Optimal Control of Heterogeneous, Spatially Distributed Systems

Ali Jadbabaie

Electrical and Systems Engineering and GRASP Laboratory
University of Pennsylvania

With Nader Motee and Alireza Tahbaz-Salehi
Complexity: dynamics vs. size

A. Jadabaie “Optimal control of spatially distributed systems”
Complexity: dynamics vs. size

System size

State dimensionality

A. Jadabaie "Optimal control of spatially distributed systems"
Complexity: dynamics vs. size

System size

A. Jadabaie "Optimal control of spatially distributed systems"
State dimensionality

Complexity: dynamics vs. size

System size

A. Jadbabaie “Optimal control of spatially distributed systems”
Complexity: dynamics vs. size

System size

State dimensionality

Heterogeneous spatially distributed systems
Novel Type of Phase Transition in a System of Self-Driven Particles

Tamás Vicsek,1,2 András Czirók,1 Eshel Ben-Jacob,3 Inon Cohen,3 and Ofer Shochet3
1Department of Atomic Physics, Eötvös University, Budapest, Puskin u 5-7, 1088 Hungary
2Institute for Technical Physics, Budapest, P.O.B. 76, 1325 Hungary
3School of Physics, Tel-Aviv University, 69978 Tel-Aviv, Israel
(Received 25 April 1994)

A simple model with a novel type of dynamics is introduced in order to investigate the emergence of self-ordered motion in systems of particles with biologically motivated interaction. In our model particles are driven with a constant absolute velocity and at each time step assume the average direction of motion of the particles in their neighborhood with some random perturbation (\(\eta\)) added. We present numerical evidence that this model results in a kinetic phase transition from no transport (zero average velocity, \(|v_\alpha| = 0\)) to finite net transport through spontaneous symmetry breaking of the rotational symmetry. The transition is continuous, since \(|v_\alpha|\) is found to scale as \((\eta_c - \eta)^\beta\) with \(\beta \approx 0.45\).

PACS numbers: 87.10.+e, 64.60.-i
Novel Type of Phase Transition in a System of Self-Driven Particles

Tamás Vicsek,1,2 András Czirók,1 Eshel Ben-Jacob,3 Inon Cohen,3 and Ofer Shochet3

1Department of Atomic Physics, Eötvös University, Budapest, Puskin u 5-7, 1088 Hungary
2Department of Technical Physics, Budapest, P.O.B. 76, 1325 Hungary
3Department of Physics, Tel-Aviv University, 69978 Tel-Aviv, Israel

(Received 25 April 1994)

A type of dynamics is introduced in order to investigate the collective motion of particles with biologically motivated interaction. In each time step each particle has a constant absolute velocity and at each time step assume the average velocity of its six nearest neighbors with some random perturbation (η) added. The model results in a kinetic phase transition from no transport (κ = 0) to transport through spontaneous symmetry breaking of the collective motion (κ > 0). The transition is continuous, since |vκ| is found to scale as (κ − η)β with β = 0.3 ± 0.05.

A. Jadabaie “Optimal control of spatially distributed systems”
Synchronization and flocking: emergence of collective behavior in homogeneous systems

Novel Type of Phase Transition in a System of Self-Driven Particles

Tamás Vicsek,1,2 András Czirók,1 Eshel Ben-Jacob,3 Inon Cohen,3 and Ofer Shochet3
1Department of Atomic Physics, Eötvös University, Budapest, Puskin u 5-7, 1088 Hungary

Birds of a feather

A. Jadabaie “Optimal control of spatially distributed systems”
A. Jadbabaie  "Optimal control of spatially distributed systems"

Synchronization and flocking: emergence of collective behavior in homogeneous systems

Mostly simulations, proofs for homogeneous case
Homogeneous systems, consensus and
\[ F_p := (D_p + I)^{-1}(A_p + I), \quad p \in \mathcal{P}, \]
\[ \theta := \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_n \end{bmatrix}', \quad \theta(k+1) = F_{\sigma(k)}\theta(k) \]
Homogeneous systems, consensus and

\[ F_p := (D_p + I)^{-1}(A_p + I), \quad p \in \mathcal{P}, \]
\[ \theta := \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_n \end{bmatrix}', \]
\[ \sigma : \{0, 1, \ldots\} \rightarrow \mathcal{P} \]
is a random switching signal
Homogeneous systems, consensus and

\[ F_p := (D_p + I)^{-1}(A_p + I), \quad p \in \mathcal{P}, \]
\[ \theta := \begin{bmatrix} \theta_1 & \theta_2 & \ldots & \theta_n \end{bmatrix}', \]
\[ \sigma : \{0, 1, \ldots \} \rightarrow \mathcal{P} \]

Repeated local averaging w/ nbrs.

\[ \theta(k + 1) = F_{\sigma(k)} \theta(k) \]

is a random switching signal
Homogeneous systems, consensus and

\[ F_p := (D_p + I)^{-1}(A_p + I), \ p \in \mathcal{P}, \]

\[ \theta := \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_n \end{bmatrix}', \]

\[ \sigma : \{0, 1, \ldots\} \rightarrow \mathcal{P} \]

is a random switching signal

Repeated local averaging w/ nbrs.

