Nonlinear Consensus Model and Lohe Hierarchy (From Bacteria to Tensors)

Seung Yeal Ha

Department of Mathematical Sciences
Seoul National University

Nov. 26th, 2019

Outline

The first Story

Nonlinear Consensus Model

The second story

Aggregation of tensors

The third story

Consensus-based optimization algorithm

Three questions to be addressed

- (Q1): Is there any "universal design principle" for collective dynamics?
- (Q3): Can we design an aggregation model on the space of tensors?
- (Q3): Can we use aggregation models for optimization?

Outline

The first Story

Nonlinear Consensus Model

The second story

Aggregation of tensors

The third story

Consensus-based optimization algorithm

The first story

Is there any possible universal design principle for collective dynamics?

Collective behaviors of biological systems

- Aggregation, flocking and synchronization



Aggregation of bacteria

The Keller-Segel model :Patlak (1953), E. Keller and L. Segel (1970s)

$$\partial_t \rho + \nabla \cdot (\rho \nabla c) = \sigma \Delta \rho, \quad -\Delta c = \rho,$$

$$\begin{aligned} \rho = \rho(t,x) &: \text{ local mass density of bacteria,} \\ c = c(t,x) &: \text{ density of chemotactic substance.} \end{aligned}$$

Paricle Keller-Segel model:

$$dx_i(t) = -\frac{\kappa}{N} \sum_{j \neq i}^{N} \nabla \phi(x_j(t) - x_i(t)) dt + \sqrt{2\sigma} dB_i(t),$$

$$\begin{aligned} x_i(t) &: \text{ Position process of the } i\text{-th bacteria at time } t, \\ F = -\nabla \phi &: \text{ Couloumb's force, e.g., } \phi(x) = \frac{1}{|x|}, \quad \text{for } d = 3. \end{aligned}$$

In the absence of stochastic noise $\sigma = 0$, particle Keller-Segel model becomes

$$\dot{x}_i = \frac{\kappa}{N} \sum_{j \neq i}^{N} \frac{x_j - x_i}{|x_j - x_i|^3}.$$

cf. N-body system under gravitational force in $\mathbb{R}^3$:

$$\ddot{x}_i = \frac{\kappa}{N} \sum_{j \neq i}^{N} \frac{x_j - x_i}{|x_j - x_i|^3}.$$

Flocking of Cucker-Smale particles

- Dynamics observables:

$$x_i: \text{ position}, \qquad v_i: \text{ velocity}.$$

The Cucker-Smale model (2007) IEEE Trans. Automat. Control (2007):

$$\frac{dx_i}{dt} = v_i, \qquad \frac{dv_i}{dt} = \frac{\kappa}{N}\sum_{j=1}^{N}\psi(|x_j - x_i|)(v_j - v_i).$$

where ψ is a communication rate (modeling issue), e.g.,

$$\psi(|x_i - x_j|) = \frac{1}{(1 + |x_i - x_j|^2)^{\beta}}, \qquad \beta \geq 0.$$

Synchronization of Kuramoto oscillators

- Dynamic observables:

$$\theta_i : \text{ phase}, \qquad \dot{\theta}_i : \text{ frequency}.$$

The Kuramoto model (1975):

$$\frac{d\theta_i}{dt} = \nu_i + \frac{\kappa}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \quad i = 1, \cdots, N.$$

Corresponding PDE models

- The Keller-Segel model

$$\partial_t \rho + \nabla \cdot (\rho \nabla c) = \sigma \Delta \rho, \quad -\Delta c = \rho,$$

- The hydrodynamic Cucker-Smale model

$$\begin{aligned} &\partial_t \rho + \nabla_x \cdot (\rho u) = 0, \\ &\partial_t (\rho u) + \nabla_x \cdot (\rho u \otimes u) \\ &\qquad = -\kappa \int_{R^d} \psi(|x-y|)(u(y)-u(x))\rho(x)\rho(y)dy. \end{aligned}$$

- The kinetic Kuramoto model

$$\begin{aligned} &\partial_t F + \partial_\theta (\omega[F]F) = 0, \\ &\omega[F](\theta,\nu,t) := \nu - \kappa \int_0^{2\pi} \int_R \sin(\theta_* - \theta) F(\theta_*, \nu_*, t) d\nu_* d\theta. \end{aligned}$$

At PDE level, PDE models look different.

Particle models

- The Keller-Segel model in $\mathbb{R}^3$

$$\dot{x}_i = \frac{\kappa}{N} \sum_{k \neq i} \frac{x_k - x_i}{|x_k - x_i|^3}.$$

- The Cucker-Smale model

$$\dot{x}_i = v_i, \quad \dot{v}_i = \frac{\kappa}{N} \sum_{k=1}^{N} \psi_{cs}(x_k - x_i)(v_k - v_i).$$

- The Kuramoto model

$$\dot{\theta}_i = \nu_i + \frac{\kappa}{N} \sum_{k=1}^{N} \sin(\theta_k - \theta_i).$$

First-order formulation of the C-S model on a line

- The C-S model in $1D$: H-Kim-Park-Zhang '19 ARMA

$$\dot{x}_i = v_i, \qquad \dot{v}_i = \frac{\kappa}{N}\sum_{k=1}^{N}\psi(x_k - x_i)(v_k - v_i).$$

Idea

$$\psi(x_k - x_i)(v_k - v_i) = \frac{d}{dt}\int_0^{x_k - x_i}\psi(s)ds =: \frac{d}{dt}\Psi_{cs}(x_k - x_i).$$

Then, C-S flocking becomes a first-order consensus model:

$$\dot{x}_i = \nu_i(X^0, V^0) + \frac{\kappa}{N}\sum_{k=1}^{N}\Psi_{cs}(x_k - x_i),$$
$$\nu_i(X^0, V^0) := v_i^0 - \frac{\kappa}{N}\sum_{j=1}^{N}\psi(x_k^0 - x_i^0).$$

cf. KM $\Longrightarrow$ CS: H-Lattanzio-Rubino-Slemrod '11.

