

333 Note that the definition of \mathcal{X}_p does not depend on ν . When we do not mention
 334 the p index we mean $p = 2$, i.e, $\mathcal{X} = \mathcal{X}_2$. We introduce on \mathcal{X}_p the distance:

$$\forall \gamma_0, \gamma_1 \in \mathcal{X}_p : d_{\mathcal{X}, p}(\gamma_0, \gamma_1) = \left(\int_X \mathbb{W}_p(\gamma_{0,x}, \gamma_{1,x})^p \mu(dx) \right)^{1/p}. \quad (48)$$

335 Note that from (70) using the Holder inequality we obtain a similar rela-
 336 tion for $d_{\mathcal{X},p}(\gamma_0, \gamma_1)$:

$$\forall 1 \leq p_1 \leq p_2, \gamma_0, \gamma_1 \in \mathcal{X}_{p_1} : d_{\mathcal{X},p_1}(\gamma_0, \gamma_1) \leq d_{\mathcal{X},p_2}(\gamma_0, \gamma_1). \quad (49)$$

337 With these provisions we can state the following result (compare with [7,
 338 Corrolary 5.2 page 425]):

339 **Proposition 1** *Suppose Y is the closure of some open set in \mathbb{R}^d . Let V_0 be a*
 340 *smooth function satisfying the hypothesis of Lemma 2 (in the appendix) and*
 341 *$\phi : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ a function with continuous second derivatives everywhere.*
 342 *Then, if $\epsilon > 0$ is small enough, the game with payoff*

$$C(x, y, \gamma) = \frac{|x - y|^2}{2} + V_0(y) + \epsilon \int_{X \times Y} \phi(y, y_2) \gamma(dx, dy_2), \quad (50)$$

343 *admits a unique Nash-MFG equilibrium and any EVI-equilibrium flow (as in*
 344 *theorem 2) converges exponentially fast to it.*

345 **Proof** We derive first a duality formula for the distance $d_{\mathcal{X},1}$; for any $g :$
 346 $X \times Y \rightarrow \mathbb{R}$, $\gamma_0, \gamma_1 \in \mathcal{X}_1$, we obtain from (71):

$$\begin{aligned} \int_{X \times Y} g(x, y) (\gamma_0(dx, dy) - \gamma_1(dx, dy)) &= \int_X \left(\int_Y g(x, y) [\gamma_{0,x}(dy) - \gamma_{1,x}(dy)] \right) \mu(dx) \\ &\leq \int_X \|g(x, \cdot)\|_{Lip} \cdot \mathbb{W}_1(\gamma_{0,x}, \gamma_{1,x}) \mu(dx) \leq \left[\sup_{x \in X} \|g(x, \cdot)\|_{Lip} \right] \cdot d_{\mathcal{X},1}(\gamma_0, \gamma_1). \end{aligned} \quad (51)$$

347 We use theorem 2 on $(\mathcal{X}, d_{\mathcal{X},2})$ and check all assumptions.

348 • First, since V_0 and ϕ are smooth and X, Y are compact one obtains as-
 349 sumption **(A₁)**.

350 • The assumption **(A₂)** is also verified because the function $C(x, y, \gamma)$ is
 351 $1 + \lambda_V - \epsilon C_2$ -convex (C_2 being a constant that only depends on ϕ and Y). In
 352 this case the geodesics between γ_0 and γ_1 are the curves γ_t such that $t \mapsto \gamma_{t,x}$
 353 is the generalized geodesic (with some fixed base ν , $\nu \in \mathcal{P}(Y)$, ν being
 354 absolutely continuous with respect to the Lebesgue measure dx) having end
 355 points $\gamma_{0,x}$ and $\gamma_{1,x}$. Here $(\gamma_{t,x})_{x \in X}$ is the disintegration of γ_t for any $t \in [0, 1]$.

356 • Assumption **(A₄)** is obvious since \mathcal{X} is compact.