Theorem (Tahbaz-Salehi & Jadbabaie ‘08): When the graph process changes randomly (but stationary & ergodic), reaching consensus is a trivial event, i.e., either it happens almost surely or almost never. This is true even when edges are dependent.
Homogeneous systems, consensus and

\[ F_p := (D_p + I)^{-1}(A_p + I), \quad p \in \mathcal{P}, \]

\[ \theta := \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_n \end{bmatrix}', \quad \theta(k + 1) = F_{\sigma(k)}\theta(k) \]

\[ \sigma : \{0, 1, \ldots\} \rightarrow \mathcal{P} \]
is a random switching signal

Theorem (Tahbaz-Salehi & Jadbabaie ‘08): When the graph process changes randomly (but stationary & ergodic), reaching consensus is a trivial event, i.e., either it happens almost surely or almost never. This is true even when edges are dependent

Theorem (Tahbaz Salehi and Jadbabaie ‘08): necessary and sufficient condition for almost sure convergence is \( \lambda_2(E(F)) < 1 \), i.e., the average system needs to reach deterministic agreement.
A general class of systems comprised of subsystems that are:

- countably large (possibly infinite)
- coupled together through:
  - their dynamics
  - a common objective function
- Typical dynamics

\[ \dot{x}_k = A_{kk} x_k + B_{kk} u_k + \sum_{i=1}^{N} A_{ki} x_i \]

With quadratic cost function

Question: When is the optimal control localized in space?
Spatially Distributed Dynamical Systems

- Infinite or finite number of subsystems.
- Infinite-dimensional abstractions allows for a precise mathematical Analysis.
- Our focus will be on spatially distributed linear systems:

\[
\frac{d}{dt} \psi(i, t) = (A\psi)(i, t) + (Bu)(i, t)
\]
\[
y(i, t) = (C\psi)(i, t) + (Du)(i, t)
\]

\(\psi, u, y\) : state, input, and output variables

\(i\) : spatial variable

\(t\) : temporal variable

\(A, B, C, D\) : infinite-dimensional matrices
Distributed Optimal Control

Structural Properties of Spatially Distributed Systems:

\[
\begin{align*}
\text{minimize} & \quad \int_0^\infty \langle \psi, Q\psi \rangle + \langle u, Ru \rangle dt \\
\text{subject to:} & \quad \frac{d\psi}{dt} = A\psi + Bu \\
& \quad u = K\psi \\
& \quad K \in S
\end{align*}
\]

- $S$ : subspace constraint.
- Standard LQR assumptions hold.
- No known tractable algorithm.
- Radner, Witsenhausen, Ho and Chu, Fan, Speyer and Jaensch, Voulgaris, Bamieh, D’Andrea, Dullerud, Tsitsiklis, Blondel, Rantzer, Rotkowitz and Lall, Seiler, Swaroop, Hedrick

A. Jadabaie “Optimal control of spatially distributed systems”
Each block element is a dynamical subsystem.

- The group operation naturally induces a translation operator $T$.
- Fourier transform can be defined on the spatial domain $\mathbb{G}$.
- The spatial structure has lots of symmetries.
- Fourier analysis tools can be used.

**Idea:** Make infinite-dimensional problems look like finite-dimensional ones.

Figure: R. D’Andrea
Optimal Control of spatially invariant systems:  
(Bamieh, Paganini, Dahleh ’02)

Theorem: \( A, B, Q, R \in S \Rightarrow K \in S \)

Locality features: For system defined on a one-dimensional lattice

\[
K = \sum_{n \in \mathbb{Z}} K_n T^n, \quad \|K_n\| \leq \mu e^{-\lambda |n|}, \quad \mu, \lambda > 0
\]
What is the fundamental mathematical difficulty? Asymmetry!

- Heterogeneous subsystems and interconnections.
- The temporal analogy is linear time-invariant (LTI) systems vs. linear time-varying (LTV) systems where Fourier analysis cannot be directly used.
- Develop operator theoretic tools for analysis of spatially-varying distributed systems.

