

# Contraction Theory for Networked Optimization and Control



Francesco Bullo

Center for Control,  
Dynamical Systems & Computation  
University of California at Santa Barbara  
<https://fbullo.github.io/ctds>

Workshop on Emerging Challenges of Network-Enabled Control and Optimization  
CDC Workshop, Singapore, Dec 12, 2023

# Acknowledgments



Veronica Centorrino  
Scuola Sup Meridionale



Lily Cothren  
UC Boulder



Emiliano Dall'Anese  
UC Boulder



Alex Davydov  
UC Santa Barbara



Anand Gokhale  
UC Santa Barbara

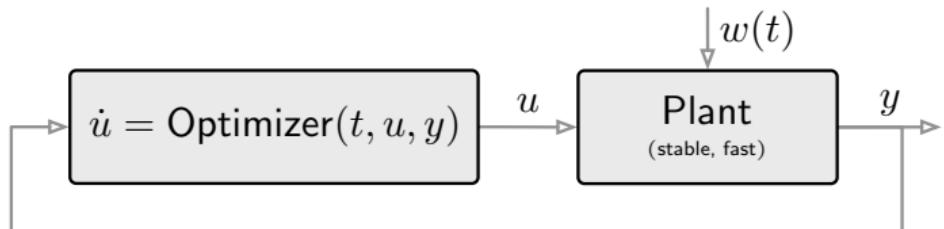


Giovanni Russo  
Univ Salerno

- A. Davydov, V. Centorrino, A. Gokhale, G. Russo, and F. Bullo. Contracting dynamics for time-varying convex optimization. *IEEE Transactions on Automatic Control*, June 2023. Submitted
- L. Cothren, F. Bullo, and E. Dall'Anese. Singular perturbation via contraction theory. *IEEE Transactions on Automatic Control*, Oct. 2023. Submitted

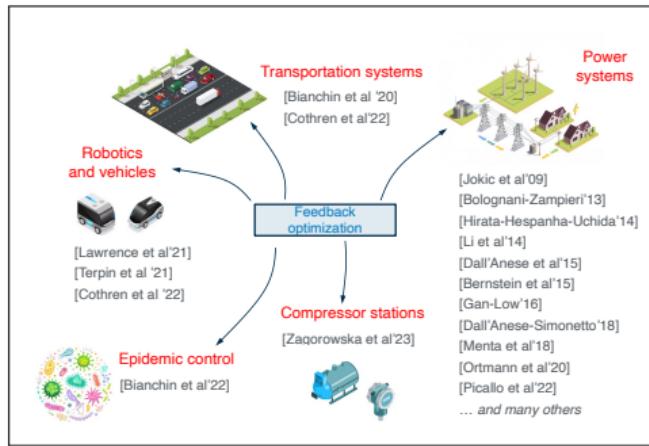
# Solving optimization problems via dynamical systems

- studies in linear and nonlinear programming (Arrow, Hurwicz, and Uzawa 1958)
- neural networks (Hopfield and Tank 1985) and analog circuits (Kennedy and Chua 1988)
- optimization on manifolds (Brockett 1991)
- ...



## Continuous-time optimizations solvers

- ① distributed optimization
- ② online feedback optimization
- ③ parametric convex optimization
- ④ model predictive control
- ⑤ control barrier functions