357 • Remain only assumptions **(A₆)** and **(A₈)**. We only prove **(A₈)**, the other
 358 being easier to obtain (and uses essentially the same arguments). The idea
 359 is to prove the upper bound for the distance $d_{\mathcal{X},1}$ invoking (51) then use the

360 ordering in (49). Let $\gamma_0, \gamma_1, \gamma_2 \in \mathcal{X}$. Then

$$\begin{aligned}
& |\mathcal{C}(\gamma_0, \gamma_1) + \mathcal{C}(\gamma_1, \gamma_2) - \mathcal{C}(\gamma_0, \gamma_2) - \mathcal{C}(\gamma_1, \gamma_1)| = \\
& \epsilon \left| \int_{X \times Y} \int_{X \times Y} \phi(y_1, y_2) [\gamma_1(dx, dy_2) - \gamma_2(dx, dy_2)] [\gamma_0(dx, dy_1) - \gamma_1(dx, dy_1)] \right| \\
& \leq \epsilon \left\| \int_{X \times Y} \phi(\cdot, y_2) [\gamma_1(dx, dy_2) - \gamma_2(dx, dy_2)] \right\|_{Lip} d_{\mathcal{X},1}(\gamma_0, \gamma_1), \tag{52}
\end{aligned}$$

361 where (51) has been used. It remains to see that the Lipschitz norm of
362 $\int_{X \times Y} \phi(\cdot, y_2) [\gamma_1(dx, dy_2) - \gamma_2(dx, dy_2)]$ can be upper bounded by $d_{\mathcal{X},1}(\gamma_1, \gamma_2)$.
363 But for any $y, \tilde{y} \in Y$ with $y \neq \tilde{y}$:

$$\begin{aligned}
& \left| \int_{X \times Y} \phi(y, y_2) [\gamma_1(dx, dy_2) - \gamma_2(dx, dy_2)] - \int_{X \times Y} \phi(\tilde{y}, y_2) [\gamma_1(dx, dy_2) - \gamma_2(dx, dy_2)] \right| \\
& = \left| \int_{X \times Y} [\phi(y, y_2) - \phi(\tilde{y}, y_2)] [\gamma_1(dx, dy_2) - \gamma_2(dx, dy_2)] \right| \\
& \leq |y - \tilde{y}| \cdot d_{\mathcal{X},1}(\gamma_1, \gamma_2) \left[\sup_{y, \tilde{y} \in Y} \left\| \frac{\phi(y, \cdot) - \phi(\tilde{y}, \cdot)}{|y - \tilde{y}|} \right\|_{Lip} \right]. \tag{53}
\end{aligned}$$

364 Since ϕ is twice differentiable, assumption **(A₈)** follows with $C_L = \epsilon C_3$ where
365 C_3 is a constant depending only on ϕ and the space Y . ■

366 3.2.1 Congestion Mean Field Games

367 We consider next a congestion mean field game. Consider a population that
368 has to choose a place on a real line (an adaptation of the "beach example" in
369 P.L. Lions' lectures at Collège de France). The strategy of any individual is
370 a probability law ξ on \mathbb{R} . The individuals prefer to be at the origin and the
371 cost increases with the distance from the origin; this is modeled by a term
372 $\int_{\mathbb{R}} F_0(x) \xi(dx)$ with, for instance, $F_0(x) = c_0 x^2/2$, $c_0 > 0$. The individuals
373 dislike congestion which means that there is a penalty to choose a location
374 where the local density of others is large. If the mean field density of others
375 is $m(dx)$ the term that models congestion is $\int_{\mathbb{R}} \int_{\mathbb{R}} \rho_\sigma(y-x) m(dx) \xi(dy)$ where
376 ρ_σ is a given kernel, for instance $\rho_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$, for some $\sigma \geq 0$
377 (σ models the range of interaction). Thus the cost functional is:

$$\mathcal{C}(\xi, m) = \int_{\mathbb{R}} F_0(y) \xi(dy) + \int_{\mathbb{R}} \int_{\mathbb{R}} \rho_\sigma(y-x) m(dx) \xi(dy). \tag{54}$$

