Multi-agent System Learning and Control

Massimo Fornasier

Fakultät für Mathematik
Technische Universität München
massimo.fornasier@ma.tum.de
http://www-m15.ma.tum.de/

European Research Council

"Collective dynamics, control and imaging"
Institute for Theoretical Studies (ETHZ) – April 28, 2017

Joint work with M. Bongini, M. Caponigro, M. Hansen, D. Kalise, M. Maggioni, B. Piccoli, F. Rossi, F. Solombrino, E. Trelat

What is a self-organizing system?

Social dynamics

We consider large particle systems of form:

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (S + K * \mu_N)(x_i, v_i), \\ i = 1, \dots N, \quad t \in [0, T], \\ \text{where, } \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{(x_j, v_j)}, \end{cases}$$

Social dynamics

We consider large particle systems of form:

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (S + K * \mu_N)(x_i, v_i), \\ i = 1, \dots N, \quad t \in [0, T], \\ \text{where, } \mu_N = \frac{1}{N} \sum_{j=1}^{N} \delta_{(x_j, v_j)}, \end{cases}$$

"Social forces" encoded in S and K :

- Repulsion-attraction
- Alignment
- Self-propulsion/friction
- ....

Possible noise/uncertainty by adding stochastic terms.

Mills in nature and in our simulations. J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, *Particle, kinetic, hydrodynamic models of swarming*, Birkhäuser 2010.

Social dynamics

We consider large particle systems of form:

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (S + K * \mu_N)(x_i, v_i), \\ i = 1, \dots N, \quad t \in [0, T], \\ \text{where, } \mu_N = \frac{1}{N} \sum_{j=1}^{N} \delta_{(x_j, v_j)}, \end{cases}$$

Mills in nature and in our simulations. J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, *Particle, kinetic, hydrodynamic models of swarming*, Birkhäuser 2010.

"Social forces" encoded in S and K :

- Repulsion-attraction
- Alignment
- Self-propulsion/friction
- ....

Possible noise/uncertainty by adding stochastic terms.

Understanding how superposition of re-iterated binary "social forces" yields global self-organization.

Models for social dynamics - Cucker-Smale

For $S \equiv 0$ and $K(x, v) = a(\|x\|)(-v)$ we get the Cucker-Smale model

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \frac{1}{N}\sum_{j=1}^{N} a(\|x_i - x_j\|)(v_j - v_i), \quad i = 1, \dots N, \end{cases}$$

Usually, $a(r) = \frac{1}{(1+r^2)^{\beta}}$, where $\beta \in [0, +\infty]$.

Models for social dynamics - Cucker-Smale

For $S \equiv 0$ and $K(x, v) = a(\|x\|)(-v)$ we get the Cucker-Smale model

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = \frac{1}{N}\sum_{j=1}^{N} a(\|x_i - x_j\|)(v_j - v_i), \quad i = 1, \dots N, \end{cases}$$

Usually, $a(r) = \frac{1}{(1+r^2)^{\beta}}$, where $\beta \in [0, +\infty]$.

If $\beta < \frac{1}{2}$, alignment occurs everytime

If $\beta \geq \frac{1}{2}$, alignment occurs only for certain initial data

Models for social dynamics - Cucker-Dong

For $S(x,v) = -bv$ and $K(x,v) = a(\|x\|^2)(-x) - f(\|x\|^2)(-x)$ we get the Cucker-Dong model

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = -bv_i + \frac{1}{N}\sum_{j=1}^{N}(a-f)(\|x_i - x_j\|^2)(x_j - x_i), \quad i = 1, \dots N, \end{cases}$$

Models for social dynamics - Cucker-Dong

For $S(x,v) = -bv$ and $K(x,v) = a(\|x\|^2)(-x) - f(\|x\|^2)(-x)$ we get the Cucker-Dong model

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = -bv_i + \frac{1}{N}\sum_{j=1}^N (a-f)(\|x_i - x_j\|^2)(x_j - x_i), \quad i = 1, \dots N, \end{cases}$$

Usually, $a(r) = \frac{1}{(1+r^2)^\beta}$ and $f(r) = \frac{1}{r^\delta}$. A sufficient condition for self-organization is given by the quantities

$$E(t) = \sum_{i=1}^N \|v_i(t)\|^2 + \frac{1}{2N}\sum_{i<j}^N \left(\int_0^{\|x_i(t)-x_j(t)\|^2} a(r)dr + \int_{\|x_i(t)-x_j(t)\|^2}^{+\infty} f(r)dr \right)$$

$$\theta = \frac{N-1}{2}\int_0^{+\infty} a(r)dr$$

Models for social dynamics - Cucker-Dong

If $E(0) \leq \theta$, cohesiveness occurs everytime

If $E(0) > \theta$, cohesiveness occurs only for certain initial data

Models for social dynamics - D'Orsogna-Bertozzi et al.

For $S(x,v) = (\alpha - \beta\|v\|^2)v$ and $K(x,v) = -\nabla U(\|x\|)\frac{x}{\|x\|}$ we get the D'Orsogna-Bertozzi et al. model

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (\alpha - \beta\|v_i\|^2)v_i - \frac{1}{N}\sum_{j=1}^{N} \nabla U(\|x_i - x_j\|)\frac{x_i - x_j}{\|x_i - x_j\|}, \quad i = 1, \dots N, \end{cases}$$

Models for social dynamics - D'Orsogna-Bertozzi et al.

For $S(x, v) = (\alpha - \beta \|v\|^2)v$ and $K(x, v) = -\nabla U(\|x\|)\frac{x}{\|x\|}$ we get the D'Orsogna-Bertozzi et al. model

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (\alpha - \beta \|v_i\|^2)v_i - \frac{1}{N}\sum_{j=1}^{N} \nabla U(\|x_i - x_j\|)\frac{x_i - x_j}{\|x_i - x_j\|}, \quad i = 1, \ldots N, \end{cases}$$

Usually, $U(r) = -C_A e^{-r/\ell_A} + C_R e^{-r/\ell_R}$.

If $\frac{C_R}{C_A}\left(\frac{\ell_R}{\ell_A}\right)^d < 1$*, cristalline structures appear*

If $\frac{C_R}{C_A}\left(\frac{\ell_R}{\ell_A}\right)^d \geq 1$*, mill patterns arise*

Split coherence in homophilious societies: government?

- A society is said to be *homophilious* whenever its agents are more influenced by near agents than far ones;

Split coherence in homophilious societies: government?

- A society is said to be *homophilious* whenever its agents are more influenced by near agents than far ones;
- As seen in the CS and CD models, in homophilious societies global self-organization can be expected as soon as enough initial coherence is reached;

Split coherence in homophilious societies: government?

- A society is said to be *homophilious* whenever its agents are more influenced by near agents than far ones;
- As seen in the CS and CD models, in homophilious societies global self-organization can be expected as soon as enough initial coherence is reached;
- However, it is common experience that coherence in a homophilious society can be lost, leading sometimes to dramatic consequences, questioning strongly the role and the effectiveness of governments.

Split coherence in homophilious societies: government?

- A society is said to be *homophilious* whenever its agents are more influenced by near agents than far ones;
- As seen in the CS and CD models, in homophilious societies global self-organization can be expected as soon as enough initial coherence is reached;
- However, it is common experience that coherence in a homophilious society can be lost, leading sometimes to dramatic consequences, questioning strongly the role and the effectiveness of governments.

Question: can a government endowed with limited resources rescue/stabilize a society by minimal interventions? Which ones?

