Weak and Semi-contraction Theory for Network Systems

Saber Jafarpour

Center for Control, Dynamical Systems & Computation
University of California at Santa Barbara

April 12, 2021

Acknowledgment

Francesco Bullo
UCSB

Alexander Davydov
UCSB

Pedro Cisneros-Velarde
UCSB

SJ and P. Cisneros-Velarde and F. Bullo. Weak and Semi-Contraction for Network Systems and Diffusively-Coupled Oscillators. IEEE Transactions on Automatic Control, accepted, Mar. 2021.

P. Cisneros-Velarde and SJ and F. Bullo. Distributed and time-varying primal-dual dynamics via contraction analysis. Submitted, May 2020.

SJ and A. Davydov and F. Bullo. Non-Euclidean Contraction Theory for Monotone and Positive Systems. Submitted, Apr. 2021.

Introduction: Large-scale Nonlinear Networks

Power grids

Brain neural network

Transportation network

Nonlinearity:

- Multiple equilibria
- Transient stability
- Cluster synchronization

Large-scale:

- Stochastic
- Distributed

- *"... As [power] systems become more heavily loaded, nonlinearities play an increasingly important role in power system behavior ... "* [I. Hiskens,1995]
- *"... in Oahu, Hawaii, at least 800,000 micro-inverters interconnect photovoltaic panels to the grid... "* [IEEE Spectrum, 2015]

Why stability wrt *non-Euclidean* metrics?

1. more systematic and efficient stability analysis:
    - conservation law
    - symmetry
2. ℓ_2-norm stability analysis gives conservative estimates:
    - synchronization tests for networks of oscillators
    - distance of power grids to frequency instability
3. error analysis for large-scale networks:

$$x \in \mathcal{N}(0, \sigma I_n) \quad \Longrightarrow \quad \mathbb{E}(\|x\|_2^2) = n\sigma^2, \quad \mathbb{E}(\|x\|_\infty^2) \sim 2\ln(n)\sigma^2$$

Presentation outline

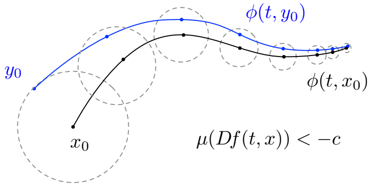
- review of contraction theory
  - definitions and Historical notes
  - results and computations
  - challenges for studying networks
- weakly-contracting systems
  - definition
  - dichotomy in asymptotic behavior
  - example: distributed primal-dual
- semi-contracting systems
  - definition
  - results and computations
  - example: diffusively-coupled oscillators

Definition: Contracting systems

$\dot{x} = f(t, x)$ is contracting wrt to $\|\cdot\|$ with rate $c > 0$:

$$\|x(t) - y(t)\| \leq e^{-ct}\|x(0) - y(0)\|.$$

Contracting system: flow is a contracting map.



Contraction theory: a brief review
Historical notes

- D. C. Lewis. Metric properties of differential equations.
  American Journal of Mathematics, 71(2):294–312, 1949
- B. P. Demidovich. Dissipativity of a nonlinear system of differential equations.
  Uspekhi Matematicheskikh Nauk, 16(3(99)):216, 1961
- **Application in control theory**: W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems.
  Automatica, 34(6):683–696, 1998
- **Differential framework**: F. Forni and R. Sepulchre. A differential Lyapunov framework for contraction analysis.
  IEEE Trans. Autom. Control, 59(3):614–628, 2014
- **Review**: J. Jouffroy and T. I. Fossen. A tutorial on incremental stability analysis using contraction theory.
  Modeling, Identification and Control, 31(3):93–106, 2010
- **Review**: Z. Aminzare and E. D. Sontag. Contraction methods for nonlinear systems: A brief introduction and some open problems.
  In *Proc CDC*, pages 3835–3847, Dec. 2014

Highly ordered asymptotic and transient behaviors:

1. initial conditions are forgotten
2. time-invariant f: unique globally stable equilibrium
3. periodic f: unique globally stable periodic solution
4. robustness properties: input-to-state stability
5. finite input-state gain in the presence of (Lipschitz continuous) unmodeled dynamics.

Combines in unified coherent framework results from:

1. stability notions and Lyapunov functions
2. Banach contraction and Brouwer fixed point theorems,
3. monotone systems theory, and
4. geometry of Banach, Riemannian and Finsler spaces

Contraction theory includes results about (1) the existence of equilibria (not just their stability), (2) entrainment to periodic solutions, and (3) the dichotomy for weakly-contracting systems.