378 This can be put into the framework of the previous section taking X to
379 be a singleton (all individuals are alike) but here Y is potentially the whole

380 real line (thus in particular not compact). However, since the individual
 381 minimizes his / her cost, any Nash-MFG equilibrium strategy ξ , if it exists,
 382 must satisfy

$$\text{supp}(\xi) \subset \text{argmin}_{y \in \mathbb{R}} \left\{ F_0(y) + \int_{\mathbb{R}} \rho_{\sigma}(y-x) \xi(dx) \right\}. \quad (55)$$

383 In particular any y in the support of ξ will satisfy $c_0 y^2 / 2 \leq 1 / (\sigma \sqrt{2\pi})$, thus
 384 the support of ξ is included in $Y = [-M, M]$ with M a constant depending
 385 only on c_0 and σ . Thus we can apply the Proposition 1 to conclude that, as
 386 soon as c_0 is large enough, theorem 2 can be applied which gives the existence
 387 and uniqueness of the EVI equilibrium flow (1) for this choice of \mathcal{C} .

388 **Remark 8** When $\sigma \rightarrow 0$ the equilibrium is the semicircle law [3]: $\xi(x) =$
 389 $\frac{2}{\pi r_0^2} (r_0^2 - x^2)_+ dx$ for some constant $r_0 > 0$. Note however that the approach
 390 above does not guarantee convexity for any choice of c_0 and in particular when
 391 $\sigma \rightarrow 0$ no value of c_0 is large enough. On the contrary, a metric similar to
 392 that in section 3.4 is 0-convex for any c_0 and allows to use theorem 2.

393 3.3 Non-smooth interaction Mean Field Games with 394 multiple equilibria

395 Consider now the cost

$$\mathcal{C}(\xi, \eta) = \int_{\mathbb{R}} F_0(x) \xi(dx) + \int_{\mathbb{R}} \int_{\mathbb{R}} |x-y| \xi(dx) \eta(dy). \quad (56)$$

396 This is a MFG model similar to that in sections 3.1 and 3.2.1 where individ-
 397 uals prefer to be at the origin but here they also prefer to be close together.

398 With the same technique as in (55) we can prove that it is enough to
 399 restrict to the domain $[-R, R]$ with $R = \sqrt{\frac{2M_1(\xi)}{c_0}}$, where $M_1(\xi) = \int_{\mathbb{R}} |x| \xi(dx)$.

400 But then

$$M_1(\xi) = \int_{\mathbb{R}} |x| \xi(dx) = \int_{-R}^R |x| \xi(dx) \leq R = \sqrt{\frac{2M_1(\xi)}{c_0}}, \quad (57)$$

401 thus $R \leq 2/c_0$. Consider $\Omega = [-2/c_0, 2/c_0]$, $(\mathcal{X}, d) = (\mathcal{P}_2(\Omega), \mathbb{W}_2)$ and the
 402 cost functional:

$$\mathcal{C}(\xi, \eta) = \int_{\Omega} F_0(x) \xi(dx) + \int_{\Omega} \int_{\Omega} |x-y| \xi(dx) \eta(dy). \quad (58)$$

403 We start checking the assumptions of the theorem 2:

- 404 • (\mathbf{A}_1) is verified with $C_1 = 0$;
 405 • (\mathbf{A}_2) is verified with $\lambda = c_0$ taking as interpolating curves the generalized
 406 geodesics with some arbitrary, fixed, base; note that there is no restriction
 407 on c_0 except strict positivity, in particular it does need to be large;
 408 • (\mathbf{A}_6) is verified for $L = 1$, using arguments similar to those in the proof of
 409 Proposition 1 (i.e., bounding first by the \mathbb{W}_1 distance);
 410 • (\mathbf{A}_4) is also satisfied.