Particle Pictures

q_i : generalized position of the i-th particle.

- The deterministic Keller-Segel model in 3d

$$\dot{q}_i = \nu_i + \frac{\kappa}{N}\sum_{k=1}^{N}\Psi_a(q_k - q_i), \quad \Psi_a(q) = \frac{q}{|q|^3}.$$

- The Cucker-Smale model in 1d

$$\dot{q}_i = \nu_i(q^0, p^0) + \frac{\kappa}{N}\sum_{k=1}^{N}\Psi_{cs}(q_j - q_i), \quad \Psi_{cs}(q) = \int^{q}\psi_{cs}(y)dy.$$

- The Kuramoto model

$$\dot{q}_i = \nu_i + \frac{\kappa}{N}\sum_{k=1}^{N}\Psi_k(q_k - q_i), \quad \Psi_k(q) = \sin q.$$

Summary

- Master Particle model for collective dynamics

$$\dot{q}_i = \nu_i + \frac{\kappa}{N} \sum_{k=1}^{N} \Psi(q_k - q_i), \quad i = 1, \cdots N, \quad q_i \in \mathcal{M}.$$

In other words, there exists a kind of triality relation:

Keller-Segel aggregation

$\Longleftrightarrow$ *1D CS flocking*

$\Longleftrightarrow$ Kuramoto synchronization.

Outline

The first Story

Nonlinear Consensus Model

The second story

Aggregation of tensors

The third story

Consensus-based optimization algorithm

The second story

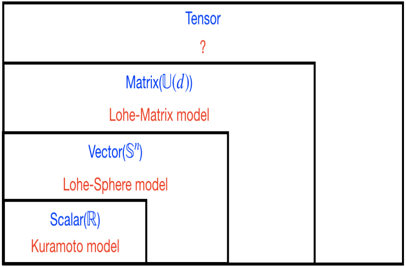
Can we design an aggregation model on the space of tensors ?

There are several first-order aggregation models for the collections of real numbers, real vectors in $\mathbb{R}^d$ and unitary matrices $\mathbb{U}(d)$.

- Are there aggregation models for non square matrices, for example $\mathbb{R}^{n \times m}$ with $n \neq m$?
- Can we design a first-order aggregation model on a space of tensors ?

Lohe Hiearchy

An aggregation model for an ensemble of tensors



This is a joint project with Hansol Park (Ph.D. student in SNU).

What is a tensor?

In high school or linear algebra class in college, you might learn "matrix" as a rectangular array of (real or complex) numbers. Note that 1×1 matrix is simply a number and $n \times 1$ matrix can be interpreted as a vector in $\mathbb{R}^n$ or $\mathbb{C}^n$ depending on your scalar field. Thus matrix includes numbers and vectors.

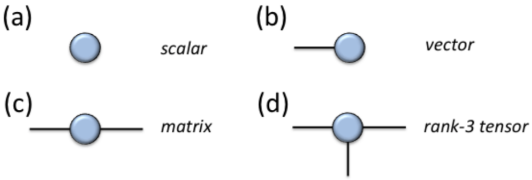
Then you also might have a chance to think of the following question what is a high-dimensional generalization of a matrix?

◇ Mathematical Definition: Let V and V^* be a vector space and dual spacer over a scalar field F. Then, a tensor is a scalar valued multi-linear map with variables in both V and V^*.

◇ Physical Definition: Tensor is a multi-dimensional array of scalar values, and the rank of a tensor is the number of indices.

◇ Remark: We denote a set of all rank-m tensors with size $d_1 \times \cdots \times d_m$ by $\mathcal{T}_m(\mathbb{C}; d_1 \times \cdots d_m)$. Then, the set $\mathcal{T}_m(\mathbb{C}; d_1 \times \cdots d_m)$ is a vector space over $\mathbb{C}$.

Pictorial representation for tensors



Scalar: rank-0 tensor, Vector: rank-1tensor, Matrix: rank-2 tensor

Thus, a tensor is a multi-dimensional generalization of a scalar value, vector and matrix

Goal of the second story

In this talk, I would like to propose a new aggregation model on $\mathcal{T}_m(\mathbb{C}; d_1 \times \cdots d_m)$

The Kuramoto model $\Longrightarrow$ The Lohe sphere model $\Longrightarrow$ The Lohe matrix model $\Longrightarrow$ The Lohe tensor model

Existing aggregation models for low-rank tensors

- The Lohe matrix model for complex-valued rank-2 tensors:

U_i : $d \times d$ unitary matrix, H_i : $d \times d$ Hermitian matrix.

$$\mathrm{i}\dot{U}_i U_i^\dagger = H_i + \frac{\mathrm{i}\kappa}{2N} \sum_{k=1}^{N} \left(U_i U_j^\dagger - U_j U_i^\dagger \right).$$

- The Lohe sphere model for real-valued rank-1 tensors:

x_i : a real vector in $\mathbb{R}^d$, Ω_i : $d \times d$ skew-symmetric matrix.