A parametric model of homo-to-hetero-philia

The Cucker-Smale model:

$$\begin{cases} \dot{x}_i = v_i \in \mathbb{R}^d \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} a(\|x_j - x_i\|)(v_j - v_i) \in \mathbb{R}^d \end{cases},$$

where $a(t) := a_\beta(t) = \frac{1}{(1+t^2)^\beta}$, $\beta > 0$ governs the rate of communication.

[1] The *Laplacian* L of A is given by $L = D - A$, with $D = \mathrm{diag}(d_1, \dots, d_N)$ and $d_k = \sum_{j=1}^{N} a_{kj}$

A parametric model of homo-to-hetero-philia

The Cucker-Smale model:

$$\begin{cases} \dot{x}_i = v_i \in \mathbb{R}^d \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} a(\|x_j - x_i\|)(v_j - v_i) \in \mathbb{R}^d \end{cases},$$

where $a(t) := a_\beta(t) = \frac{1}{(1+t^2)^\beta}$, $\beta > 0$ governs the rate of communication. In matrix notation:

$$\begin{cases} \dot{x} = v \\ \dot{v} = -L_x v \end{cases}$$

where L_x is the Laplacian of the matrix[1] $(a(\|x_j - x_i\|)/N)_{i,j=1}^N$ and depends on x.

[1] The *Laplacian* L of A is given by $L = D - A$, with $D = \mathrm{diag}(d_1, \ldots, d_N)$ and $d_k = \sum_{j=1}^N a_{kj}$

A parametric model of homo-to-hetero-philia

The Cucker-Smale model:

$$\begin{cases} \dot{x}_i = v_i \in \mathbb{R}^d \\ \dot{v}_i = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} a(\|x_j - x_i\|)(v_j - v_i) \in \mathbb{R}^d \end{cases},$$

where $a(t) := a_\beta(t) = \frac{1}{(1+t^2)^\beta}$, $\beta > 0$ governs the rate of communication. In matrix notation:

$$\begin{cases} \dot{x} = v \\ \dot{v} = -L_x v \end{cases}$$

where L_x is the Laplacian of the matrix[1] $(a(\|x_j - x_i\|)/N)_{i,j=1}^N$ and depends on x.

- Mean-velocity conservation: $\frac{d}{dt}\bar{v}(t) = \frac{1}{N}\sum_{i=1}^N \dot{v}_i(t) = \frac{1}{N^2}\sum_{i=1}^N \sum_{j=1}^N \frac{v_j - v_i}{(1+\|x_j-x_i\|^2)^\beta} \equiv 0.$

[1] The *Laplacian* L of A is given by $L = D - A$, with $D = \mathrm{diag}(d_1, \ldots, d_N)$ and $d_k = \sum_{j=1}^N a_{kj}$

Conditional consensus emergence for a generic communication rate $a(\cdot)$

Consider the symmetric bilinear form

$$B(u, v) = \frac{1}{2N^2} \sum_{i,j} \langle u_i - u_j, v_i - v_j \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle u_i, v_i \rangle - \langle \bar{u}, \bar{v} \rangle,$$

and

$$X(t) = B(x(t), x(t)), \quad V(t) = B(v(t), v(t)).$$

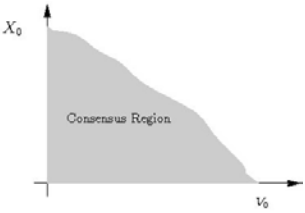
Conditional consensus emergence for a generic communication rate $a(\cdot)$

Consider the symmetric bilinear form

$$B(u,v) = \frac{1}{2N^2}\sum_{i,j}\langle u_i - u_j, v_i - v_j\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle u_i, v_i\rangle - \langle \bar{u}, \bar{v}\rangle,$$

and

$$X(t) = B(x(t), x(t)), \quad V(t) = B(v(t), v(t)).$$



Theorem (Ha-Ha-Kim)

Let $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ be such that $X_0 = B(x_0, x_0)$ and $V_0 = B(v_0, v_0)$ satisfy

$$\gamma(X_0) := \sqrt{N}\int_{\sqrt{NX_0}}^{\infty} a(\sqrt{2}r)dr > \sqrt{V_0}\,.$$

Then the solution with initial data (x_0, v_0) tends to consensus.

Non-consensus events

Consider $\beta = 1$ and $x(t) = x_1(t) - x_2(t)$, $v(t) = v_1(t) - v_2(t)$ relative pos. and vel. of two agents on the line:

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\dfrac{v}{1 + x^2} \end{cases}$$

with initial conditions $x(0) = x_0$ and $v(0) = v_0 > 0$.

Non-consensus events

Consider $\beta = 1$ and $x(t) = x_1(t) - x_2(t)$, $v(t) = v_1(t) - v_2(t)$ relative pos. and vel. of two agents on the line:

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\dfrac{v}{1+x^2} \end{cases}$$

with initial conditions $x(0) = x_0$ and $v(0) = v_0 > 0$.
By direct integration

$$v(t) = -\arctan x(t) + \arctan x_0 + v_0.$$

Hence, if $\arctan x_0 + v_0 > \pi/2 + \varepsilon$ we have

$$v(t) > \pi/2 + \varepsilon - \arctan x(t) > \varepsilon, \quad \forall t \in \mathbb{R}_+.$$

Self-organization Vs organization by intervention

We introduce the notion of *organization via intervention*.

Self-organization Vs organization by intervention

We introduce the notion of *organization via intervention*.

Admissible controls: measurable functions $u = (u_1, \ldots, u_N) : [0, +\infty) \to \mathbb{R}^N$ such that $\sum_{i=1}^N \|u_i(t)\| \leq M$ for every $t > 0$, for a given constant M:

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{j=1}^N a(\|x_j - x_i\|)(v_j - v_i) + u_i \end{cases}$$

for $i = 1, \ldots, N$, and $x_i \in \mathbb{R}^d$, $v_i \in \mathbb{R}^d$.

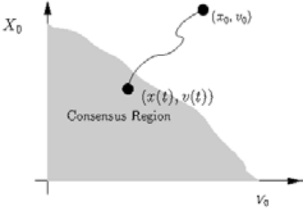
Self-organization Vs organization by intervention

We introduce the notion of *organization via intervention*.

Admissible controls: measurable functions $u = (u_1, \ldots, u_N) : [0, +\infty) \to \mathbb{R}^N$ such that $\sum_{i=1}^N \|u_i(t)\| \leq M$ for every $t > 0$, for a given constant M:

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} a(\|x_j - x_i\|)(v_j - v_i) + u_i \end{cases}$$

for $i = 1, \ldots, N$, and $x_i \in \mathbb{R}^d$, $v_i \in \mathbb{R}^d$.



Our aim is then to find admissible controls steering the system to the consensus region.

Total control

Proposition (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and* $u : [0, T] \to (\mathbb{R}^d)^N$*, with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated solution reaches the consensus region.*

Total control

Proposition (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and* $u : [0, T] \to (\mathbb{R}^d)^N$, *with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated solution reaches the consensus region.*

Proof.

Consider a solution of system with initial data (x_0, v_0) associated with a feedback control $u = -\alpha(v - \bar{v})$, with $0 < \alpha \leq M/(N\sqrt{B(v_0, v_0)})$.

Total control

Proposition (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and* $u : [0, T] \to (\mathbb{R}^d)^N$, *with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated solution reaches the consensus region.*

Proof.