411 This allows to use the first point of the theorem 2 to ensure the conver-
 412 gence of numerical schemes to a EVI equilibrium flow. Note that (\mathbf{A}_8) is not
 413 valid here. Based on section 3.1 we do not expect to have uniqueness of the
 414 Nash-MFG equilibria. Indeed one can check that any δ_{x^*} with $|x^*| \leq 1/c_0$ is
 415 a Nash-MFG equilibrium:

$$\begin{aligned} \forall \xi \in \mathcal{X} : \mathcal{C}(\xi, \delta_{x^*}) &= \int_{\Omega} (F_0(x) + |x - x^*|) \xi(dx) \geq \int_{\Omega} F_0(x^*) \xi(dx) \\ &= F_0(x^*) = \int_{\Omega} (F_0(x) + |x - x^*|) \delta_{x^*}(dx) = \mathcal{C}(\delta_{x^*}, \delta_{x^*}). \end{aligned} \quad (59)$$

416 3.4 Vaccination Mean Field Games

417 An important class of MFG examples that go beyond potential or interaction
 418 terms are vaccination mean field games (see [13, 4, 5, 23] and more generally
 419 [10, 15, 34] for vaccination coverage dynamics). In this framework the indi-
 420 vidual can choose his / her vaccination strategy which is a probability law
 421 set on $\Omega = [0, T] \cup \{\infty\}$ (the point at infinity meaning no vaccination). Thus
 422 \mathcal{X} is the space of probability measures on Ω ; we can also see it as a subset of
 423 $\dot{H}^{-1}(\Omega)$ (the dual of the space of zero-mean $H^1(\Omega)$ -Sobolev functions) with
 424 metric:

$$\begin{aligned} d^2(m_1, m_2) &= \sum_{\omega \in \{0, T, \infty\}} [m_1(\{\omega\}) - m_2(\{\omega\})]^2 \\ &+ \sup_{\|\nabla \phi\|_{L^2([0, T]; \mathbb{R})} \leq 1, \int_0^T \phi = 0} \left\{ \int_0^T \phi(m_1 - m_2) \right\}^2. \end{aligned} \quad (60)$$

425 This metric can also be given another expression (see [31, Section 5.5.2
 426 page 210]: denote by u_{m_1, m_2} the solution on $]0, T[$ of the Laplace problem
 427 $-\Delta u_{m_1, m_2} = m_1 - m_2$ with homogeneous Neumann boundary conditions

428 $\frac{\partial u}{\partial n} = 0$ in 0 and T ; note that u_{m_1, m_2} depends linearly of $m_1 - m_2$. Then:

$$d^2(m_1, m_2) = \left(\sum_{\omega \in \{0, T, \infty\}} [m_1(\{\omega\}) - m_2(\{\omega\})]^2 \right) + \|\nabla u_{m_1, m_2}\|_{L^2([0, T]; \mathbb{R})}^2. \quad (61)$$

429 The cost can be expressed, for some functional $F : (\mathcal{X}, d) \rightarrow C(\Omega; \mathbb{R})$, as

$$\mathcal{C}(\xi, m) = \int_{\Omega} F(m)(x) \xi(dx). \quad (62)$$

430 The functional F is highly non-linear and depends on the solution of a system
 431 of measure-driven ODEs, see [23, equations (1) and (4)] for details. More
 432 precisely, for some constants $\beta, \gamma > 0$, $0 < r_V < r_I < \infty$, $S_0, I_0 \geq 0$,
 433 $S_0 + I_0 < 1$, consider the system (see also [9, Theorem 10.2.3 page 246] for
 434 the well-posedness):

$$\begin{cases} dS^m(t) = -\beta S^m(t)I^m(t)dt - dm(t), & S^m(0^-) = S_0, \\ dI^m(t) = (\beta S^m(t) - \gamma)I^m(t)dt, & I^m(0^-) = I_0, \\ d\varphi^m(t) = \beta I^m(t)(1 - \varphi^m(t))dt, & \varphi^m(0^-) = 0, \end{cases} \quad (63)$$