$$\dot{x}_i = \Omega_i x_i + \frac{\kappa}{N} \sum_{k=1}^{N} (\langle x_i, x_i \rangle x_k - \langle x_k, x_i \rangle x_i),$$

- The Kuramoto model for real-valued rank-0 tensors:

θ_i : real number, ν_i : real number.

$$\dot{\theta}_i = \nu_i + \frac{\kappa}{N} \sum_{k=1}^{N} \sin(\theta_k - \theta_i).$$

Hierarchy relations

- Lohe matrix model $\Longrightarrow$ Lohe sphere model: For $d = 2$, we set

$$U_i := \mathrm{i}\sum_{k=1}^{3} x_i^k \sigma_k + x_i^4 I_2 = \begin{pmatrix} x_i^4 + \mathrm{i}x_i^1 & x_i^2 + \mathrm{i}x_i^3 \\ -x_i^2 + \mathrm{i}x_i^3 & x_i^4 - \mathrm{i}x_i^1 \end{pmatrix},$$

$$H_i = \sum_{k=1}^{3} \omega_i^k \sigma_k + \nu_i I_2,$$

where

$$I_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$$||x_i||^2 \dot{x}_i = \Omega_i x_i + \frac{K}{N} \sum_{k=1}^{N} (||x_i||^2 x_k - \langle x_i, x_k \rangle x_i),$$

where Ω_i is a real 4×4 antisymmetric matrix:

$$\Omega_i := \begin{pmatrix} 0 & -\omega_i^3 & \omega_i^2 & -\omega_i^1 \\ \omega_i^3 & 0 & -\omega_i^1 & -\omega_i^2 \\ -\omega_i^2 & \omega_i^1 & 0 & -\omega_i^3 \\ \omega_i^1 & \omega_i^2 & \omega_i^3 & 0 \end{pmatrix}.$$

• Lohe sphere model $\Longrightarrow$ Kuramoto model:

We set

$$d = 2, \quad x_i = \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix}, \Omega_i = \begin{bmatrix} 0 & -\nu_i \\ \nu_i & 0 \end{bmatrix},$$

Then, x^1 and x^2 components of

$$\dot{x}_i = \Omega_i x_i + \frac{\kappa}{N} \sum_{k=1}^{N} (\langle x_i, x_i \rangle x_k - \langle x_k, x_i \rangle x_i),$$

reduce to

$$\dot{\theta}_i = \nu_i + \frac{\kappa}{N} \sum_{k=1}^{N} \sin(\theta_k - \theta_i).$$

Problem statement

On the space of rank-m tensors $\mathcal{T}_m(\mathbb{C})$, we would like to design a new aggregation model with the following two minimal properties:

- Emergent collective behavior under suitable conditions
- Reductions to Lohe type low-rank aggregation models for special cases

Lessons from existing models

For a given tensor $T \in \mathcal{T}_m(\mathbb{C}; d_1 \times \cdots d_m)$ and $\alpha \in \Pi_{i=1}^m \{1, \cdots, d_i\}$, we denote $[T]_\alpha$ to be the α-th component of T.

- The Lohe sphere model in vector form

$$\dot{x}_i = \Omega_i x_i + \frac{\kappa}{N}\sum_{k=1}^{N}(\langle x_i, x_i\rangle x_k - \langle x_i, x_k\rangle x_i).$$

◇ The Lohe sphere model in component form

$$\begin{aligned}\frac{d}{dt}[x_i]_\alpha &= [\Omega_i x_i]_\alpha + \kappa([x_i]_\beta[x_i]_\beta[x_c]_\alpha - [x_i]_\beta[x_c]_\beta[x_i]_\alpha)\\ &= [\Omega_i]_{\alpha\beta}[x_i]_\beta + \kappa([x_i]_\beta[x_i]_\beta[x_c]_\alpha - [x_i]_\beta[x_c]_\beta[x_i]_\alpha)\end{aligned}$$

where $x_c = \frac{1}{N}\sum_{k=1}^N x_k$.

- The Lohe matrix model in matrix form

$$\mathrm{i}\dot{U}_j U_j^* = H_j + \frac{\mathrm{i}\kappa}{2N}\sum_{k=1}^{N}(U_k U_j^* - U_j U_k^*).$$

or equivalently

$$\dot{U}_j = -\mathrm{i}H_j U_j + \frac{\kappa}{2}(U_c - U_j U_c^* U_j).$$

or

$$\dot{U}_j = -\mathrm{i}H_j U_j + \frac{\kappa}{2}(U_j U_j^* U_c - U_j U_c^* U_j).$$

◇ The Lohe matrix model in component form

$$\frac{d}{dt}[U_j]_{\alpha\beta} = [-\mathrm{i}H_i U_j]_{\alpha\beta} + \frac{\kappa}{2}([U_j]_{\alpha\gamma}[\bar{U}_j]_{\delta\gamma}[U_c]_{\delta\beta} - [U_j]_{\alpha\gamma}[\bar{U}_c]_{\delta\gamma}[U_j]_{\delta\beta})$$

Next, we interpret the free flow term $[-\mathrm{i}H_j U_j]_{\alpha\beta}$ as a contraction of rank-4 tensor A_j and rank-2 tensor U_j. For this, we define rank-4 tensor A_j as follows:

$$[A_j]_{\alpha\beta\gamma\delta} := [-\mathrm{i}H_j]_{\alpha\gamma}\delta_{\beta\delta} \quad \text{and} \quad \delta_{\beta\delta} := \begin{cases} 1, & \beta = \delta, \\ 0, & \beta \neq \delta. \end{cases}$$