Consider a solution of system with initial data (x_0, v_0) associated with a feedback control $u = -\alpha(v - \bar{v})$, with $0 < \alpha \leq M/(N\sqrt{B(v_0, v_0)})$. Then

$$\begin{aligned}
\frac{d}{dt}V(t) &= \frac{d}{dt}B(v(t), v(t)) \\
&= -2B(L_x v(t), v(t)) + 2B(u(t), v(t)) \\
&\leq 2B(u(t), v(t)) = -2\alpha B(v - \bar{v}, v - \bar{v}) = -2\alpha V(t).
\end{aligned}$$

Total control

Proposition (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and* $u : [0, T] \to (\mathbb{R}^d)^N$, *with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated solution reaches the consensus region.*

Proof.

Consider a solution of system with initial data (x_0, v_0) associated with a feedback control $u = -\alpha(v - \bar{v})$, with $0 < \alpha \leq M/(N\sqrt{B(v_0, v_0)})$. Then

$$
\begin{aligned}
\frac{d}{dt}V(t) = \frac{d}{dt}B(v(t), v(t)) \\
&= -2B(L_x v(t), v(t)) + 2B(u(t), v(t)) \\
&\leq 2B(u(t), v(t)) = -2\alpha B(v - \bar{v}, v - \bar{v}) = -2\alpha V(t).
\end{aligned}
$$

Therefore $V(t) \leq e^{-2\alpha t} V(0)$ and $V(t)$ tends to 0 exponentially fast as $t \to \infty$. Moreover $\sum_{i=1}^N \|u_i\| \leq M$. $\square$

More economical choices?

We wish to make

$$\begin{aligned}\frac{d}{dt}V(t) &= \frac{d}{dt}B(v(t), v(t)) \\ &= -2B(L_x v(t), v(t)) + 2B(u(t), v(t))\end{aligned}$$

the smallest possible and use the minimal amount of intervention:

More economical choices?

We wish to make

$$\frac{d}{dt}V(t) = \frac{d}{dt}B(v(t), v(t))$$
$$= -2B(L_x v(t), v(t)) + 2B(u(t), v(t))$$

the smallest possible and use the minimal amount of intervention: minimize $B(u(t), v(t))$ with additional sparsity constraints.

Greedy sparse control

Theorem (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and a sparse control* $u : [0, T] \to (\mathbb{R}^d)^N$*, with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated AC solution reaches the consensus region.*

Greedy sparse control

Theorem (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and a sparse control* $u : [0, T] \to (\mathbb{R}^d)^N$, *with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated AC solution reaches the consensus region. More precisely, we can choose adaptively the control law explicitly as one of the solutions of the variational problem*

$$\min B(v, u) + \frac{\gamma(x)}{N} \sum_{i=1}^N \|u_i\| \quad \textit{subject to } \sum_{i=1}^N \|u_i\| \leq M \,,$$

where $\gamma(x) = \sqrt{N} \int_{\sqrt{NB(x,x)}}^{\infty} a(\sqrt{2}r) dr$ *threshold by Ha-Ha-Kim.*

Greedy sparse control

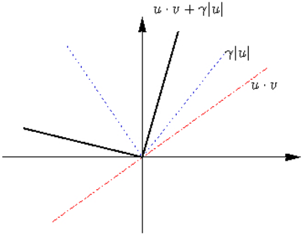
Theorem (Caponigro-F.-Piccoli-Trélat)

For every initial condition $(x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ *and* $M > 0$ *there exist* $T > 0$ *and a sparse control* $u : [0, T] \to (\mathbb{R}^d)^N$*, with* $\sum_{i=1}^N \|u_i(t)\| \leq M$ *for every* $t \in [0, T]$ *such that the associated AC solution reaches the consensus region. More precisely, we can choose adaptively the control law explicitly as one of the solutions of the variational problem*

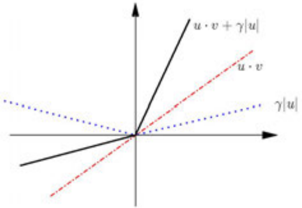
$$\min B(v, u) + \frac{\gamma(x)}{N} \sum_{i=1}^N \|u_i\| \quad \textit{subject to} \; \sum_{i=1}^N \|u_i\| \leq M \,,$$

where $\gamma(x) = \sqrt{N} \int_{\sqrt{NB(x,x)}}^{\infty} a(\sqrt{2}r)dr$ *threshold by Ha-Ha-Kim.*
This choice of the control makes $V(t) = B(v(t), v(t))$ *vanishing in finite time, hence there exists* T *such that* $B(v(t), v(t)) \leq \gamma(x)^2$*,* $t \geq T$*.*

Geometrical interpretation in the scalar case



For $|v| \leq \gamma$ the minimal solution $u \in [-M, M]$ is zero.



For $|v| > \gamma$ the minimal solution $u \in [-M, M]$ is $|u| = M$.

Explicit sparse control

Denote $v_\perp = v - \bar{v}$. We construct the control law from the variational problem.

Explicit sparse control

Denote $v_\perp = v - \bar{v}$. We construct the control law from the variational problem.

If $\|v_{\perp_i}\| \leq \gamma(x)$ for every $i = 1, \ldots, N$, then

$$u_1 = \cdots = u_N = 0$$

Explicit sparse control

Denote $v_\perp = v - \bar{v}$. We construct the control law from the variational problem.
If $\|v_{\perp_i}\| \leq \gamma(x)$ for every $i = 1, \ldots, N$, then

$$u_1 = \cdots = u_N = 0 \Rightarrow \text{reached consensus region}.$$

Explicit sparse control

Denote $v_\perp = v - \bar{v}$. We construct the control law from the variational problem.
If $\|v_{\perp_i}\| \leq \gamma(x)$ for every $i = 1, \ldots, N$, then

$$u_1 = \cdots = u_N = 0 \Rightarrow \text{reached consensus region}.$$

Otherwise there exists a "best index" $i \in \{1, \ldots, N\}$ such that

$$\|v_{\perp_i}\| > \gamma(x) \quad \text{and} \quad \|v_{\perp_i}\| \geq \|v_{\perp_j}\| \quad \text{for every } j = 1, \ldots, N.$$

Therefore we can choose $i \in \{1, \ldots, N\}$ satisfying it, and a control law

$$u_i = -M \frac{v_{\perp_i}}{\|v_{\perp_i}\|}, \quad \text{and} \quad u_j = 0, \quad \text{for every } j \neq i.$$

Explicit sparse control

Denote $v_\perp = v - \bar{v}$. We construct the control law from the variational problem.
If $\|v_{\perp_i}\| \leq \gamma(x)$ for every $i = 1, \dots, N$, then

$$u_1 = \cdots = u_N = 0 \Rightarrow \text{reached consensus region}.$$

Otherwise there exists a "best index" $i \in \{1, \dots, N\}$ such that

$$\|v_{\perp_i}\| > \gamma(x) \quad \text{and} \quad \|v_{\perp_i}\| \geq \|v_{\perp_j}\| \quad \text{for every } j = 1, \dots, N.$$

Therefore we can choose $i \in \{1, \dots, N\}$ satisfying it, and a control law

$$u_i = -M \frac{v_{\perp_i}}{\|v_{\perp_i}\|}, \quad \text{and} \quad u_j = 0, \quad \text{for every } j \neq i.$$

Hence the control acts on the most "stubborn". We may call this control the "shepherd dog strategy".

