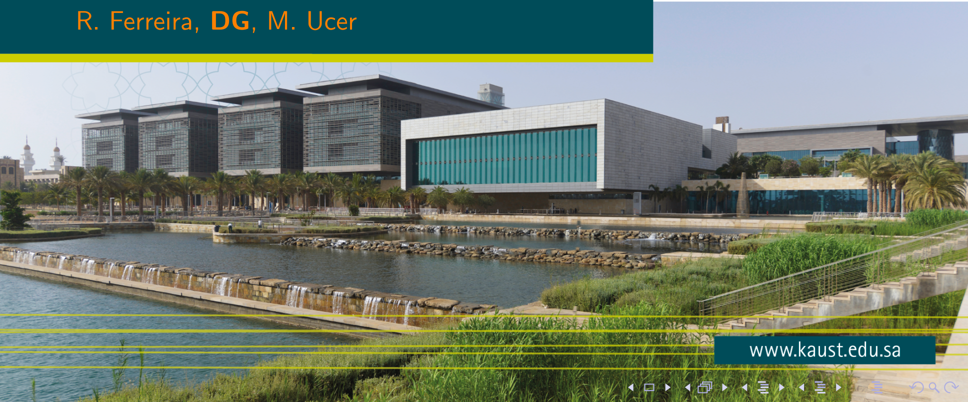
جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Monotonicity Methods for Mean Field Games

A Functional Analytic Perspective

R. Ferreira, **DG**, M. Ucer



www.kaust.edu.sa

Motivation: Digital Twins for Large Populations

- Digital twins (mobility, energy, crowds) involve massive numbers of interacting agents.
- **The Problem:** Agent-based models scale poorly.
- **The Solution (MFG):** A continuum limit ($N \to \infty$) modeling the population density.
- **Key Advantage:** Complexity becomes independent of the population size N.

MFG in One Slide (Stationary, Periodic)

- **Context:** Digital twins for large populations (traffic, energy, crowds, markets).
- **Model:** Representative agent on the torus $\mathbb{T}^d$.

$$\begin{cases} \text{(HJ):} & -u - H\big(x, Du, m\big) + V(x) = 0, \\ \text{(Transport):} & m - \operatorname{div}\big(m\, D_p H(x, Du, m)\big) - 1 = 0. \end{cases}$$

Interpretation:

- **HJ:** optimality condition for one agent (cost depends on crowd m).
- **Transport:** stationary distribution when everyone follows the optimal strategy Du.
- **Together:** a PDE-based digital twin where microscopic decisions and macroscopic density are consistent.

Separable Case & Crowd Aversion

- **Separable structure:** $H(x, p, m) = H(x, p) - f(m)$, with H convex in p.

$$\boxed{\begin{cases} -u - H(x, Du) + f(m) = 0, \\ m - \operatorname{div}\big(m\, D_p H(x, Du)\big) - 1 = 0. \end{cases}}$$

- **Monotonicity / crowd aversion:** if f is **increasing**, higher m raises the cost so agents avoid crowded regions (congestion, bottlenecks).

Talk roadmap

We use monotonicity methods to give a unified analysis of MFG-based digital twins:

1. **Structure:** identify variational (loss), saddle-point, and monotone-operator formulations, which support stable learning algorithms (primal–dual / GAN-like).
2. **Existence:** prove existence of **strong** solutions (not just weak ones) via a low-order regularization of the MFG system.
3. **Uniqueness:** establish **weak–strong uniqueness**, ensuring that numerical / ML solvers converge to the physically meaningful solution.

Variational Principle

Separable systems admit a variational formulation via a suitable convex functional F.

- **Variational Principle:** The MFG system is the Euler-Lagrange equation for

$$\mathcal{J}(u) = \int_{\mathbb{T}^d} \left\{ F\big(u + H(x, Du)\big) - u \right\} dx.$$

Saddle Formulation

- Using convex duality, rewrite the minimization as

$$\min_u \mathcal{J}(u) = \min_u \max_{m \geq 0} \int_{\mathbb{T}^d} \Big[-F^*(m) - m\big(u + H(x, Du)\big) - u \Big] dx,$$

  where F^* is the convex conjugate of F.
- This is a **convex–concave** problem; its skew gradient is **monotone**.

MFGs and Machine Learning

- **Variational view as a loss:** use $\mathcal{J}(u_\theta)$ as a loss for a neural network u_θ.
- **Saddle formulation:** The min–max problem in (u, m) is amenable to primal-dual / adversarial training algorithms (GAN-like).

Running Example: Quadratic H and Power Coupling

The Quadratic Model

$$\begin{cases} -u - \frac{1}{2}|Du|^2 + f(m) = 0, \\ m - \operatorname{div}(m\,Du) - 1 = 0, \end{cases}$$

Increasing coupling: $f(m) \sim m^\beta$ with $\beta > 0$.

- **Interpretation:** agents trade kinetic effort against density penalty $f(m)$ – a standard crowd-averse model.

The Monotone Operator and Functional Setting

Write the MFG system as $A(m,u)=0$.