- **Lemma**: Let A_j be a rank-4 tensor defined in previous slide. Then, the following relations hold:

$$[\bar{A}_j]_{\gamma\delta\alpha\beta} = -[A_j]_{\alpha\beta\gamma\delta} \quad \text{and} \quad [A_j]_{\alpha\beta\gamma\delta}[U_j]_{\gamma\delta} = [-\mathrm{i}H_jU_j]_{\alpha\beta}.$$

Proof: For the first identity, we use defining relation for a rank-4 tensor A_j, $H_j^* = H_j$ and $\delta_{\delta\beta} = \delta_{\beta\delta}$ to get

$$[\bar{A}_j]_{\gamma\delta\alpha\beta} = [\mathrm{i}\bar{H}_j]_{\gamma\alpha}\delta_{\delta\beta} = [\mathrm{i}H_j]_{\alpha\gamma}\delta_{\delta\beta} = -[-\mathrm{i}H_j]_{\alpha\gamma}\delta_{\beta\delta} = -[A_j]_{\alpha\beta\gamma\delta}.$$

For the second identity, one has

$$[A_j]_{\alpha\beta\gamma\delta}[U_j]_{\gamma\delta} = [-\mathrm{i}H_j]_{\alpha\gamma}\delta_{\beta\delta}[U_j]_{\gamma\delta} = [-\mathrm{i}H_j]_{\alpha\gamma}[U_j]_{\gamma\beta} = [-\mathrm{i}H_jU_j]_{\alpha\beta}.$$

Finally, one has

$$\frac{d}{dt}[U_j]_{\alpha\beta} = [A_j]_{\alpha\beta\gamma\delta}[U_j]_{\gamma\delta}+\frac{\kappa}{2}\left[[U_c]_{\alpha\gamma}[U_j^*]_{\gamma\delta}[U_j]_{\delta\beta} - [U_j]_{\alpha\gamma}[U_c^*]_{\gamma\delta}[U_j]_{\delta\beta}\right].$$

Lesson from previous models

Consider an ensemble $\{T_j\}_{j=1}^N$ of rank-m tensors over complex field $\mathbb{C}$, and for notational simplicity, we set

$$\alpha_* = (\alpha_1, \cdots, \alpha_m), \qquad \beta_* = (\beta_1, \cdots, \beta_m).$$

Then, we begin with following structure:

$$\frac{d}{dt}[T_j]_{\alpha_*} = \text{free flow} + \text{cubic interactions}.$$

- (Modeling of free flow)

Contraction of rank-$2m$ tensor A_j and rank-m tensor T_j:

$$\text{free flow part} = [A_j]_{\alpha_* \beta_*}[T_j]_{\beta_*}.$$

- (Modeling of cubic interactions): for a dummy variable β,

$$[T_c]_{i_1}[\bar{T}_j]_\beta[T_j]_{i_2} - [T_j]_{i_1}[\bar{T}_c]_\beta[T_j]_{i_2}.$$

- **Definition**:

We define the inner product of size $N_1 \times N_2 \times \cdots \times N_m$ as follows.

$$\langle T_i, T_j \rangle_F := [\bar{T}_i]_{\alpha_*}[T_j]_{\alpha_*}, \quad i, j = 1, \cdots, N.$$

Generalized Lohe tensor model

$$\frac{d}{dt}[T_i]_{\alpha_{10}\alpha_{20}\cdots\alpha_{m0}} = [A_i]_{\alpha_{10}\alpha_{20}\cdots\alpha_{m0}\beta_1\beta_2\cdots\beta_m}[T_i]_{\beta_1\beta_2\cdots\beta_m}$$
$$+ \sum_{(i_1,i_2,\cdots,i_m)\in\{0,1\}^m} \kappa_{i_1\cdots i_m}\Big([T_c]_{\alpha_{1i_1}\cdots\alpha_{mi_m}}[\bar{T}_i]_{\alpha_{11}\alpha_{21}\cdots\alpha_{m1}}[T_i]_{\alpha_{1(1-i_1)}\cdots\alpha_{m(1-i_m)}}$$
$$- [T_i]_{\alpha_{1i_1}\alpha_{2i_2}\cdots\alpha_{mi_m}}[\bar{T}_c]_{\alpha_{11}\alpha_{21}\cdots\alpha_{m1}}[T_i]_{\alpha_{1(1-i_1)}\alpha_{2(1-i_2)}\cdots\alpha_{m(1-i_m)}}\Big),$$

where A_i satisfies

$$[\bar{A}_i]_{\alpha_1\alpha_2\cdots\alpha_m\beta_1\beta_2\cdots\beta_m} = -[A_i]_{\beta_1\beta_2\cdots\beta_m\alpha_1\alpha_2\cdots\alpha_m}.$$

- $\alpha_{10}, \alpha_{20}, \cdots, \alpha_{m0}$ are fixed indices.
- $\alpha_{11}, \alpha_{21}, \cdots, \alpha_{m1}$ are dummy variables.
- A_i are generalization of skew-hermitian matrices.
- T_i have size $d_1 \times d_2 \times \cdots \times d_m$.(Rank m-tensor)
- A_i has size $d_1 \times d_2 \times \cdots \times d_m \times d_1 \times d_2 \times \cdots \times d_m$.(Rank $2m$-tensor)