Explicit sparse control

Denote $v_\perp = v - \bar{v}$. We construct the control law from the variational problem.
If $\|v_{\perp_i}\| \leq \gamma(x)$ for every $i = 1, \ldots, N$, then

$$u_1 = \cdots = u_N = 0 \Rightarrow \text{reached consensus region.}$$

Otherwise there exists a "best index" $i \in \{1, \ldots, N\}$ such that

$$\|v_{\perp_i}\| > \gamma(x) \quad \text{and} \quad \|v_{\perp_i}\| \geq \|v_{\perp_j}\| \quad \text{for every } j = 1, \ldots, N.$$

Therefore we can choose $i \in \{1, \ldots, N\}$ satisfying it, and a control law

$$u_i = -M \frac{v_{\perp_i}}{\|v_{\perp_i}\|}, \quad \text{and} \quad u_j = 0, \quad \text{for every } j \neq i.$$

Hence the control acts on the most "stubborn". We may call this control the "shepherd dog strategy". This choice of the control makes $V(t) = B(v(t), v(t))$ vanishing in finite time, hence there exists T such that $B(v(t), v(t)) \leq \gamma(x)^2$, $t \geq T$.

Instantaneous optimality of the greedy strategy

Consider generic control u (solution of the variation problem) of components

$$u_i(x, v) = \begin{cases} 0 & \text{if } v_{\perp_i} = 0 \\ -\alpha_i \dfrac{v_{\perp_i}}{\|v_{\perp_i}\|} & \text{if } v_{\perp_i} \neq 0 \end{cases}$$

where $\alpha_i \geq 0$ such that $\sum_{i=1}^{N} \alpha_i \leq M$.

Instantaneous optimality of the greedy strategy

Consider generic control u (solution of the variation problem) of components

$$u_i(x, v) = \begin{cases} 0 & \text{if } v_{\perp_i} = 0 \\ -\alpha_i \dfrac{v_{\perp_i}}{\|v_{\perp_i}\|} & \text{if } v_{\perp_i} \neq 0 \end{cases}$$

where $\alpha_i \geq 0$ such that $\sum_{i=1}^N \alpha_i \leq M$.

Theorem (Caponigro-F.-Piccoli-Trélat)

The 1-sparse control is the minimizer of

$$\mathcal{R}(t, u) := \mathcal{R}(t) = \frac{d}{dt} V(t),$$

among all the control of the previous form.

∗ A policy maker, who is not allowed to have prediction on future developments, should always consider more favorable to intervene with stronger actions on the fewest possible instantaneous optimal leaders than trying to control more agents with minor strength.
∗ Homophilious society can be stabilized by parsiminious interventions!

Homophilious societies are sparsely stabilizable – 1

If we allow external intervention, the CS system in the homophilious regime ($\beta > \frac{1}{2}$)

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \frac{1}{N} \sum_{j=1}^{N} a(|x_i - x_j|)(x_j - x_i) + u_i \end{cases}$$

can be stabilized for any initial condition by using only sparse controls, i.e., zero for almost every agent [2].

[2] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat, *Sparse stabilization and control of alignment models*, Math. Models Methods Applied Sciences, '14

Homophilious societies are sparsely stabilizable – 1

If we allow external intervention, the CS system in the homophilious regime ($\beta > \frac{1}{2}$)

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{j=1}^{N} a(|x_i - x_j|)\,(x_j - x_i) \; + \; u_i \end{cases}$$

can be stabilized for any initial condition by using only sparse controls, i.e., zero for almost every agent [2].

The sparse control acts on the most “stubborn” agent at every time, like the “shepherd dog strategy”.

[2] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat, *Sparse stabilization and control of alignment models*, Math. Models Methods Applied Sciences, ’14

Homophilious societies are sparsely stabilizable – 2

For the CD system in the homophilious regime (that is $E(0) \geq \theta$), the shepard dog strategy stabilization of

$$\begin{cases} \dot{x}_i &= v_i, \\ \dot{v}_i &= -bv_i + \frac{1}{N}\sum_{j=1}^{N}(a-f)(|x_i - x_j|)(x_j - x_i) + u_i, \end{cases}$$

with $b \equiv 0$ occurs under the additional hypothesis

$$\theta > E(0) \exp\left(-\frac{2\sqrt{3}}{9}\frac{N\|\bar{v}(0)\|^3}{E(0)^{3/2}}\right)$$

Homophilious societies are sparsely stabilizable – 2

For the CD system in the homophilious regime (that is $E(0) \geq \theta$), the shepard dog strategy stabilization of

$$\begin{cases} \dot{x}_i &= v_i, \\ \dot{v}_i &= -bv_i + \frac{1}{N}\sum_{j=1}^{N}(a-f)(|x_i - x_j|)(x_j - x_i) + u_i, \end{cases}$$

with $b \equiv 0$ occurs under the additional hypothesis

$$\theta > E(0)\exp\left(-\frac{2\sqrt{3}}{9}\frac{N\|\bar{v}(0)\|^3}{E(0)^{3/2}}\right)$$

The shepherd dog strategy does not work for every initial condition!

Observing the future: sparse optimal control

The problem is to minimize, for a given $\gamma > 0$

$$\mathcal{J}(u) = \int_0^T \frac{1}{N} \sum_{i=1}^N \left(\left(v_i(t) - \frac{1}{N} \sum_{j=1}^N v_j(t) \right)^2 \quad + \quad \gamma \| u_i(t) \| \right) dt,$$

$$\text{s.t.} \ \sum_{i=1}^N \|u_i\| \leq M$$

where the state is a trajectory of the control system

$$\begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{j=1}^N a(\|x_j - x_i\|)(v_j - v_i) + u_i \end{cases}$$

with initial constraint

$$(x(0), v(0)) = (x_0, v_0) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N.$$

Beyond a greedy approach: sparse optimal control

Theorem (Caponigro-F.-Piccoli-Trélat)

For every (x_0, v_0) *in* $(\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$*, for every* $M > 0$*, and for every* $\gamma > 0$ *the optimal control problem has an optimal solution. The optimal control* $u(t)$ *is "usually" instantaneously a vector with at most one nonzero coordinate.*

Beyond a greedy approach: sparse optimal control

Theorem (Caponigro-F.-Piccoli-Trélat)

For every (x_0, v_0) in $(\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$, for every $M > 0$, and for every $\gamma > 0$ the optimal control problem has an optimal solution. The optimal control $u(t)$ is "usually" instantaneously a vector with at most one nonzero coordinate.

The PMP ensures the existence of $\lambda \geq 0$ and of a nontrivial covector $(p_x, p_v) \in (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$ satisfying the adjoint equations, for $i = 1, \ldots, N$,

$$\begin{cases} \dot{p}_{x_i} = \frac{1}{N} \sum_{j=1}^{N} \frac{a(\|x_j - x_i\|)}{\|x_j - x_i\|} \langle x_j - x_i, v_j - v_i \rangle (p_{v_j} - p_{v_i}) \\ \dot{p}_{v_i} = -p_{x_i} - \frac{1}{N} \sum_{j \neq i} a(\|x_j - x_i\|)(p_{v_j} - p_{v_i}) - 2\lambda v_i + \frac{2\lambda}{N} \sum_{j=1}^{N} v_j. \end{cases}$$

The application of the PMP leads to minimize

$$\min \sum_{i=1}^{N} \langle p_{v_i}, u_i \rangle + \lambda\gamma \sum_{i=1}^{N} \|u_i\|, \quad \text{subject to } \sum_{i=1}^{N} \|u_i\| \leq M.$$

What if the population is too large $N \approx \infty$?

Mean-field (sparse) optimal control?

What if $N \to \infty$?

Mean-field (sparse) optimal control?