Idea

Extend the notion of analytic continuation from translation-invariant operators to arbitrary operators.
Fig. 1: Distributed System on a one-dimensional lattice

$\text{dis}(k, i) = |k - i|$ 

Fig. 2: A Distributed System with an arbitrary interconnection

A. Jadbabaie “Optimal control of spatially distributed systems”
Spatially Decaying (SD) Operators

Linear operator $A$ is spatially-decaying (SD) if $\tilde{A}(\alpha)$ is bounded on all $\ell_p$ spaces where

$$[\tilde{A}(\alpha)]_{ki} := [A]_{ki} \chi_{\alpha}(\text{dis}(k, i))$$

for all $0 \leq \alpha < \tau$. 
Linear operator $\tilde{A}(\alpha)$ is spatially-decaying (SD) if $\tilde{A}(\alpha)$ is bounded on all $\ell_p$ spaces where

$$[\tilde{A}(\alpha)]_{ki} := [A]_{ki} \chi_\alpha(\text{dis}(k, i))$$

for all $0 \leq \alpha < \tau$.

**Examples:**

- **Exponentially-decaying if:**
  $$\chi_\alpha(x) = e^{\alpha x}$$

- **Algebraically-decaying if:**
  $$\chi_\alpha(x) = (1 + \lambda x)^\alpha, \quad \lambda > 0$$

- **Product of coupling functions, e.g.:**
  $$\chi_\alpha(x) = e^{\alpha x}(1 + \lambda x)^\alpha$$
Properties of a coupling characteristic function:

- $\chi_\alpha(0) = 1$ for all $\alpha \geq 0$ and $\chi_0(x) = 1$ for all $x \geq 0$.

- Continuous and nondecreasing in $x$.

- $\chi_\alpha(x + y) \leq \chi_\alpha(x) \chi_\alpha(y)$ (submultiplicative)

- If $\alpha \leq \beta$ then $\chi_\alpha(x) \leq \chi_\beta(x)$ for all $x \geq 0$.

- $\chi_\alpha(x) \chi_\beta(x) = \chi_{\alpha + \beta}(x)$ for all $\alpha, \beta, x \geq 0$. 
Subspace of Spatially Decaying Operators

- Consider the following subspace of linear operators:

\[ S^\infty_\tau(\mathcal{C}) = \left\{ A : \sup_k \sum_i \|[A]_{ki}\| \chi_\alpha(\text{dis}(k, i)) < \infty \quad \text{for all} \quad 0 \leq \alpha < \tau \right\} \]

- An operator norm can be defined:

\[ \|\|A\|\| = \sup_{\alpha \in [0, \tau)} \sup_k \sum_i \|[A]_{ki}\| \chi_\alpha(\text{dis}(k, i)) \]

- Structure of this subspace:

\((S^\infty_\tau(\mathcal{C}), \|\|\cdot\||)\) forms a Banach Algebra,

\[ \|\|AB\|\| \leq \|\|A\|\| \|\|B\|\| \]

for all \(A, B \in S^\infty_\tau(\mathcal{C}).\)
Banach Algebra of Spatially Decaying Operators

$(\mathcal{S}^\infty_T, ||| \cdot |||)$ forms a Banach Algebra

- **Properties:** For all $A, B \in \mathcal{S}^\infty_T$, it follows
  - Closed under addition: $A + B \in \mathcal{S}^\infty_T$
  - Closed under multiplication: $AB \in \mathcal{S}^\infty_T$
  - Closed under inversion: $A^{-1} \in \mathcal{S}^\infty_T$
  - Closed under limit.
Spatially Decaying Systems

- A linear control system

\[
\frac{d}{dt} \psi(i, t) = (A \psi)(i, t) + (B u)(i, t) \\
y(i, t) = (C \psi)(i, t) + (D u)(i, t)
\]

is called spatially decaying (SD) if \( A, B, C, D \) are SD operators.

- The serial, parallel, and feedback interconnections of SD systems result in SD systems.
Theorem:

Assume that the corresponding LQR problem is optimizable and \((A, Q^{1/2})\) is exponentially detectable. If \(A, B, Q, R \in S^\infty_\tau(\mathcal{H})\), then

\[
K \in S^\infty_\tau(\mathcal{H})
\]
Locality Features of the Optimal Controller