- **Limitation:** analysis of multistability
- **Limitation:** analysis of regions of attraction

Definition: matrix measure

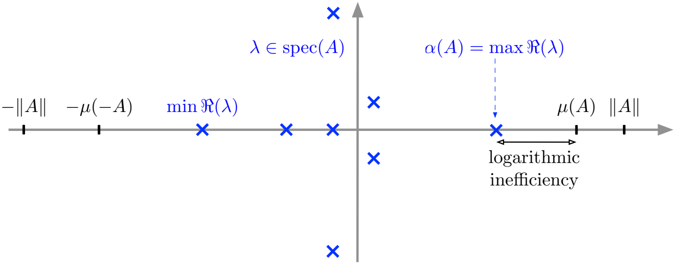
The **matrix measure** of $A \in \mathbb{R}^{n\times n}$ wrt to $\|\cdot\|$:

$$\mu_{\|\cdot\|}(A) := \lim_{h\to 0^+} \frac{\|I_n + hA\| - 1}{h}.$$

- Directional derivative of norm $\|\cdot\|$ in direction of A,
- One-sided Lipschitz constant: E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer, 1993
- Logarithmic norm: T. Ström. On logarithmic norms. *SIAM Journal on Numerical Analysis*, 12(5):741–753, 1975
- scaling property: $\mu_{\|\cdot\|}(cA) = |c|\mu_{\|\cdot\|}(\mathrm{sign}(c)A)$,
- subadditivity: $\mu_{\|\cdot\|}(A + B) \leq \mu_{\|\cdot\|}(A) + \mu_{\|\cdot\|}(B)$,

Matrix measures and contraction theory

- spectral property: $-\|A\| = \Re(\lambda) \leq \mu_{\|\cdot\|}(A) \leq \|A\|$, for every $\lambda \in \mathrm{spec}(A)$



Theorem: contraction via matrix measures

$\dot{x} = f(t, x)$ with f continuously differentiable in x is contracting wrt to $\|\cdot\|$ with rate $c > 0$ iff

$$\mu_{\|\cdot\|}(Df(t, x)) \leq -c, \qquad \text{for every } t, x$$

Computing matrix measures

For $\|\cdot\|_{2,R}$ with invertible R:

$$\mu_{2,R}(Df(t,x)) \leq -c \iff (R^\top R)Df(t,x) + Df(t,x)^\top (R^\top R) \preceq -2c(R^\top R)$$

For $\|\cdot\|_{1,[\eta]}$ with $\eta \in \mathbb{R}^n_{>0}$ and $Df(t,x)$ Metzler:

$$\mu_{1,[\eta]}(Df(t,x)) \leq -c \iff \eta^\top Df(t,x) \leq -c\eta^\top$$
$$\mu_{\infty,[\eta]^{-1}}(Df(t,x)) \leq -c \iff Df(t,x)\eta \leq -c\eta$$

If f is polynomial in t and x,

1. can be checked using SOS programming
2. search for the weights ($R^\top R \succ 0$ and $\eta \in \mathbb{R}^n_{>0}$) using SOS programming

E. M. Aylward, P. A. Parrilo, and J.-J. E. Slotine. Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming. *Automatica*, 44(8):2163–2170, 2008.

Contraction theory for networks

Challenge: many real-world networks are not contracting.


Network flow system $\dot{x} = f(x)$ preserving commodity x:

$$\begin{aligned} \text{constant} &= \mathbb{1}_n^\top x(t) \\ &\implies 0 = \mathbb{1}_n^\top \dot{x}(t) = \mathbb{1}_n^\top f(x(t)) \\ &\implies \mathbb{0}_n = \mathbb{1}_n^\top Df(x(t)) \end{aligned}$$

If additionally f has Metzler Jacobian, then $\mu_1(Df(x)) = 0$.

Definition: Weakly-contracting systems

$\dot{x} = f(t,x)$ with f continuously differentiable in x is weakly-contracting wrt $\|\cdot\|$:

$$\mu_{\|\cdot\|}(Df(t,x)) \leq 0$$

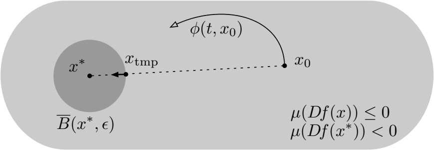
1. Lotka-Volterra population dynamics (Lotka, 1920; Volterra, 1928), (weakly-contracting wrt ℓ_1-norm)
2. Kuramoto oscillators (Kuramoto, 1975) and coupled swing equations (Bergen and Hill, 1981), (weakly-contracting wrt ℓ_1-norm and ℓ_∞-norm)
3. Daganzo's cell transmission model for traffic networks (Daganzo, 1994), (weakly-contracting wrt ℓ_1-norm)
4. compartmental systems in biology, medicine, and ecology (Sandberg, 1978; Maeda et al., 1978). (weakly-contracting wrt ℓ_1-norm)
5. saddle-point dynamics for optimization of weakly-convex functions (Arrow et al., 1958). (weakly-contracting wrt ℓ_2-norm)