Lohe tensor Model

For the handy notation, we define follows:

$$\alpha_{*0} = \alpha_{10}\alpha_{20}\cdots\alpha_{m0}, \quad \alpha_{*1} = \alpha_{11}\alpha_{21}\cdots\alpha_{m1},$$
$$\alpha_{*i_*} = \alpha_{1i_1}\alpha_{2i_2}\cdots\alpha_{mi_m}, \quad \alpha_{*(1-i_*)} = \alpha_{1(1-i_1)}\alpha_{2(1-i_2)}\cdots\alpha_{m(1-i_m)},$$
$$\beta_* = \beta_1\beta_2\cdots\beta_m, \quad i_* = i_1 i_2\cdots i_m.$$

If we use above handy notation, we can obtain

$$\frac{d}{dt}[T_i]_{\alpha_{*0}} = \underbrace{[A_i]_{\alpha_{*0}\beta_*}[T_i]_{\beta_*}}_{\text{Free Flow}} + \underbrace{\sum_{i_*\in\{0,1\}^m} \kappa_{i_*}\left([T_c]_{\alpha_{*i_*}}[\bar{T}_i]_{\alpha_{*1}}[T_i]_{\alpha_{*(1-i_*)}} - [T_i]_{\alpha_{*i_*}}[\bar{T}_c]_{\alpha_{*1}}[T_i]_{\alpha_{*(1-i_*)}}\right)}_{\text{Cubic coupling Terms}}$$

Reduction of the Lohe tensor model

- Ensemble of rank-1 tensors

$$\frac{d}{dt}[z_i]_{\alpha_{10}} = [\Omega_i]_{\alpha_{10}\beta_1}[z_i]_{\beta_1} + \kappa_0\Big([z_c]_{\alpha_{10}} \underbrace{[\bar{z}_i]_{\alpha_{11}}[z_i]_{\alpha_{11}}}_{\text{Contracted}} - [z_i]_{\alpha_{10}} \underbrace{[\bar{z}_c]_{\alpha_{11}}[z_i]_{\alpha_{11}}}_{\text{Contracted}}\Big)$$
$$+ \kappa_1\Big(\underbrace{[z_c]_{\alpha_{11}}[\bar{z}_i]_{\alpha_{11}}}_{\text{Contracted}}[z_i]_{\alpha_{10}} - \underbrace{[z_i]_{\alpha_{11}}[\bar{z}_c]_{\alpha_{11}}}_{\text{Contracted}}[z_i]_{\alpha_{10}}\Big).$$

After contractions, one has the complex analog of the Lohe sphere model:

$$\dot{z}_i = \underbrace{\Omega_i z_i}_{\text{Free Flow}} + \underbrace{\kappa_0(\langle z_i, z_i\rangle z_c - \langle z_c, z_i\rangle z_i)}_{\text{Lohe sphere coupling}} + \underbrace{\kappa_1(\langle z_i, z_c\rangle - \langle z_c, z_i\rangle) z_i}_{\text{new coupling}},$$

where inner product $\langle \cdot, \cdot \rangle$ defined as

$$\langle u, v\rangle := u^* v = [\bar{u}]_\alpha [v]_\alpha.$$

For real rank-1 tensors, the new coupling terms are zero, and we obtain the Lohe sphere model for $z_i = x_i$:

$$\dot{x}_i = \Omega_i x_i + \kappa_0(\langle x_i, x_i \rangle x_c - \langle x_c, x_i \rangle x_i).$$

- Ensemble of rank-2 tensors

$$
\begin{aligned}
[\dot{U}_i]_{\alpha_{10}\alpha_{20}} = {} & [A_i]_{\alpha_{10}\alpha_{20}\beta_1\beta_2}[U_i]_{\beta_1\beta_2} \\
& + \kappa_{00}([U_c]_{\alpha_{10}\alpha_{20}}[\bar{U}_i]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{11}\alpha_{21}} - [U_i]_{\alpha_{10}\alpha_{20}}[\bar{U}_c]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{11}\alpha_{21}}) \\
& + \kappa_{01}([U_c]_{\alpha_{10}\alpha_{21}}[\bar{U}_i]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{11}\alpha_{20}} - [U_i]_{\alpha_{10}\alpha_{21}}[\bar{U}_c]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{11}\alpha_{20}}) \\
& + \kappa_{10}([U_c]_{\alpha_{11}\alpha_{20}}[\bar{U}_i]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{10}\alpha_{21}} - [U_i]_{\alpha_{11}\alpha_{20}}[\bar{U}_c]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{10}\alpha_{21}}) \\
& + \kappa_{11}([U_c]_{\alpha_{11}\alpha_{21}}[\bar{U}_i]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{10}\alpha_{20}} - [U_i]_{\alpha_{11}\alpha_{21}}[\bar{U}_c]_{\alpha_{11}\alpha_{21}}[U_i]_{\alpha_{10}\alpha_{20}}).
\end{aligned}
$$

Where α_{11} and α_{21} are dummy variables.

After simplification, one has

$$\dot{U}_i = \underbrace{A_i U_i}_{\text{free flow}} + \underbrace{\kappa_{00}(\mathrm{tr}(U_i^* U_i) U_c - \mathrm{tr}(U_c^* U_i) U_i)}_{\text{Lohe sphere coupling}}$$
$$+ \underbrace{\kappa_{01}(U_c U_i^* U_i - U_i U_c^* U_i)}_{\text{Lohe matrix coupling}} + \underbrace{\kappa_{10}(U_i U_i^* U_c - U_i U_c^* U_i)}_{\text{Lohe matrix coupling}}$$
$$+ \underbrace{\kappa_{11}(\mathrm{tr}(U_i^* U_c) - \mathrm{tr}(U_c^* U_i)) U_i}_{\text{new coupling}}.$$

- Remark If we put $m = 2$, $\kappa_{00} = \kappa_{11} = 0$, $\kappa_{01} + \kappa_{10} = \kappa$ and $[A_i]_{\alpha\beta\gamma\epsilon} = [-\mathrm{i}H_i]_{\alpha\gamma}\delta_{\beta\epsilon}$ then we can obtain “Lohe Matrix Model".