What if $N \to \infty$? Did you know that the term "curse of dimensionality" was first introduced by Richard E. Bellman precisely for high-dimensional optimal control problems?

Mean-field (sparse) optimal control?

What if $N \to \infty$? Did you know that the term "curse of dimensionality" was first introduced by Richard E. Bellman precisely for high-dimensional optimal control problems?

How can we define an approximating "infinite dimensional" sparse optimal control?

Mean-field (sparse) optimal control?

What if $N \to \infty$? Did you know that the term "curse of dimensionality" was first introduced by Richard E. Bellman precisely for high-dimensional optimal control problems?

How can we define an approximating "infinite dimensional" sparse optimal control?

We consider a perhaps natural control problem:

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (H * \mu_N)(x_i, v_i) + u_i, i = 1, \dots N, \quad t \in [0, T], \\ \text{where } , \mu_N = \frac{1}{N} \sum_{j=1}^N \delta_{(x_j, v_j)}, \end{cases}$$

controlled by the minimizer of the cost functional

$$\mathcal{J}(u) := \int_0^T \int_{\mathbb{R}^{2d}} \left(L(x, v, \mu_N) d\mu_N(t, x, v) + \frac{1}{N} \sum_{i=1}^N \|u_i\| \right) dt,$$

Mean-field (sparse) optimal control?

What if $N \to \infty$? Did you know that the term "curse of dimensionality" was first introduced by Richard E. Bellman precisely for high-dimensional optimal control problems?

How can we define an approximating "infinite dimensional" sparse optimal control?

We consider a perhaps natural control problem:

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (H * \mu_N)(x_i, v_i) + u_i, i = 1, \dots N, \quad t \in [0, T], \\ \text{where } , \mu_N = \frac{1}{N} \sum_{j=1}^{N} \delta_{(x_j, v_j)}, \end{cases}$$

controlled by the minimizer of the cost functional

$$\mathcal{J}(u) := \int_0^T \int_{\mathbb{R}^{2d}} \left(L(x, v, \mu_N) d\mu_N(t, x, v) + \frac{1}{N} \sum_{i=1}^{N} \|u_i\| \right) dt,$$

Which topology on $\mu_N = \frac{1}{N} \sum_{j=1}^{N} \delta_{(x_j, v_j)}$? Which topology on $\nu_N = \frac{1}{N} \sum_{j=1}^{N} u_i \delta_{(x_j, v_j)}$?

Too weak convergence

The compactness of the problem is way too weak

$$\nu_N = \frac{1}{N}\sum_{j=1}^{N} u_i \delta_{(x_j,v_j)} \rightharpoonup \nu, \quad \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{(x_j,v_j)} \rightharpoonup \mu,$$

as it can happen that $\nu \perp \mu$.

Too weak convergence

The compactness of the problem is way too weak

$$\nu_N = \frac{1}{N}\sum_{j=1}^{N} u_i \delta_{(x_j,v_j)} \rightharpoonup \nu, \quad \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{(x_j,v_j)} \rightharpoonup \mu,$$

as it can happen that $\nu \perp \mu$. Natural limit equation

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu)\mu + \nu\right].$$

Too weak convergence

The compactness of the problem is way too weak

$$\nu_N = \frac{1}{N}\sum_{j=1}^{N} u_i \delta_{(x_j,v_j)} \rightharpoonup \nu, \quad \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{(x_j,v_j)} \rightharpoonup \mu,$$

as it can happen that $\nu \perp \mu$. Natural limit equation

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu)\mu + \nu \right].$$

Steering the cloud μ by means of the toothpicks ν!!

Too weak convergence

The compactness of the problem is way too weak

$$\nu_N = \frac{1}{N}\sum_{j=1}^{N} u_i \delta_{(x_j,v_j)} \rightharpoonup \nu, \quad \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{(x_j,v_j)} \rightharpoonup \mu,$$

as it can happen that $\nu \perp \mu$. Natural limit equation

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu)\mu + \nu \right].$$

Steering the cloud μ by means of the toothpicks ν!! Then relax it by $\nu = f\mu$ and obtain

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu + f)\,\mu \right],$$

Too weak convergence

The compactness of the problem is way too weak

$$\nu_N = \frac{1}{N}\sum_{j=1}^{N} u_i \delta_{(x_j,v_j)} \rightharpoonup \nu, \quad \mu_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{(x_j,v_j)} \rightharpoonup \mu,$$

as it can happen that $\nu \perp \mu$. Natural limit equation

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu)\mu + \nu\right].$$

Steering the cloud μ by means of the toothpicks ν!! Then relax it by $\nu = f\mu$ and obtain

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[\left(H * \mu + f\right)\mu\right],$$

but $f \in L^1_\mu(\mathbb{R}^{2d}, \mathbb{R}^d)$ only and no well-posedness can be expected!

Mean-field Sparse Optimal Control?

"Ultimately it would be good to have a theory that combined both the collective behaviour of a large number of "ordinary" agents with the decisions of a few key players of unusually large (relative) influence – some complicated combination of PDE and game theory, presumably – but our current mathematical technology is definitely insufficient for even a zeroth approximation to this task".

– Terry Tao, January 7, 2010

https://terrytao.wordpress.com/2010/01/07/mean-field-equations

A natural relaxation: smoother controls

Definition

For a horizon time $T > 0$, and an exponent $1 \leq q < +\infty$ we fix a control bound function $\ell \in L^q(0, T)$. The class of admissible control functions $\mathcal{F}_\ell([0, T])$ is so defined: $f \in \mathcal{F}_\ell([0, T])$ if and only if

(i) $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^d$ is a Carathéodory function,

(ii) $f(t, \cdot) \in W^{1,\infty}_{loc}(\mathbb{R}^n, \mathbb{R}^d)$ for almost every $t \in [0, T]$, and

(iii) $\|f(t, 0)\| + \mathrm{Lip}(f(t, \cdot), \mathbb{R}^d) \leq \ell(t)$ for almost every $t \in [0, T]$.

Mean-field optimal control

Theorem (F. and Solombrino)

Assume that we are given maps H, L, and ψ as in assumptions (H), (L), and (Ψ). For $N \in \mathbb{N}$ and an initial datum $((x_N^0)_1, \dots, (x_N^0)_N, (v_N^0)_1, \dots, (v_N^0)_N) \in B(0, R_0) \subset (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N$, for $R_0 > 0$ independent of N, we consider

$$\min_{f \in \mathcal{F}_\ell} \int_0^T \int_{\mathbb{R}^{2d}} [L(x, v, \mu_N(t, x, v)) + \psi(f(t, x, v))] \, d\mu_N(t, x, v) dt,$$

where $\mu_N(t, x, v) = \frac{1}{N} \sum_{j=1}^N \delta_{(x_j(t), v_j(t))}(x, v)$, constrained by being the solution of

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = (H * \mu_N)(x_i, v_i) + f(t, x_i, v_i), \quad i = 1, \dots N, \quad t \in [0, T], \end{cases}$$

with initial datum $(x(0), v(0)) = (x_N^0, v_N^0)$ and, for consistency, we set

$$\mu_N^0 = \frac{1}{N} \sum_{i=1}^M \delta_{((x_N^0)_i, (v_N^0)_i)}(x, v).$$

For all $N \in \mathbb{N}$ let us denote the function $f_N \in \mathcal{F}_\ell$ as a solution of the finite dimensional optimal control problem.