Theorem: Dichotomy for weakly-contracting systems

For a weakly-contracting system $\dot{x} = f(x)$, either

1. f has no equilibrium and every trajectory is unbounded, or
2. f has at least one equilibrium x^* and every trajectory is bounded, and:
   - (i) each equilibrium x^{**} is stable with weak Lyapunov function $x \mapsto \|x - x^{**}\|$,
   - (ii) if the norm $\|\cdot\|$ is a (p, R)-norm, $p \in \{1, \infty\}$ and f is piecewise real analytic, then every trajectory converges to the set of equilibria,
   - (iii) x^* is locally asy stable $\implies$ x^* is globally asy stable.

Idea of the proof



Optimization problem

$$\min_{x\in\mathbb{R}^k} f(x) = \min_{x\in\mathbb{R}^k} \sum_{i=1}^{n} f_i(x)$$

Distributed implementation

- n agents communicate over a undirected weighted graph G,
- agent i have access to function f_i and can exchange x_i with its neighbors.

$$\min_{x\in\mathbb{R}^k} \quad \sum_{i=1}^{n} f_i(x_i)$$
$$x_1 = x_2 = \ldots = x_n$$

In matrix form by assuming $x = (x_1^\top, \ldots, x_n^\top)^\top \in \mathbb{R}^{nk}$:

$$\min_{x\in\mathbb{R}^k} \quad \sum_{i=1}^{n} f_i(x_i)$$
$$(L \otimes I_k)x = \mathbb{0}_{nN}$$

If each f_i is continuously differentiable in x_i:

Lagrangian

$\mathcal{L}(x,\nu) = \sum_{i=1}^n f_i(x_i) + \nu^\top (L \otimes I_k)x$

Distributed primal-dual algorithm (component form):

$$\dot{x}_i = -\frac{\partial \mathcal{L}}{\partial x_i} = -\nabla f_i(x_i) - \sum_{j=1}^n a_{ij}(\nu_i - \nu_j),$$

$$\dot{\nu}_i = \frac{\partial \mathcal{L}}{\partial \nu_i} = \sum_{j=1}^n a_{ij}(x_i - x_j)$$

Distributed primal-dual algorithm (vector form):

$$\dot{x} = -\nabla f(x) - (L \otimes I_k)\nu,$$

$$\dot{\nu} = (L \otimes I_k)x$$

Example: Primal-dual algorithm
Stability and rate of convergence

Assume

1. f has a minimum $x^* \in \mathbb{R}^k$,
2. for each $i \in \{1, \dots, n\}$, f_i is twice differentiable, $\nabla^2 f_i(x) \succeq 0$ for all x, and $\nabla^2 f_i(x^*) \succ 0$, and
3. the undirected weighted graph G is connected with Laplacian L.

Theorem: Distributed primal-dual dynamics

The distributed primal-dual algorithm

1. is weakly-contracting wrt ℓ_2-norm,
2. its trajectory $(x(t), \nu(t))$ converges exponentially to $(\mathbb{1}_n \otimes x^*, \mathbb{1}_n \otimes \nu^*)$, where $\nu^* = \sum_{i=1}^n \nu_i(0)$,
3. its exponential convergence rate is $-\alpha_{\text{ess}}\left(\begin{bmatrix} -\nabla^2 f(x^*) & -L \otimes I_k \\ L \otimes I_k & \mathbb{0} \end{bmatrix} \right)$ where

$$\alpha_{\text{ess}}(A) := \max\{\Re(\lambda) \mid \lambda \in \text{spec}(A) \setminus \{0\}\}.$$

- Idea: flows converges to each other only in **certain directions**.

Definition: Semi-norms

$|||\cdot|||$ is a *semi-norm* if

1. $|||cv||| = |c|\,|||v|||$, for every $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$;
2. $|||v + w||| \leq |||v||| + |||w|||$, for every $v, w \in \mathbb{R}^n$.