Emergent aggregation estimates

Consider the Lohe tensor model.

$$\frac{d}{dt}[T_i]_{\alpha_{*0}} = \underbrace{[A_i]_{\alpha_{*0}\beta_*}[T_i]_{\beta_*}}_{\text{Free Flow}}$$
$$+ \underbrace{\sum_{i_* \in \{0,1\}^m} \left([T_c]_{\alpha_{*i_*}}[\bar{T}_i]_{\alpha_{*1}}[T_i]_{\alpha_{*(1-i_*)}} - [T_i]_{\alpha_{*i_*}}[\bar{T}_c]_{\alpha_{*1}}[T_i]_{\alpha_{*(1-i_*)}}\right)}_{\text{Coupling Term}}$$

We set

$$||T_i||_F := \sqrt{[\bar{T}_i]_{\alpha_*}[T_i]_{\alpha_*}}.$$

- **Theorem**: (Conservation law)

$$||T_i(t)||_F = ||T_i^{in}||_F, \quad t \geq 0.$$

Emergent aggregation dynamics

We set

$$\mathcal{D}(T) := \max_{i,j} ||T_i - T_j||_F, \quad \mathcal{D}(A) := \max_{i,j} ||A_i - A_j||_F, \quad \hat{\kappa}_0 := 2 \sum_{i_* \neq 0} \kappa_{i_*}.$$

- **Theorem**: H-Park '19

Suppose that the coupling strength and the initial data satisfy

$$A_j = 0, \quad \hat{\kappa}_0 < \frac{\kappa_0}{2||T_c^{in}||_F^2}, \quad ||T_j^{in}||_F = 1, \quad 0 < \mathcal{D}(T^{in}) < \frac{\kappa_0 - 2\hat{\kappa}_0||T_c^{in}||_F^2}{2\kappa_0}.$$

Then, there exist positive constants C_0 and C_1 depending on κ_{i_*} and T^{in} such that

$$C_0 e^{-\left(\kappa_0 + 2\hat{\kappa}_0||T_c^{in}||_F\right)t} \leq \mathcal{D}(T(t)) \leq C_1 e^{-\left(\kappa_0 - 2\hat{\kappa}_0||T_c^{in}||_F\right)t}, \quad t \geq 0.$$

Proof: By direct estimates, one has Gronwall differential inequality:

$$\left| \frac{d}{dt}\mathcal{D}(T) + \kappa_0 \mathcal{D}(T) \right| \leq 2\kappa_0 \mathcal{D}(T)^2 + \hat{\kappa}_0 ||T_c^{in}||_F \mathcal{D}(T), \quad \text{a.e. } t > 0.$$

Let η be the largest root of the quadratic equation:

$$2\kappa_0 x^2 + (\kappa_0 - 2\hat{\kappa}_0 ||T_c^{in}||_F^2)x = \mathcal{D}(A).$$

Then, the root η satisfies

$$0 < \eta < \frac{\kappa_0 - 2\hat{\kappa}_0 ||T_c^{in}||_F^2}{2\kappa_0}.$$

- **Theorem**: H-Park '19

Suppose that coupling strength, initial data and frequency matrices satisfy

$$\kappa_0 > 0, \quad 0 \le \mathcal{D}(T(0)) \le \eta \quad \text{and} \quad \mathcal{D}(A) < \frac{|\kappa_0 - 2\hat{\kappa}_0 ||T_c^{in}||_F^2|^2}{8\kappa_0},$$

Then practical synchronization emerges:

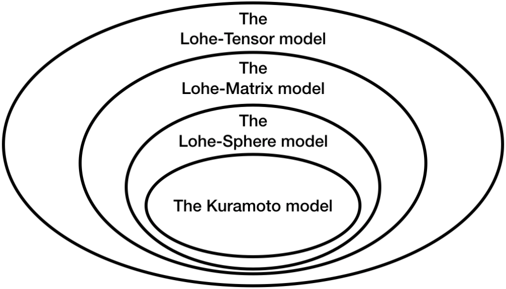
$$\lim_{D(A)/\kappa_0 \to 0+} \limsup_{t \to \infty} D(T(t)) = 0.$$

Proof: By direct estimates, one has Gronwall differential inequality:

$$\frac{d}{dt}\mathcal{D}(T) \le 2\kappa_0 \mathcal{D}(T)^2 - (\kappa_0 - 2\hat{\kappa}_0 ||T_c^{in}||_F^2)\mathcal{D}(T) + \mathcal{D}(A), \quad \text{a.e. } t > 0.$$

Summary

1. We have established Lohe Hierarchy:



2. As byproducts of our generalized approach, we have derived complex analogue for the Lohe sphere model.

Outline

The first Story

Nonlinear Consensus Model

The second story

Aggregation of tensors

The third story

Consensus-based optimization algorithm

The third story

Application of nonlinear consensus models to metaheuristic optimization algorithms

This is a joint work with Doheon Kim (KIAS) and Shi Jin (Shanghai Jiaotong Univ.)