Mean-field optimal control

If there exists a compactly supported $\mu_0 \in \mathcal{P}_1(\mathbb{R}^{2d})$ such that $\lim_{N\to\infty} \mathcal{W}_1(\mu_N^0, \mu^0) = 0$, then there exists a subsequence $(f_{N_k})_{k\in\mathbb{N}}$ and a function $f_\infty \in \mathcal{F}_\ell$ such that f_{N_k} converges to f_∞ and f_∞ is a solution of the infinite dimensional optimal control problem

$$\min_{f\in\mathcal{F}_\ell} \int_0^T \int_{\mathbb{R}^{2d}} \left[L(x, v, \mu(t, x, v)) + \psi(f(t, x, v))\right] d\mu(t, x, v)dt,$$

where $\mu : [0, T] \to \mathcal{P}_1(\mathbb{R}^{2d})$ is the unique weak solution of

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu + f)\, \mu\right],$$

with initial datum $\mu(0) := \mu^0$ and forcing term f.

Mean-field optimal control

If there exists a compactly supported $\mu_0 \in \mathcal{P}_1(\mathbb{R}^{2d})$ such that $\lim_{N\to\infty} \mathcal{W}_1(\mu_N^0, \mu^0) = 0$, then there exists a subsequence $(f_{N_k})_{k\in\mathbb{N}}$ and a function $f_\infty \in \mathcal{F}_\ell$ such that f_{N_k} converges to f_∞ and f_∞ is a solution of the infinite dimensional optimal control problem

$$\min_{f\in\mathcal{F}_\ell} \int_0^T \int_{\mathbb{R}^{2d}} \left[L(x, v, \mu(t, x, v)) + \psi(f(t, x, v))\right] d\mu(t, x, v)dt,$$

where $\mu : [0, T] \to \mathcal{P}_1(\mathbb{R}^{2d})$ is the unique weak solution of

$$\frac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * \mu + f)\, \mu\right],$$

with initial datum $\mu(0) := \mu^0$ and forcing term f.

The proof is based on the simultaneous development of the mean-field limit for the equation and the Γ-limit for the optimization of the control.

Sparse optimal control? Mixing diffuse and granular

Let us now consider a controlled system with m leaders for $m \gg N$

$$\begin{cases} \dot{y}_k = w_k, & \\ \dot{w}_k = H * \mu_N(y_k, w_k) + H * \mu_m(y_k, w_k) + u_k & k = 1, \dots m, \\ \dot{x}_i = v_i, & \\ \dot{v}_i = H * \mu_N(x_i, v_i) + H * \mu_m(x_i, v_i) & i = 1, \dots N, \end{cases}$$

Sparse optimal control? Mixing diffuse and granular

Let us now consider a controlled system with m leaders for $m \gg N$

$$\begin{cases} \dot{y}_k = w_k, \\ \dot{w}_k = H * \mu_N(y_k, w_k) + H * \mu_m(y_k, w_k) + u_k & k = 1, \dots m, \\ \dot{x}_i = v_i, \\ \dot{v}_i = H * \mu_N(x_i, v_i) + H * \mu_m(x_i, v_i) & i = 1, \dots N, \end{cases}$$

For $N \to \infty$ the limit dynamics is

$$\begin{cases} \dot{y}_k = w_k, \\ \dot{w}_k = H * (\mu + \mu_m)(y_k, w_k) + u_k, & k = 1, \dots m, \\ \partial_t \mu + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * (\mu + \mu_m))\, \mu \right], \end{cases}$$

where the weak solutions of the equations have to be interpreted in the Carathéodory sense.

Mixing diffuse and granular

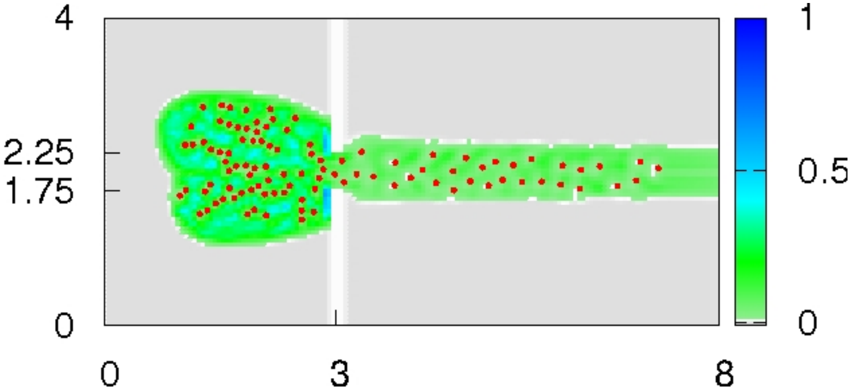


Figure : A mixed granular-diffuse crowd leaving a room through a door. This figure was kindly provided by the authors.

Main result on mean-field sparse control, I

Denote $\mathcal{X} := \mathbb{R}^{2d\times m} \times \mathcal{P}(\mathbb{R}^{2d})$.

Theorem (F., Piccoli, and Rossi)

Let H and L be maps satisfying conditions (H) and (L) respectively. Given an initial datum $(y^0, w^0, \mu^0) \in \mathcal{X}$, with μ^0 compactly supported, $\mathrm{supp}(\mu^0) \subset B(0,R)$, $R > 0$, the optimal control problem

$$\min_{u\in L^1([0,T],\mathcal{U})} \int_0^T \left\{ L(y(t), w(t), \mu(t)) + \frac{1}{m}\sum_{k=1}^m \|u_k(t)\| \right\} dt,$$

has solutions, where the triplet (y, w, μ) defines the unique solution of

$$\begin{cases} \dot{y}_k = w_k, \\ \dot{w}_k = H * (\mu + \mu_m)(y_k, w_k) + u_k, & k = 1, \dots m, \quad t \in [0,T] \\ \partial_t \mu + v \cdot \nabla_x \mu = \nabla_v \cdot \left[(H * (\mu + \mu_m))\, \mu \right], \end{cases}$$

with initial datum (y^0, w^0, μ^0) and control u, and

$$\mu_m(t) = \frac{1}{m}\sum_{k=1}^n \delta_{(y_k(t), w_k(t))}.$$

Main result on mean-field sparse control, II

Moreover, solutions to the problem can be constructed as weak limits u^* of sequences of optimal controls u_N^* of the finite dimensional problems

$$\min_{u\in L^1([0,T],\mathcal{U})} \int_0^T \left\{ L(y_N(t), w_N(t), \mu_N(t)) + \frac{1}{m}\sum_{k=1}^m \|u_k(t)\| \right\} dt,$$

where $\mu_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{(x_{i,N}(t), v_{i,N}(t))}$ and $\mu_{m,N}(t) = \frac{1}{m}\sum_{k=1}^m \delta_{(y_{k,N}(t), w_{k,N}(t))}$ are the time-dependent atomic measures supported on the trajectories defining the solution of the system

$$\begin{cases} \dot{y}_k = w_k, \\ \dot{w}_k = H * \mu_N(y_k, w_k) + H * \mu_{m,M}(y_k, w_k) + u_k & k = 1, \dots m, \quad t \in [0, T], \\ \dot{x}_i = v_i, \\ \dot{v}_i = H * \mu_N(x_i, v_i) + H * \mu_{m,M}(x_i, v_i) & i = 1, \dots N, \quad t \in [0, T], \end{cases}$$

with initial datum (y^0, w^0, x_N^0, v_N^0), control u, and $\mu_N^0 = \frac{1}{N}\sum_{i=1}^N \delta_{(x_i^0, v_i^0)}$ is such that $\mathcal{W}_1(\mu_N^0, \mu^0) \to 0$ for $N \to +\infty$.