- Define the subspace $\operatorname{Ker} |||\cdot||| = \{v \in \mathbb{R}^n \mid |||v||| = 0\}$.
- Example: for $k < n$, $R \in \mathbb{R}^{k\times n}$, and norm $\|\cdot\|$, we get $|||x|||_R = \|Rx\|$.
- Example: for a network G with edge set $\mathcal{E}$ and incidence matrix B:

$$|||x|||_{\mathcal{E}} := \max_{(i,j)\in\mathcal{E}} |x_i - x_j| = \|B^\top x\|_\infty$$

For strongly connected graphs $\operatorname{Ker} |||\cdot|||_{\mathcal{E}} = \operatorname{span}\{\mathbb{1}_n\}$

Definition: Matrix semi-measures

The **matrix semi-measure** of $A \in \mathbb{R}^{n\times n}$ wrt $|||\cdot|||$:

$$\mu_{|||\cdot|||}(A) = \lim_{h\to 0^+} \frac{|||I_n + hA||| - 1}{h}.$$

- Directional derivative of $|||\cdot|||$ in direction of A.
- scaling property: $\mu_{|||\cdot|||}(cA) = |c|\mu_{|||\cdot|||}(\mathrm{sign}(c)A)$,
- subadditivity: $\mu_{|||\cdot|||}(A+B) \le \mu_{|||\cdot|||}(A) + \mu_{|||\cdot|||}(B)$,
- if Ker $|||\cdot|||$ is invariant under A then $\Re(\lambda) \le \mu_{|||\cdot|||}(A)$, for every $\lambda \in \mathrm{spec}_{\mathrm{Ker}\,|||\cdot|||^{\perp}}(A^\top)$.

Definition: Semi-contracting systems

$\dot{x} = f(t, x)$ with f continuously differentiable in x is semi-contracting wrt the semi-norm $|||\cdot|||$ with rate $c > 0$:

$$\mu_{|||\cdot|||}(Df(t,x)) \leq -c$$

1. Kuramoto oscillators (Kuramoto, 1975) and coupled swing equations (Bergen and Hill, 1981), (semi-contracting wrt ℓ_1-norm)
2. Chua's diffusively-coupled circuits (Wu and Chua, 1995), (semi-contracting wrt ℓ_2-norm)
3. morphogenesis in developmental biology (Turing, 1952), (semi-contracting wrt ℓ_1-norm)
4. Goodwin model for oscillating auto-regulated gene (Goodwin, 1965). (semi-contracting wrt ℓ_1-norm)

Theorem: Semi-contracting systems

Consider $\dot{x} = f(t, x)$ with f continuously differentiable in x and assume

- f is semi-contracting wrt the semi-norm $|||\cdot|||$ with rate $c > 0$, and
- **(Affine invariance)**: $f(t, x^* + \operatorname{Ker} |||\cdot|||) \subseteq \operatorname{Ker} |||\cdot|||$ for every t

Then,

1. for every trajectory $x(t)$,

$$|||x(t) - x^*||| \leq e^{-ct} |||x(0) - x^*|||, \qquad \text{for every } t \geq 0.$$

2. every trajectory converges to the affine invariant subspace $x^* + \operatorname{Ker} |||\cdot|||$.

- partial contraction (only for ℓ_2-norms): W. Wang and J.-J. E. Slotine. On partial contraction analysis for coupled nonlinear oscillators. *Biological Cybernetics*, 92(1):38–53, 2005
- horizontal contraction (stronger assumptions): F. Forni and R. Sepulchre. A differential Lyapunov framework for contraction analysis. *IEEE Trans. Autom. Control*, 59(3):614–628, 2014

- For semi-norm $\|\cdot\|_{2,R}$:

$$\mu_{2,R}(Df(t,x)) \leq -c \iff$$
$$x^\top[(R^\top R)Df(t,x) + Df(t,x)^\top(R^\top R) + 2cR^\top R]x \leq 0, \quad x \in \mathrm{Ker}\, R^\perp$$

- $R \in \mathbb{R}^{k\times n}$ such that $\mathrm{Ker}(R) = \mathrm{span}\{v_1, \dots, v_{n-k}\}$
- we set $Q = \left[\mathbb{0}_{n\times(n-k)}, v_{n-k+1}, \dots, v_n\right] \in \mathbb{R}^{n\times n}$

SOS programming for checking semi-contractivity

$$y^\top Q^\top(-R^\top R Df(t,x) - (Df(t,x))^\top R^\top R - 2cR^\top R)Qy \in \mathrm{SOS}$$
$$y^\top R^\top R y \in \mathrm{SOS}$$
$$Rv_1 = Rv_2 = \cdots = Rv_{n-k} = \mathbb{0}_n$$