- **The Operator** (for $H=\frac{1}{2}|p|^2$):

$$A\begin{bmatrix} m \\ u \end{bmatrix} = \begin{bmatrix} -u-\frac{1}{2}|Du|^2+f(m) \\ m-\operatorname{div}(mDu)-1 \end{bmatrix}.$$

- **Functional setting:** $X = L^{\bar\beta}\times W^{1,\bar\gamma}$,

$$A : X \to X' = L^{\bar\beta'} \times W^{-1,\bar\gamma'}.$$

  $\bar\beta=\beta+1$, $\bar\gamma = 2\,\bar\beta'$.

- **Monotonicity:** A is the skew gradient of a convex-concave functional, hence monotone.

Strong vs. Minty Weak Solutions

Let $X = L^{\bar{\beta}} \times W^{1,\bar{\gamma}}$ and $\mathcal{K} \subset X$ convex (e.g. $m \geq 0$).

Strong solution (The "Physical" Solution)

$w = (m, u) \in \mathcal{K}$ such that

$$\langle A(w), z - w \rangle \geq 0, \qquad \forall z \in \mathcal{K}.$$

If $w \in \operatorname{int} \mathcal{K}$, then $A(w) = 0$.

Weak solution (Minty)

$w \in \overline{\mathcal{K}}$ such that

$$\langle A(z), z - w \rangle \geq 0 \quad \forall z \in \mathcal{K}.$$

Challenge: Numerical methods often converge to weak solutions. We need to ensure these are physical.

Abstract Existence (Browder–Minty / Debrunner–Flor)

Theorem

If $A : \mathcal{K} \to X'$ is **monotone**, **hemicontinuous**, and **coercive** on a convex $\mathcal{K} \subset X$, then there exists a Minty weak solution $w^* \in \overline{\mathcal{K}}$.

- Under mild conditions, w^* can be upgraded to a strong solution.
- **Issue:** A is monotone and hemicontinuous but **not** coercive on X, so we need a workaround (next slide).

Low-Order Regularization (on u only)

To enforce coercivity, we regularize using a $\bar{\gamma}$-Laplacian:

$$A_\varepsilon[m,u] = A[m,u] + \varepsilon \begin{bmatrix} 0 \\ -\operatorname{div}\left(|Du|^{\bar{\gamma}-2} Du\right) - |u|^{\bar{\gamma}-2} u \end{bmatrix}.$$

- **Adds coercivity in u:** the pairing generates $\varepsilon \|u\|_{W^{1,\bar{\gamma}}}^{\bar{\gamma}}$ (analogous to regularizers in ML).
- Coercivity in m comes from the growth of $f(m)$.
- **Variational meaning:**

$$\mathcal{J}(u) \;\mapsto\; \mathcal{J}(u) + \frac{\varepsilon}{\bar{\gamma}} \int_{\mathbb{T}^d} \left(|Du|^{\bar{\gamma}} + |u|^{\bar{\gamma}}\right) dx.$$

A Priori Bounds: The Energy Trick

- **Energy Identity:** Testing the system $A_\varepsilon(w_\varepsilon) = 0$ against the solution w_ε yields uniform bounds on the density and the regularized energy.

$$\|m_\varepsilon\|_{L^{\bar{\beta}}} + \varepsilon\|u_\varepsilon\|_{W^{1,\bar{\gamma}}} \leq C.$$

- **The Challenge:** The bound on u degenerates as $\varepsilon \to 0$.
- **Fix:** Use the Hamilton–Jacobi equation: from $|Du|^2 \lesssim f(m) + |u|$ and the density bound, we obtain a uniform (in ε) bound on Du.

From Regularization to Strong Solutions

Strategy to recover the physical solution:

1. **Approximate:** Solve the regularized system $A_\varepsilon(w_\varepsilon) = 0$.
2. **Compactness:** Uniform estimates allow us to extract a weak limit $(m_\varepsilon, u_\varepsilon) \rightharpoonup (m, u)$.
3. **Identification:** Using the monotonicity of A, we identify the limit as a weak solution.
4. **Regularity:** Hemicontinuity upgrades the weak limit to a **strong solution** $A(m, u) = 0$.

Other Directions

Non-separable $H(x, Du, m)$: variational structure is lost, but monotone-operator ideas can still apply:

- **General power growth** $H \sim |p|^\alpha - m^\beta$: same blueprint works
- **Congestion** $H \sim |p|^\alpha / m^\tau$: solve on a wedge $(m > 0)$, then upgrade to strong solutions; important for density-dependent mobility (crowds, traffic).
- **Weak growth** (no upper bound in p): requires inf-convolution regularization H^ε; typically yields only Minty weak solutions.

Weak–Strong Uniqueness (Idea)

- **Question:** Do the numerical (weak) solutions coincide with the physical (strong) PDE solutions?
- **Setup:** Let w be a strong solution and $\tilde{w}$ be a Minty weak solution.
- **Mechanism / result:** If the linearized operator is strictly positive, then $w = \tilde{w}$.
- **Impact:** We do not need to worry about "spurious" weak solutions.

Conclusions & Non-Monotone Couplings

- **Summary:** Monotonicity methods plus low-order regularization yield strong solutions and weak–strong uniqueness for MFG-based digital twins

- **The Frontier: Non-Monotone Couplings** (e.g., crowd attraction).
  - Monotonicity and the saddle structure may fail.
  - Consequences: possible multiplicity of equilibria, loss of stability, and highly non-convex optimization landscapes for ML and numerical methods.
  - Critical for applications with attraction/clustering (agglomeration, social influence, market formation); largely open and calling for new analytical and ML tools.

Thank you.