435 where m is prolonged by 0 on $]T, \infty[$. Note that $S^m(t)$, $I^m(t)$ and $\varphi^m(t)$ take
 436 values in $[0, 1]$. The cost is:

$$\mathcal{C}(\xi, m) = -\xi(\{\infty\})r_I\varphi^m(\infty) + \int_0^T \left(r_V + (r_I - r_V)\varphi^m(t) \right) \xi(dt), \quad (64)$$

437 and thus

$$F(m)(t) = r_V + (r_I - r_V)\varphi^m(t) \text{ for } t < \infty, \quad F(m)(\infty) = -r_I\varphi^m(\infty). \quad (65)$$

- 438 • In particular F is bounded from below which implies hypothesis **(A₁)**.
- 439 • The motivation for the distance (61) appears when trying to fulfill the
 440 assumption **(A₂)**. Indeed, there is no apparent convexity in the function
 441 $x \mapsto F(m)(x)$. But, the cost \mathcal{C} is linear in the first variable with respect
 442 to affine combinations of measures. Therefore we choose as curves in the
 443 assumption **(A₂)** the segments $\gamma(t) = (1 - t)\gamma(0) + t\gamma(1)$. With this choice
 444 the distance (61) satisfies the equation (20) (in fact we have equality) while
 445 \mathcal{C} will satisfy (19) with $\lambda = 0$ thus assumption **(A₂)** is valid.
- 446 • Assumption **(A₄)** is immediate, the space being compact.
- 447 • Assumption **(A₆)** follows immediately after noting that $\frac{d}{dt}F(m)(t)$ is
 448 bounded independently on m and t because $\frac{d}{dt}\varphi^m(t)$ is.
- 449 • Assumption **(A₈)** requires longer computations and is left as an exercise
 450 for the reader.

451 **Remark 9** *Within the vaccination MFG framework one can encounter also*
 452 *different, more exotic, metric spaces, see for instance [18] where:*

$$\mathcal{X} = \{u : [0, \infty[\rightarrow \mathbb{R} \mid u(0) = 0, \forall b \geq a \geq 0 : |u(a) - u(b)| \leq |a - b|, u(b) \geq u(a)\}, \quad (66)$$

453 *the metric $d^2(f, g) = \sum_{n \in \mathbb{N}^*} \frac{\|f - g\|_{L^2([0, n])}^2}{2^n}$, and the cost functional*

$$\mathcal{C}(\xi, m) = \int_{\mathbb{R}} F(m)(t) \dot{u}(t) e^{-rt} dt, \quad (67)$$

454 *for some functional $F : \mathcal{X} \rightarrow C(\mathbb{R}; \mathbb{R})$.*

455 *In all cases, the numerical simulations show that the algorithm (2) con-*
 456 *verges to a MFG equilibrium.*

457 Acknowledgements

458 G.T. acknowledges support from the Agence Nationale de la Recherche (ANR),
 459 projects *EMAQS* (ANR-2011-BS01-017-01), *CINE-PARA* and *MFG* (*ANR-*
 460 *16-CE40-0015-01*).

461 A Notations

462 We recall below some notations and results used in the paper.

463 For any set Ω we denote by $\mathcal{P}(\Omega)$ the ensemble of probability laws on Ω .
 464 For a general application f defined on Ω with values in a measure space, $f_{\#}\nu$
 465 is the push-forward (image) measure of $\nu \in \mathcal{P}(\Omega)$ through f characterized
 466 by $f_{\#}\nu(A) = \nu(f^{-1}(A))$ for any measurable set A in the image. When Ω is
 467 a metric space, for $p \geq 1$ and some $x_0 \in \Omega$:

$$\mathcal{P}_p(\Omega) = \left\{ \nu \in \mathcal{P}(\Omega) \mid \int_{\Omega} d(x, x_0)^p \nu(dx) < \infty \right\}, \quad (68)$$

468 i.e., $\mathcal{P}_p(\Omega)$ is the set of probability laws on Ω with finite p -th moment. Note
 469 that the definition does not depend on the choice of x_0 .