Example: Diffusively-coupled oscillators

- n agents connected by a weighted undirected graph G,
- identical internal dynamics $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^k \to \mathbb{R}^k$

$$\dot{x}_i = f(t, x_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j), \quad i \in \{1, \dots, n\}$$

- **synchronization**:

$$\lim_{t\to\infty} \|x_i - x_j\| = 0 \qquad \text{for every } i, j$$

- synchronization of diffusively-coupled oscillators:

  1. contractivity of the internal dynamics
  2. strength of the diffusive coupling

Example: Diffusively-coupled oscillators

Introduce a **local-global mixed norm:** $(2,p)$-tensor norm on $\mathbb{R}^{nk} = \mathbb{R}^n \otimes \mathbb{R}^k$

$$\|u\|_{(2,p)} = \inf \left\{ \Big(\sum_{i=1}^{r} \|v^i\|_2^2 \|w^i\|_p^2 \Big)^{\frac{1}{2}} \mid u = \sum_{i=1}^{r} v^i \otimes w^i \right\}.$$

- closely related to, but different from, the projective tensor product norm

  R. A. Ryan. *Introduction to Tensor Products of Banach Spaces.*
  Springer, 2002

- different from the mixed global norm

  G. Russo, M. Di Bernardo, and E. D. Sontag. A contraction approach to the hierarchical analysis and design of networked systems.
  IEEE Trans. Autom. Control, 58(5):1328–1331, 2013

- **Global norm**: ℓ_2-norm for the interactions between agents
- **Local norm**: ℓ_p-norm for internal dynamics of each agent

Example: Diffusively-coupled oscillators

Introduce $R \in \mathbb{R}^{(n-1)\times n}$

$$R = [v_2, v_3, \dots, v_n]^\top$$

where $\{v_1 = \mathbb{1}_n, v_2, \dots, v_n\}$ are eigenvectors of the Laplacian matrix L

- $(R \otimes I_k)x$ measures **dissimilarity** of the states x_i

$$x = \mathbb{1}_n \otimes x^* \implies$$

$$(R \otimes I_k)x = (R \otimes I_k)(\mathbb{1}_n \otimes x^*) = R\mathbb{1}_n \otimes x^* = \mathbb{0}_{(n-1)\times k}.$$

Example: Diffusively-coupled oscillators

- G is an undirected weighted graph with Laplacian L,
- $p \in [1, \infty]$, $Q \in \mathbb{R}^{k \times k}$

$$\dot{x}_i = f(t, x_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j), \quad i \in \{1, \dots, n\}$$

Theorem: diffusively-coupled oscillators are semi-contracting

Suppose that

$$\mu_{p,Q}(Df(t,x)) \leq \lambda_2(L) - c, \qquad \text{for every } t, x$$

then

1. for every trajectory $x(t)$,

$$\|x(t) - \mathbb{1}_n \otimes x_{\text{ave}}(t)\|_{2,p,(R \otimes Q)} \leq e^{-ct} \|x(0) - \mathbb{1}_n \otimes x_{\text{ave}}(0)\|_{2,p,(R \otimes Q)}.$$

2. the system achieves synchronization: $\lim_{t \to \infty} x(t) = \mathbb{1}_n \otimes x_{\text{ave}}(t)$

where $x_{\text{ave}}(t) = \frac{1}{n} \sum_{i=1}^{n} x_i(t)$

Example: Diffusively-coupled oscillators

$$\mu_{p,Q}(Df(t,x)) \leq \lambda_2(L) - c, \qquad \text{for every } t, x$$

- trade off between **internal dynamics** and **coupling strength**
- f time-invariant: every trajectory converges to the unique equilibrium point.
- f periodic: every trajectory converges to the unique periodic orbit.
- Unlike small-gain theorems where coupling has destabilizing effect: diffusive coupling helps synchronization.
- Unstable dynamics f, sufficiently strong coupling $\implies$ $\lambda_2(L)$ large $\implies$ the network synchronizes.

Summary

- Review contraction theory and matrix measures
- Two extensions of classical contraction:
  - weak contraction
  - semi-contraction
- Properties of weakly-contracting and semi-contracting systems

Future research

- Contraction-based compositional analysis of interconnected systems
  - scalable stability certificates using non-Euclidean contraction.
- Optimization algorithms using contraction theory
  - extension to gradient descent algorithms and time-varying algorithms.
  - connection with discrete-time algorithms for optimization.
- Learning stable system from trajectory data
  - use contraction condition as side-information in the optimization problem.