Evacuating an unknown environment

Simulations I

Simulations II

Learning the dynamics

Consider the dynamics

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j \neq i} a(\|x_i - x_j\|)(x_j - x_i), \quad i = 1, \dots, N.$$

with $a \in X = \left\{ b : \mathbb{R}_+ \to \mathbb{R} \,|\, b \in L^\infty(\mathbb{R}_+) \cap W^{1,\infty}_{loc}(\mathbb{R}_+) \right\}$.

Learning the dynamics

Consider the dynamics

$$\dot{x}_i(t) = \frac{1}{N}\sum_{j\neq i} a(\|x_i - x_j\|)(x_j - x_i), \quad i = 1, \ldots, N.$$

with $a \in X = \left\{b : \mathbb{R}_+ \to \mathbb{R} \mid b \in L^\infty(\mathbb{R}_+) \cap W^{1,\infty}_{loc}(\mathbb{R}_+)\right\}$. Can we "learn" the interaction function a from observations of the dynamics?

A least square functional

As an approximation to a we seek for a minimizer of the following *discrete error functional*

$$\mathcal{E}_N(\widehat{a}) = \frac{1}{T}\int_0^T \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{N}\sum_{j=1}^{N}\left(\widehat{a}(\|x_i(t)-x_j(t)\|)(x_i(t)-x_j(t)) - \dot{x}_i(t)\right)\right\|^2 dt,$$

among all functions $\widehat{a} \in X$.

A least square functional

As an approximation to a we seek for a minimizer of the following *discrete error functional*

$$\mathcal{E}_N(\widehat{a}) = \frac{1}{T}\int_0^T \frac{1}{N}\sum_{i=1}^N \left\| \frac{1}{N}\sum_{j=1}^N \left(\widehat{a}(\|x_i(t)-x_j(t)\|)(x_i(t)-x_j(t)) - \dot{x}_i(t)\right)\right\|^2 dt,$$

among all functions $\widehat{a} \in X$. In particular, given a finite dimensional space $V \subset X$, we consider the minimizer:

$$\widehat{a}_{N,V} = \arg\min_{\widehat{a}\in V} \mathcal{E}_N(\widehat{a}).$$

A least square functional

As an approximation to a we seek for a minimizer of the following *discrete error functional*

$$\mathcal{E}_N(\widehat{a}) = \frac{1}{T}\int_0^T \frac{1}{N}\sum_{i=1}^N \left\| \frac{1}{N}\sum_{j=1}^N (\widehat{a}(\|x_i(t) - x_j(t)\|)(x_i(t) - x_j(t)) - \dot{x}_i(t)) \right\|^2 dt,$$

among all functions $\widehat{a} \in X$. In particular, given a finite dimensional space $V \subset X$, we consider the minimizer:

$$\widehat{a}_{N,V} = \arg\min_{\widehat{a} \in V} \mathcal{E}_N(\widehat{a}).$$

The fundamental question is

(Q) For which choice of the approximating spaces $V \in \Lambda$ (we assume here that Λ is a countable family of invading subspaces of X) does $\widehat{a}_{N,V} \to a$ for $N \to \infty$ and $V \to X$ and in which topology should this convergence hold?

Mean-field limit

The empirical measure $\mu_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}$ weakly converges for $N \to \infty$ to the probability measure valued trajectory $t \to \mu(t)$ satisfying weakly the equation

$$\partial_t \mu(t) = -\nabla \cdot ((H[a] * \mu(t))\mu(t)), \quad \mu(0) = \mu^0.$$

where $H[a](x) = -a(\|x\|)x$, for $x \in \mathbb{R}^d$.

Mean-field limit

The empirical measure $\mu_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}$ weakly converges for $N \to \infty$ to the probability measure valued trajectory $t \to \mu(t)$ satisfying weakly the equation

$$\partial_t \mu(t) = -\nabla \cdot ((H[a] * \mu(t))\mu(t)), \quad \mu(0) = \mu^0.$$

where $H[a](x) = -a(\|x\|)x$, for $x \in \mathbb{R}^d$. We define

$$\mathcal{E}(\widehat{a}) = \frac{1}{T}\int_0^T \int_{\mathbb{R}^d} \left\| \big(H[\widehat{a}] - H[a]\big) * \mu(t) \right\|^2 d\mu(t)(x) dt,$$

Coercivity property I

By Jensen inequality

$$\begin{aligned}
\mathcal{E}(\widehat{a}) &\leq \frac{1}{T}\int_0^T \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\widehat{a}(\|x-y\|) - a(\|x-y\|)|^2 \|x-y\|^2 d\mu(t)(x) d\mu(t)(y) dt \\
&= \frac{1}{T}\int_0^T \int_{\mathbb{R}_+} \left|\widehat{a}(s) - a(s)\right|^2 s^2 d\varrho(t)(s) dt \qquad (1)
\end{aligned}$$

where $\varrho(t) = (\|x-y\|_{\#}\mu_x(t) \otimes \mu_y(t))$.

Coercivity property I

By Jensen inequality

$$\begin{aligned}\mathcal{E}(\widehat{a}) &\leq \frac{1}{T}\int_0^T \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\widehat{a}(\|x-y\|) - a(\|x-y\|)|^2 \|x-y\|^2 d\mu(t)(x) d\mu(t)(y) dt \\ &= \frac{1}{T}\int_0^T \int_{\mathbb{R}_+} \left|\widehat{a}(s) - a(s)\right|^2 s^2 d\varrho(t)(s) dt \end{aligned} \tag{1}$$

where $\varrho(t) = (\|x-y\|_{\#}\mu_x(t) \otimes \mu_y(t))$. We define the prob. measure

$$\bar{\rho} := \frac{1}{T}\int_0^T \varrho(t) dt.$$

Coercivity property I

By Jensen inequality

$$\begin{aligned}\mathcal{E}(\widehat{a}) &\leq \frac{1}{T}\int_0^T \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\widehat{a}(\|x-y\|) - a(\|x-y\|)|^2 \|x-y\|^2 d\mu(t)(x) d\mu(t)(y) dt \\ &= \frac{1}{T}\int_0^T \int_{\mathbb{R}_+} \left|\widehat{a}(s) - a(s)\right|^2 s^2 d\varrho(t)(s) dt \qquad (1)\end{aligned}$$

where $\varrho(t) = (\|x-y\|_{\#}\mu_x(t)\otimes\mu_y(t))$. We define the prob. measure

$$\bar{\rho} := \frac{1}{T}\int_0^T \varrho(t) dt.$$

Finally we define the weighted measure

$$\rho(A) := \int_A s^2 d\bar{\rho}(s),$$

Coercivity property I

By Jensen inequality

$$\begin{aligned}\mathcal{E}(\widehat{a}) &\leq \frac{1}{T}\int_0^T \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\widehat{a}(\|x-y\|) - a(\|x-y\|)|^2 \|x-y\|^2 d\mu(t)(x) d\mu(t)(y) dt \\ &= \frac{1}{T}\int_0^T \int_{\mathbb{R}_+} \left|\widehat{a}(s) - a(s)\right|^2 s^2 d\varrho(t)(s) dt \qquad (1)\end{aligned}$$

where $\varrho(t) = (\|x-y\|_{\#}\mu_x(t) \otimes \mu_y(t))$. We define the prob. measure

$$\bar{\rho} := \frac{1}{T}\int_0^T \varrho(t) dt.$$

Finally we define the weighted measure

$$\rho(A) := \int_A s^2 d\bar{\rho}(s),$$

Then one can reformulate (1) in a very compact form as follows

$$\mathcal{E}(\widehat{a}) \leq \int_{\mathbb{R}_+} \left|\widehat{a}(s) - a(s)\right|^2 d\rho(s) = \|\widehat{a} - a\|^2_{L^2(\mathbb{R}_+,\rho)}.$$

Coercivity property II

To establish coercivity of the learning problem it is essential to assume that there exists $c_T > 0$ such that also the following additional lower bound holds

$$c_T \|\widehat{a} - a\|^2_{L_2(\mathbb{R}_+,\rho)} \leq \mathcal{E}(\widehat{a}),$$

for all relevant $\widehat{a} \in X \cap L^2(\mathbb{R}_+, \rho)$. This crucial assumption eventually determines also the natural space $X \cap L^2(\mathbb{R}_+, \rho)$ for the solutions.