470 When Ω is a tensor product $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_M$, and $\gamma \in \mathcal{P}(\Omega)$ we
 471 denote $\pi_{\#}^1 \gamma \in \mathcal{P}(\Omega_1)$ the first marginal, $\pi_{\#}^2 \gamma \in \mathcal{P}(\Omega_2)$ the second marginal,
 472 $\pi_{\#}^{1,2} \gamma \in \mathcal{P}(\Omega_1 \times \Omega_2)$ the marginal with respect to the first two coordinates,
 473 etc.

474 The set $\mathcal{P}_p(\Omega)$ can be given the structure of a metric space with the
 475 p -Wasserstein metric [31, Section 5] denoted \mathbb{W}_p with

$$\mathbb{W}_p(\nu_1, \nu_2) = \left(\inf \left\{ \int_{\Omega \times \Omega} d(x, y)^p \gamma(x, y) \mid \gamma \in \mathcal{P}(\Omega \times \Omega), \pi_{\#}^1 \gamma = \nu_1, \pi_{\#}^2 \gamma = \nu_2 \right\} \right)^{1/p}. \quad (69)$$

476 The set of γ that realize the "inf" in (69) is denoted $\Gamma_o(\nu_1, \nu_2)$.

477 Proving that \mathbb{W}_p is a distance can be performed with usual techniques [31,
 478 Lemmas 5.4 and 5.5 page 182]; also standard is to note that these distances
 479 are ordered i.e.

$$\mathbb{W}_{p_1}(\gamma_0, \gamma_1) \leq \mathbb{W}_{p_2}(\gamma_0, \gamma_1), \quad \forall 1 \leq p_1 \leq p_2, \quad \forall \gamma_0, \gamma_1 \in \mathcal{P}_{p_1}(\Omega). \quad (70)$$

480 Recall that the norm $\|f\|_{Lip}$ of a function f is its smallest Lipschitz con-
 481 stant (and $+\infty$ if no such constant exists). Then (see [33, Remark 6.5 page
 482 95]):

$$\mathbb{W}_1(\gamma_0, \gamma_1) = \sup \left\{ \int_{\Omega} f(\gamma_0 - \gamma_1) \mid \|f\|_{Lip} \leq 1 \right\}. \quad (71)$$

483 In order to exploit convexity, we need to define the notion of (generalized)
 484 geodesics (see [2, Definition 9.2.2. page 207]): let $\nu_0, \nu_1, \nu \in \mathcal{P}_p(\Omega)$ and $\xi \in$
 485 $\mathcal{P}(\Omega \times \Omega \times \Omega)$ such that $\pi_{\#}^{1,2} \xi \in \Gamma_0(\nu, \nu_0)$, $\pi_{\#}^{1,3} \xi \in \Gamma_0(\nu, \nu_1)$. The generalized
 486 geodesic between ν_0 and ν_1 with base ν is the curve $(\nu^\xi(t))_{t \in [0,1]} \subset \mathcal{P}_p(\Omega)$
 487 with

$$\nu^\xi(t) = [(y_1, y_2, y_3) \mapsto ((1-t)y_2 + ty_3)]_{\#} \xi. \quad (72)$$

488 In particular $\nu^\xi(0) = \nu_0$, $\nu^\xi(1) = \nu_1$. When $\nu = \nu_0$ the generalized geodesic
 489 is a (ordinary) geodesic in the space $\mathcal{P}_p(\Omega)$; in this case we do not mention
 490 the base any more.