- The state feedback is **SD**: \( K \in S_\tau^\infty (\mathcal{C}) \)

\[
\sup_k \sum_i \| [K]_{ki} \| \chi_\alpha (\text{dis}(k, i)) < \infty
\]

\[
\lim_{\text{dis}(k,i) \to \infty} \| [K]_{ki} \| \chi_\alpha (\text{dis}(k, i)) = 0
\]

\[\chi_\alpha(x) = (1 + \lambda x)^\alpha, \quad \lambda > 0\]

\[\chi_\alpha(x) = e^{\alpha x}\]
Extension to Finite Spatial Dimensions

- The SD condition holds trivially for finite-dimensional matrices:
  \[
  \sup_k \sum_i \|[A]_{ki}\| \chi_\alpha(\text{dis}(k, i)) < \infty
  \]

- Consider the following subspace of finite-dimensional matrices:
  \[
  S_T^N(C) = \{ A : \|[A]_{ki}\| \leq C \chi_\alpha(\text{dis}(k, i))^{-1} \text{ for some } C > 0, \ 0 < \alpha < \tau \}
  \]

- This subspace has the following properties:
  - Closed under addition
  - Closed under multiplication
  - Closed under inversion

- Our previous results directly extend to finite dimensional case.
Definition: A vector space of bounded linear operators $S$ is called an operator algebra if it is closed in the norm topology of operators, $I \in S$, and for every $A, B \in S$,

$$A + B \in S$$
$$AB \in S.$$

Examples include:

- Infinite-Dimensional Spatially Decaying Operators
- Finite-Dimensional Spatially Decaying Operators
- Lower and Upper Triangular Matrices
- Circulant Matrices
**Definition:** A vector space of bounded linear operators $S$ is called an operator algebra if it is closed in the norm topology of operators, $I \in S$, and for every $A, B \in S$,

- $A + B \in S$
- $AB \in S$.

Examples include:

- Infinite-Dimensional Spatially Decaying Operators
- Finite-Dimensional Spatially Decaying Operators
- Lower and Upper Triangular Matrices
- Circulant Matrices
Theorem: Suppose that $S$ is an operator algebra and map $F : \mathbb{R} \times \mathcal{B}(\ell_p) \to \mathcal{B}(\ell_p)$ is continuous on $\mathbb{R}$. Assume that under appropriate conditions the following initial value problem

$$\frac{d}{dt}X(t) = F(t, X(t)) , \quad X(0) = X_0, \quad (1)$$

has a unique solution. If

$S$ is $F$-invariant,

$X_0 \in S$,

then $X(t) \in S$ for all $t \geq 0$ where $X(t)$ is the unique solution of problem (1).
Theorem: Suppose that $S$ is an operator algebra and map $F : \mathbb{R} \times \mathcal{B}(\ell_p) \to \mathcal{B}(\ell_p)$ is continuous on $\mathbb{R}$. Assume that under appropriate conditions the following initial value problem
\[ \frac{d}{dt} X(t) = F(t, X(t)) \ , 
X(0) = X_0, \] (1)
has a unique solution. If

- $S$ is $F$-invariant,
- $X_0 \in S$,

then $X(t) \in S$ for all $t \geq 0$ where $X(t)$ is the unique solution of problem (1).
Simulations

Systems marked by ‘*’:

\[ A_{kk} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad B_{kk} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

Systems marked by ‘o’:

\[ A_{kk} = \begin{bmatrix} -2 & 1 \\ 1 & -3 \end{bmatrix}, \quad B_{kk} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

Coupling matrix:

\[ A_{ki} = \frac{1}{\chi_\alpha(\text{dis}(k, i))} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

- Coupled systems:
  \[ \dot{x}_k = A_{kk} x_k + B_{kk} u_k + \sum_{i=1}^{N} A_{ki} x_i, \quad N = 200 \]

- In quadratic cost functional, the weighting matrices are defined as:
  \[ Q_{ij} = \begin{cases} -1 & \text{if } i \sim j \\ d_{ii} & \text{if } i = j \end{cases} \]
  \[ R = I \]
Systems marked by ‘*’: 
\[ A_{kk} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad B_{kk} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

Systems marked by ‘o’: 
\[ A_{kk} = \begin{bmatrix} -2 & 1 \\ 1 & -3 \end{bmatrix}, \quad B_{kk} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

Coupling matrix:
\[ A_{ki} = \frac{1}{\chi_\alpha(\text{dis}(k, i))} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

- Coupled systems: 
  \[ \dot{x}_k = A_{kk} x_k + B_{kk} u_k + \sum_{i=1}^{N} A_{ki} x_i, \quad N = 200 \]

- In quadratic cost functional, the weighting matrices are defined as:
  \[ Q_{ij} = \begin{cases} -1 & \text{if } i \sim j \\ d_{ii} & \text{if } i = j \end{cases}, \quad R = I \]
The coupling function is $\chi_\alpha(x) = e^{\alpha x}$ where $\alpha = 0.1823$.