Uniform approximation property

For $M > 0$ and an interval $K = [0, 2R]$ define the set

$$X_{M,K} = \left\{ b \in W^{1,\infty}(K) : \|b\|_{L^\infty(K)} + \|b'\|_{L^\infty(K)} \leq M \right\}.$$

Additionally for every $N \in \mathbb{N}$, let V_N be a closed subset of $X_{M,K}$ w.r.t. the uniform convergence on K with the following *uniform approximation property*: for all $b \in X_{M,K}$ there exists a sequence $(b_N)_{N \in \mathbb{N}}$ converging uniformly to b on K and such that $b_N \in V_N$ for every $N \in \mathbb{N}$.

Main learnability result I

Theorem (Bongini, F., Hansen, Maggioni, 2015)

Fix $M \geq \|a\|_{L^\infty(K)} + \|a'\|_{L^\infty(K)}$ for $K = [0, 2R]$, for $R > 0$ large enough. For every $N \in \mathbb{N}$, let V_N be a closed subset of $X_{M,K}$ w.r.t. the uniform convergence on K with the uniform approximation property. Then the minimizers

$$\widehat{a}_N \in \arg\min_{\widehat{a} \in V_N} \mathcal{E}_N(\widehat{a}).$$

converge uniformly up to subsequ. for $N \to \infty$ to a continuous function $\widehat{a} \in X_{M,K}$ such that $\mathcal{E}(\widehat{a}) = 0$.

Main learnability result II

If we additionally assume the coercivity condition, then $\widehat{a} = a$ *in* $L_2(\mathbb{R}_+, \rho)$*. Moreover, in this latter case, if there exist rates* $\alpha, \beta > 0$*, constants* $C_1, C_2 > 0$*, and a sequence* $(a_N)_{N \in \mathbb{N}}$ *of elements* $a_N \in V_N$ *such that*

$$\|a - a_N\|_{L^\infty(K)} \leq C_1 N^{-\alpha},$$

and

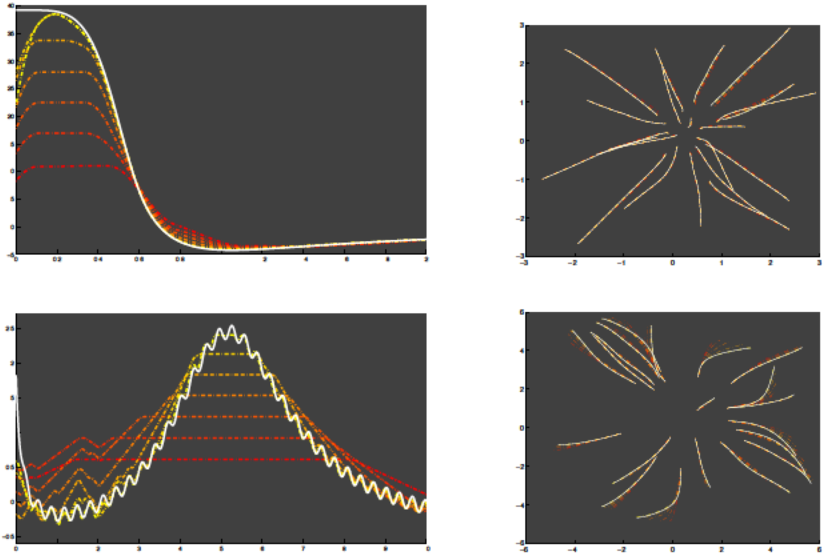
$$\mathcal{W}_1(\mu_0^N, \mu_0) \leq C_2 N^{-\beta},$$

then there exists a constant $C_3 > 0$ *such that*

$$\|a - \widehat{a}_N\|^2_{L^2(\mathbb{R}_+,\rho)} \leq C_3 N^{-\min\{\alpha,\beta\}},$$

for all $N \in \mathbb{N}$*. In particular, in this case, it is the entire sequence* $(\widehat{a}_N)_{N \in \mathbb{N}}$ *(and not only subsequences) to converge to* a *in* $L^2(\mathbb{R}_+, \rho)$*.*

Numerical experiments



Conclusion

- We presented dynamical systems with self-organization features $\Rightarrow$ *organization by external intervention*.

Conclusion

- We presented dynamical systems with self-organization features $\Rightarrow$ *organization by external intervention*.
- We proved that the most effective *greedy strategy* is by instantaneous 1-sparse controls.

Conclusion

- We presented dynamical systems with self-organization features $\Rightarrow$ *organization by external intervention*.
- We proved that the most effective *greedy strategy* is by instantaneous 1-sparse controls.

Conclusion

- We presented dynamical systems with self-organization features $\Rightarrow$ *organization by external intervention*.
- We proved that the most effective *greedy strategy* is by instantaneous 1-sparse controls.
- We presented relaxations of the (sparse) finite dimensional optimal control problems and a general technique to derive their mean-field limits.

Conclusion

- We presented dynamical systems with self-organization features $\Rightarrow$ *organization by external intervention*.
- We proved that the most effective *greedy strategy* is by instantaneous 1-sparse controls.
- We presented relaxations of the (sparse) finite dimensional optimal control problems and a general technique to derive their mean-field limits.
- We showed recent results on learnability of systems modeling social dynamics.

A few info

- **WWW:** http://www-m15.ma.tum.de/
- **References**:

  - M. Bongini, M. Fornasier, F. Rossi, and F. Solombrino, *Mean-field Pontryagin maximum principle*, submitted
  - G. Albi, M. Bongini, E. Cristiani, and D. Kalise, *Invisible control of self-organizing agents leaving unknown environments*, SIAM J. Appl. Math.
  - M. Bongini and M. Fornasier, *Sparse stabilization of dynamical systems driven by attraction and avoidance forces*, Networks and Heterogeneous Media, 2014
  - M. Bongini, M. Fornasier, M. Hansen, and M. Maggioni, *Inferring interaction rules from observations of evolutive systems I: the variational approach*, to appear in M3AS
  - M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat, *Sparse stabilization and optimal control of the Cucker-Smale model*, M3AS, 2014
  - M. Fornasier, F. Rossi, and B. Piccoli, *Mean field sparse optimal control*, Phil. Trans. Royal Soc. A, 2014
  - M. Fornasier and F. Solombrino, *Mean field optimal control*, ESAIM: Control, Optimization, and Calculus of Variations, 2014

References 2

- **WWW:** http://www-m15.ma.tum.de/

- **References**:

  - M. Fornasier, Learning and sparse control of multiagent systems, In Proc. 7thECM, pp. 1-31, 2016.
  - G. Albi, Y.-P. Choi, M. Fornasier and D. Kalise. Mean Field Control Hierarchy, to appear in Applied Mathematics and Optimization (special issue on Mean-Field Games)