491 This definition allows to state the following result (for the proof see [2,
 492 Lemma 9.2.1 page 206, Proposition 9.3.2 page 210]):

493 **Lemma 2** Take $p = 2$, $\lambda_V \geq 0$ and $V : \Omega \rightarrow \mathbb{R}$ be a λ_V -convex function in
 494 the sense that for any $y_1, y_2 \in \Omega$, $t \in [0, 1]$:

$$V((1-t)y_1 + ty_2) \leq (1-t)V(y_1) + tV(y_2) - \frac{\lambda_V}{2} d^2(y_1, y_2). \quad (73)$$

495 Then the functional $\mathcal{V} : \mathcal{P}_p(\Omega) \rightarrow \mathbb{R}$:

$$\mathcal{V}(\gamma) = \int_{\Omega} V(y) \gamma(dy), \quad (74)$$

496 is λ_V -convex on $(\mathcal{P}_p(\Omega), \mathbb{W}_p)$ in the sense of the assumption **(A₂)** on the
 497 generalized geodesics with base ν for any $\nu \in \mathcal{P}_p(\Omega)$.

References

- 499 [1] Luigi Ambrosio and Nicola Gigli. *Modelling and Optimisation of Flows*
500 *on Networks: Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel*
501 *Rasche*, chapter A User’s Guide to Optimal Transport, pages 1–155.
502 Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- 503 [2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in*
504 *metric spaces and in the space of probability measures. 2nd ed.* Basel:
505 Birkhäuser, 2nd ed. edition, 2008.
- 506 [3] Z. D. Bai and Y. Q. Yin. Convergence to the semicircle law. *Ann.*
507 *Probab.*, 16(2):863–875, 1988.
- 508 [4] Chris T. Bauch and David J. D. Earn. Vaccination and the theory of
509 games. *Proc. Natl. Acad. Sci. USA*, 101(36):13391–13394 (electronic),
510 2004.
- 511 [5] Chris T. Bauch, Alison P. Galvani, and David J. D. Earn. Group interest
512 versus self-interest in smallpox vaccination policy. *Proceedings of the*
513 *National Academy of Sciences*, 100(18):10564–10567, 2003.
- 514 [6] Alain Benoussan, Jens Frehse, and Phillip Yam. *Mean Field Games and*
515 *Mean Field Type Control Theory*. Springer-Verlag New York, 2013.
- 516 [7] A. Blanchet and G. Carlier. Remarks on existence and uniqueness of
517 cournot–nash equilibria in the non-potential case. *Mathematics and Fi-*
518 *nancial Economics*, 8(4):417–433, 2014.
- 519 [8] Adrien Blanchet and Guillaume Carlier. From Nash to Cournot-Nash
520 equilibria via the Monge-Kantorovich problem. *Philos. Trans. R. Soc.*
521 *Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2028):20130398, 11, 2014.
- 522 [9] Alberto Bressan and Benedetto Piccoli. *Introduction to the mathematical*
523 *theory of control*, volume 2 of *AIMS Series on Applied Mathematics*.
524 American Institute of Mathematical Sciences (AIMS), Springfield, MO,
525 2007.
- 526 [10] Bruno Buonomo, Alberto d’Onofrio, and Deborah Lacitignola. Global
527 stability of an SIR epidemic model with information dependent vacci-
528 nation. *Mathematical Biosciences*, 216(1):9 – 16, 2008.
- 529 [11] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric*
530 *geometry*, volume 33 of *Graduate Studies in Mathematics*. American
531 Mathematical Society, Providence, RI, 2001.