A typical optimization problem

For a given objective function $L : \mathcal{S} \to \mathbb{R}$, we would like to find a global minimum $X^* \in \mathcal{S}$ such that

$$X^* \in \operatorname{argmin}_{X \in \mathcal{S}} L(X),$$

where we do not assume L is neither convex nor smooth, and $\beta > 0$

Consensus-based optimization(CBO) algorithm

- Introduced by a series of papers: Pinnau-Totzeck-Tse-Martin ('17), Carrillo-Choi-Totzeck-Tse ('18), Carrillo-Jin-Li-Zhu ('19)

$$
\begin{cases}
dX_t^i = -\lambda(X_t^i - \bar{X}_t^*)dt + \sigma \displaystyle\sum_{l=1}^{d} (x_t^{i,l} - \bar{x}_t^{*,l}) dW_t^l e_l, \\
\bar{X}_t^* = (x_t^{*,1}, \cdots, x_t^{*,d}) := \dfrac{\sum_{l=1}^{N} X_t^l e^{-\beta L(X_t^l)}}{\sum_{l=1}^{N} e^{-\beta L(X_t^l)}},
\end{cases}
$$

- (Emergence of a global consensus): For a.s. $\omega \in \Omega$, is there a global consensus state $X_\infty(\omega) \in^d$ such that

$$\lim_{t\to\infty} \|X_t^i(\omega) - X_t^j(\omega)\|_{\ell^2} = 0, \quad i, j = 1, \cdots, N,$$

?

- (Convergence of CBO algorithm): If the constant consensus state X_∞ exists, then under what condition how the consensus state X_∞ is close to the global minimum X^* of L?

CBO as a nonlinear consensus model

$$dX_t^i = \lambda \sum_{k=1}^{N} \psi_t^k (X_t^k - X_t^i) dt + \sigma \sum_{k=1}^{N} \sum_{l=1}^{d} \psi_t^k (x_t^{k,l} - x_t^{i,l}) dW_t^l e_l, \quad t > 0,$$

where $\psi_t^k := \psi^k(\mathcal{X}, t)$ is the communication weight function:

$$\psi_t^k := \frac{e^{-\beta L(X_t^k)}}{\sum_{l=1}^{N} e^{-\beta L(X_t^l)}}, \quad t \geq 0, \quad k = 1, \cdots, N.$$

Note that

$$(i)\ \psi_t^k \geq 0, \quad 1 \leq k \leq N, \quad \sum_{k=1}^{N} \psi_t^k = 1 \quad \text{for all } t \geq 0,$$

(ii) Dependence only on the state of source sample point,

Emergence of a global consensus

Note that $x_t^{ij,l} := x_t^{i,l} - x_t^{j,l}$ satisfies

$$\begin{cases} dx_t^{ij,l} = -\lambda x_t^{ij,l} dt - \sigma x_t^{ij,l} dW_t^l, \quad t > 0, \\ x_t^{ij,l}\Big|_{t=0} = x_0^i - x_0^j. \end{cases}$$

By Ito's formula, one has

$$x_t^{ij,l} = x_0^{ij,l} \exp\left[-\left(\lambda + \frac{\sigma^2}{2}\right)t + \sigma W_t^l \right], \quad t \geq 0.$$

- **Theorem**: H-Jin-Kim '19

Let $\{X_t^i\}$ be a solution process. Then,

$$\lim_{t\to\infty} |x_t^{i,k} - x_t^{j,k}| = 0, \quad \text{a.s.}$$
$$\lim_{t\to\infty} \mathbb{P}\Big(|x_t^{i,k} - x_t^{j,k}|^2 > \varepsilon \Big) = 0, \quad \text{for any } \varepsilon > 0.$$

cf. No restrictions on initial data

Convergence Analysis of CBO

Recall that X_t^i satisfies

$$dX_t^i = -\lambda(X_t^i - \bar{X}_t^*)dt + \sigma \sum_{l=1}^{d} (x_t^{i,l} - \bar{x}_t^{*,l})dW_t^l e_l,$$

and we introduce an ensemble average:

$$\bar{X}_t := \frac{1}{N} \sum_{i=1}^{N} X_t^i = (\bar{x}_t^1, \cdots, \bar{x}_t^d).$$

- **Lemma**: Let $\{X_t^i\}_{1 \le i \le N}$ be a solution.

$$(i)\ |X_t^i - \bar{X}_t|^2 = \sum_{l=1}^{d} (x_0^{i,l} - \bar{x}_0^l)^2 \exp\left[-\left(2\lambda + \sigma^2\right)t + 2\sigma W_t^l\right].$$

$$(ii)\ |\bar{X}_t - \bar{X}_t^*|^2 \le \max_{1 \le i \le N} |X_t^i - \bar{X}_t|^2.$$

$$(iii)\ \frac{1}{N} \sum_{i=1}^{N} |X_t^i - \bar{X}_t^*|^2 \le 2 \sum_{l=1}^{d} \left(\max_{1 \le i \le N} (x_0^{i,l} - \bar{x}_0^l)^2 \right) \exp\left[-\left(2\lambda + \sigma^2\right)t + 2\sigma W_t^l\right].$$

- **Lemma**: Let $\{X_t^i\}_{1\le i\le N}$ be a solution.

$$(i)\ \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}|X_t^i-\bar{X}_t^*|^2\le 2e^{-(2\lambda-\sigma^2)t}\sum_{l=1}^{d}\mathbb{E}\Big[\max_{1\le i\le N}(x_0^{i,l}-\bar{x}_0^l)^2\Big].$$

(ii) If $2\lambda>\sigma^2$, then there exists a random vector X_∞ such that

$$\lim_{t\to\infty}X_t^i=X_\infty\ \text{a.s.},\ 1\le i\le N.$$

Proof. For $i=1,\cdots,N$ and $l=1,\cdots,d$,

$$x_t^{i,l}=x_0^{i,l}-\lambda\int_0^t(x_s^{i,l}-\bar{x}_s^{*,l})ds+\sigma\int_0^t(x_s^{i,l}-\bar{x}_s^{*,l})dW_s^l=:x_0^{i,l}-\lambda\mathcal{I}_{11}+\sigma\mathcal{I}_{12}.$$

Thus, it suffices to check that convergence of $\mathcal{I}_{11}$ and $\mathcal{I}_{12}$.