Optimal state-feedback: 
$$u_k = K_{kk} \ x_k + \sum_{i \neq k} K_{ki} \ x_i$$
The coupling function is $\chi_\alpha(x) = (1 + 0.1x)^\alpha$ where $\alpha = 4$

Optimal state-feedback: $u_k = K_{kk} x_k + \sum_{i\neq k} K_{ki} x_i$
Spatial truncation of the optimal controller:

\[
[K_T]_{ki} = \begin{cases} 
[K]_{ki} & \text{if } \text{dis}(k, i) \leq T \\
0 & \text{if } \text{dis}(k, i) > T.
\end{cases}
\]

- Stabilizing truncation length:
  - Exp. decaying: \( T_s = 7.9785 \)
  - Algeb. decaying: \( T_s = 2.9603 \)
  - Nearest Neighbor: \( T_s = 15.0934 \)

Performance criteria:

\[
\frac{\text{Trace}(P_T) - \text{Trace}(P)}{\text{Trace}(P)} \times 100
\]

where

\[
(A+BK_T)^*P_T + P_T(A+BK_T) + Q + K_T^*RK_T = 0
\]
Distributed Quadratic Programming

\[
\min_x \frac{1}{2} \langle x, Qx \rangle + \langle c, x \rangle \\
\text{subject to: } Ax \leq b
\]

- Assumptions: \( Q > 0, \quad Q, A \in S_T^N(\mathbb{C}), \quad N \in \mathbb{N} \cup \{\infty\} \).
- Think of \( b \) and \( c \) as the parameters of the problem.
- Multi-parametric quadratic programming (MPQP): finds the optimal solution as a function of the parameters.
- This problem is motivated by the receding horizon control of spatially distributed dynamical systems.
Locality Feature of Quadratic Programming

Theorem:

Assume that $Q, A \in S^N_\tau(\mathbb{C})$, $N \in \mathbb{N} \cup \{\infty\}$. Let $\mathbb{B}$ be the set of parameters $b$ for which some combination of constraints are active at the optimal solution. Then the optimal solution is:

- An affine map of $b$ over $\mathbb{B}$,

$$x_k^* = \sum_i [K]_{ki} b_i + \sum_i [K_0]_{ki} c_i$$

- Spatially distributed,

$$K, K_0 \in S^N_\tau(\mathbb{C}) \text{ for some } 0 < \hat{\tau} < \tau$$
Locality Feature of Quadratic Programming

The state feedback law is localized:

\[ K, K_0 \in S^N_{\hat{\tau}}(\mathcal{C}) \text{ for some } 0 < \hat{\tau} < \tau \]

\[ \|K_{ki}\| \leq \kappa \chi_{\alpha}(\text{dis}(k, i))^{-1}, \quad \|[K_0]_{ki}\| \leq \kappa_0 \chi_{\alpha}(\text{dis}(k, i))^{-1} \]

\[ \chi_{\alpha}(x) = (1 + \lambda x)^{\alpha}, \quad \lambda > 0 \]

\[ \chi_{\alpha}(x) = e^{\alpha x} \]
Simulations

- Consider the QP problem:

\[
\min_x \frac{1}{2} \langle x, Qx \rangle + \langle c, x \rangle
\]
subject to: \( Ax \leq b \)

where

\[
Q_{ki} = \chi_\alpha(\text{dis}(k, i))^{-1}
\]

\[
A_{ki} = (-1)^{|k-i|} \chi_\alpha(\text{dis}(k, i))^{-1}
\]

and

\[
b, c = [1, 1, \ldots, 1]^T
\]

- Optimal solution can be represented as:

\[
x_k^* = \sum_{i=1}^{50} [K]_{ki} b_i + \sum_{i=1}^{50} [K_0]_{ki} c_i
\]
Exponentially Decaying Couplings

- The coupling function is $\chi_\alpha(s) = e^{\alpha s}$ where $\alpha = 1$

- For this case, the optimal solution is:
  $$x_k^* = \sum_{i=1}^{50} [K_0]_{ki} c_i$$
The coupling function is $\chi_\alpha(s) = (1 + s)^\alpha$ where $\alpha = 2$.

For this case, the optimal solution is:

$$x^*_k = \sum_{i=1}^{50} [K_0]_{ki} c_i$$
What’s next?

- Incorporating design: How to “patch” local controllers to achieve a global one?
- Appropriate blending of local controllers will result in a good global solution, by spatial gain-scheduling.

- How does the performance deteriorate with the support of the controller? Where is a good place to clip?
- How do take network topology change into account?