- 532 [12] Lucas C. F. Ferreira and Julio C. Valencia-Guevara. Gradient flows of
533 time-dependent functionals in metric spaces and applications to pdes.
534 *Monatshefte für Mathematik*, pages 1–38, 2017.
- 535 [13] Paul E. M Fine and Jacqueline A Clarkson. Individual versus public
536 priorities in the determination of optimal vaccination policies. *American*
537 *Journal of Epidemiology*, 124(6):1012–1020, 1986.
- 538 [14] Drew Fudenberg and David K. Levine. *The theory of learning in games*,
539 volume 2 of *MIT Press Series on Economic Learning and Social Evolu-*
540 *tion*. MIT Press, Cambridge, MA, 1998.
- 541 [15] Sebastian Funk, Marcel Salathé, and Vincent A. A. Jansen. Modelling
542 the influence of human behaviour on the spread of infectious diseases: a
543 review. *J. R. Soc. Interface*, 7:1247–1256, 2010.
- 544 [16] Minyi Huang, Roland P. Malhamé, and Peter E. Caines. Nash equi-
545 libria for large-population linear stochastic systems of weakly coupled
546 agents. In Elkébir Boukas and Roland P. Malhamé, editors, *Analysis,*
547 *Control and Optimization of Complex Dynamic Systems*, pages 215–252.
548 Springer US, 2005.
- 549 [17] Minyi Huang, Roland P. Malhamé, and Peter E. Caines. Large popula-
550 tion stochastic dynamic games: closed-loop mckean-vlasov systems and
551 the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–
552 252, 2006.
- 553 [18] Emma Hubert and Gabriel Turinici. Nash-MFG equilibrium in a SIR
554 model with time dependent newborn vaccination. [https://hal.archives-](https://hal.archives-ouvertes.fr/hal-01389584)
555 [ouvertes.fr/hal-01389584](https://hal.archives-ouvertes.fr/hal-01389584), October 2016.
- 556 [19] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational for-
557 mulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–
558 17, 1998.
- 559 [20] E. Kopfer. Gradient flow for the Boltzmann entropy and Cheeger’s en-
560 ergy on time-dependent metric measure spaces. *ArXiv e-prints*, Novem-
561 ber 2016.
- 562 [21] E. Kopfer and K.-T. Sturm. Heat Flows on Time-dependent Metric
563 Measure Spaces and Super-Ricci Flows. *ArXiv e-prints*, November 2016.
- 564 [22] S. N. Kružkov. First order quasilinear equations with several indepen-
565 dent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.

- 566 [23] Laetitia Laguzet and Gabriel Turinici. Individual vaccination as Nash
567 equilibrium in a SIR model with application to the 2009–2010 in-
568 fluenza A (H1N1) epidemic in France. *Bulletin of Mathematical Biology*,
569 77(10):1955–1984, 2015.
- 570 [24] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I: Le
571 cas stationnaire. *C. R., Math., Acad. Sci. Paris*, 343(9):619–625, 2006.
- 572 [25] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. II:
573 Horizon fini et contrôle optimal. *C. R., Math., Acad. Sci. Paris*,
574 343(10):679–684, 2006.
- 575 [26] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese*
576 *Journal of Mathematics*, 2(1):229–260, 2007.
- 577 [27] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Variational
578 convergence of gradient flows and rate-independent evolutions in metric
579 spaces. *Milan J. Math.*, 80(2):381–410, 2012.
- 580 [28] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré. Nonsmooth
581 analysis of doubly nonlinear evolution equations. *Calc. Var. Partial*
582 *Differential Equations*, 46(1-2):253–310, 2013.
- 583 [29] John Nash. Non-cooperative games. *Ann. of Math. (2)*, 54:286–295,
584 1951.
- 585 [30] Riccarda Rossi, Alexander Mielke, and Giuseppe Savaré. A metric ap-
586 proach to a class of doubly nonlinear evolution equations and applica-
587 tions. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 7(1):97–169, 2008.
- 588 [31] Filippo Santambrogio. *Optimal transport for applied mathematicians*.
589 Progress in Nonlinear Differential Equations and their Applications, 87.
590 Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and
591 modeling.
- 592 [32] Cédric Villani. *Optimal transport. Old and new*, volume 338 of
593 *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.
- 594 [33] Cédric Villani. *Optimal transport, Old and new*, volume 338. Springer-
595 Verlag, Berlin, 2009.
- 596 [34] Zhen Wang, Chris T. Bauch, Samit Bhattacharyya, Alberto d’Onofrio,
597 Piero Manfredi, Matjaž Perc, Nicola Perra, Marcel Salathé, and Dawei
598 Zhao. Statistical physics of vaccination. *Physics Reports*, 664:1 – 113,
599 2016.