• Case A (Almost sure convergence of $\mathcal{I}_{11}$): We first show that there exist positive random functions $C_i = C_i(\omega),\ i = 1, 2$ such that

$$|x_t^{i,l} - \bar{x}_t^{*,l}| \leq C_1 e^{-C_2 t}, \quad \text{a.s. } \omega \in \Omega,$$

where C_1 and C_2 are positive constants. We set

$$\mathcal{J}_{11} := \mathcal{I}_{11} - \int_0^t C_1 e^{-C_2 s} ds = \int_0^t \underbrace{\left(x_s^{i,l} - \bar{x}_s^{*,l} - C_1 e^{-C_2 s}\right)}_{\leq 0} ds.$$

Since the integrand is nonpositive a.s., $\mathcal{J}_{11}$ is non-increasing in t a.s. Then, we show

$$\mathcal{J}_{11} \geq -\frac{2C_1}{C_2}.$$

• Case B (Almost sure convergence of $\mathcal{I}_{12}$): We show that the term $\mathcal{I}_{12}$ is martingale and uniformly bounded in L^2. By direct calculation, one has

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \left(x_s^{i,l} - \bar{x}_s^{*,l}\right) dW_s^l\right]^2 \\
&= \mathbb{E}\int_0^t (x_s^{i,l} - \bar{x}_s^{*,l})^2 ds \le \int_0^t \sum_{i=1}^N \mathbb{E}|X_s^i - \bar{X}_s^*|^2 ds \\
&\le 2N\left(\int_0^t e^{-(2\lambda-\sigma^2)s} ds\right) \sum_{l=1}^d \left(\mathbb{E}\max_{1\le i\le N}(x_0^{i,l} - \bar{x}_0^l)^2\right) \\
&\le \frac{2N}{2\lambda - \sigma^2} \sum_{l=1}^d \left(\mathbb{E}\max_{1\le i\le N}(x_0^{i,l} - \bar{x}_0^l)^2\right).
\end{aligned}$$

Let $L = L(x)$ be a C_b^2-objective function satisfying the following relations:

$$L_m := \inf_{x \in R^d} L(x) > 0 \quad \text{and} \quad C_L := \max\left\{ \sup_{x\in R^d} \|\nabla^2 L(x)\|_2, \max_{1\le l\le d} \sup_{x\in R^d} |\partial_l^2 L(x)| \right\}$$

where $\|\cdot\|_2$ denotes the spectral norm.

- **Theorem**: H-Jin-Kim '19

Suppose that λ, σ and $\{X_0^i\}$ satisfy

$$2\lambda > \sigma^2, \quad X_0^i : i,i.d,, \quad X_0^i \sim X^{in} \text{ for some random variable } X^{in},$$

$$(1-\varepsilon)\mathbb{E}\Big[e^{-\beta L(X^{in})}\Big] \geq \frac{2\lambda+\sigma^2}{2\lambda-\sigma^2} C_L \beta e^{-\beta L_m} \sum_{l=1}^{d} \mathbb{E}\Big[\max_{1\leq i\leq N} (x_0^{i,l} - \bar{x}_0^l)^2\Big],$$

for some $0 < \varepsilon < 1$. Then, one has

$$\text{ess inf}_{\omega\in\Omega} L(X^\infty(\omega)) \leq \text{ess inf}_{\omega\in\Omega} L(X^{in}(\omega)) + \mathcal{O}\left(\frac{1}{\beta}\right), \quad \text{for } \beta \gg 1.$$

Consequently, if the global minimizer X^* of L is contained in $\text{supp}\,\text{law}(X^{in})$, then

$$\text{ess inf}_{\omega\in\Omega} L(X^\infty(\omega)) \leq L_m + \mathcal{O}\left(\frac{1}{\beta}\right), \quad \text{for } \beta \gg 1.$$

Idea of Proof: By technical calculations, one can derive

$$-\frac{1}{\beta}\log \mathbb{E} e^{-\beta L(X^{\infty})} \leq -\frac{1}{\beta}\log \mathbb{E} e^{-\beta L(X^{in})} - \frac{1}{\beta}\log \varepsilon.$$

Now we use Laplace's principle in the limit $\beta \to \infty$ to get

$$\text{ess inf}_{\omega\in\Omega} L(X^{\infty}(\omega)) \leq \text{ess inf}_{\omega\in\Omega} L(X^{in}(\omega)) + O\left(\frac{1}{\beta}\right) \quad \text{for } \beta \gg 1.$$

Summary

1. We provided a convergence analysis for the CBO algorithm under some conditions on system parameters and initial data.
2. Our theoretical analysis might be used for the convergence analysis for biologically motivated metaheuristic algorithms, e.g., Particle Swarm Optimization.

Thank you for your attention !!!