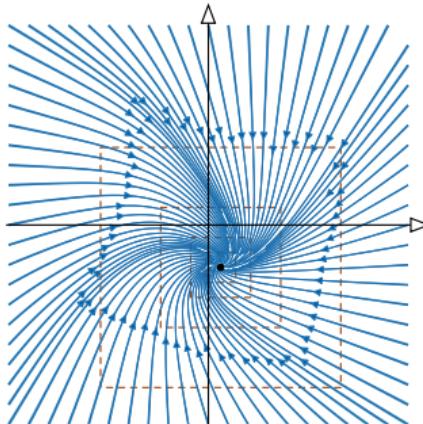


Online feedback optimization

Slide courtesy of Emiliano Dall'Anese, University of Colorado Boulder

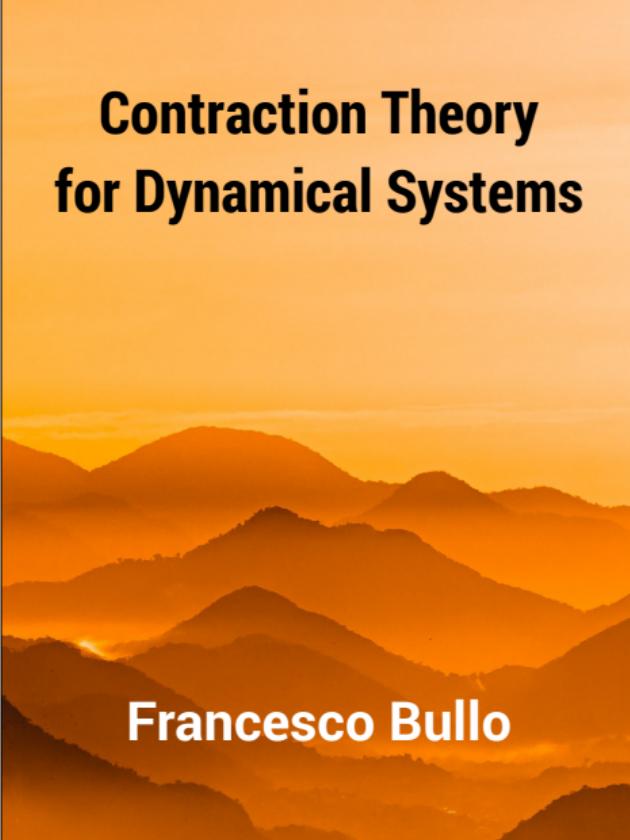
**contractivity = robust computationally-friendly stability**

fixed point theory + Lyapunov stability theory + geometry of metric spaces



**search for** contraction properties  
**design** engineering systems to be contracting  
**verify** correct/safe behavior via known Lipschitz constants

- 2023 ACC Workshop on "Contraction Theory for Systems, Control, and Learning"  
<http://motion.me.ucsb.edu/contraction-workshop-2023>
- Tutorial session: <https://sites.google.com/view/contractiontheory> "Contraction Theory for Machine Learning" (PDFs and youtube videos) at the 2021 IEEE CDC conference, by Soon-Jo Chung, Jean-Jacques Slotine, and Hiroyasu Tsukamoto
- Plenary presentation: (Slides  
<https://fbullo.github.io/talks/2022-12-FBullo-ContractionSystemsControl-CDC.pdf>)  
"Contraction Theory in Systems and Control" by Francesco Bullo at the 2022 IEEE CDC
- Comprehensive tutorial slides: <https://fbullo.github.io/ctds>
- Youtube lectures: "Minicourse on Contraction Theory" by Francesco Bullo, Fall 2023. Youtube lectures <https://youtu.be/FQV5PrRHks8> 12h in 6 lectures, with chapters
- Three CDC2023 invited sessions on **Contraction Theory for Analysis, Synchronization and Regulation**, tomorrow Wednesday!



# Contraction Theory for Dynamical Systems

Francesco Bullo

**Contraction Theory for Dynamical Systems**, Francesco Bullo, KDP,  
1.1 edition, 2023, ISBN 979-8836646806

- ➊ Textbook with exercises and answers. Format: textbook, slides, and paperbook
- ➋ Content:
  - Fixed point theory
  - Theory of contracting dynamics on vector spaces
  - Applications to nonlinear and interconnected systems
- ➌ Self-Published and Print-on-Demand at:  
<https://www.amazon.com/dp/B0B4K1BTF4>
- ➍ PDF Freely available at  
<https://fbullo.github.io/ctds>
- ➎ Future version to include: systems on Riemannian manifolds, homogeneous spaces, and solid cones
  - "Continuous improvement is better than delayed perfection" **Mark Twain**

# Outline

§1. Introduction

§2. Basic contractivity concepts

§3. Gradient systems defined by strongly convex functions are contracting

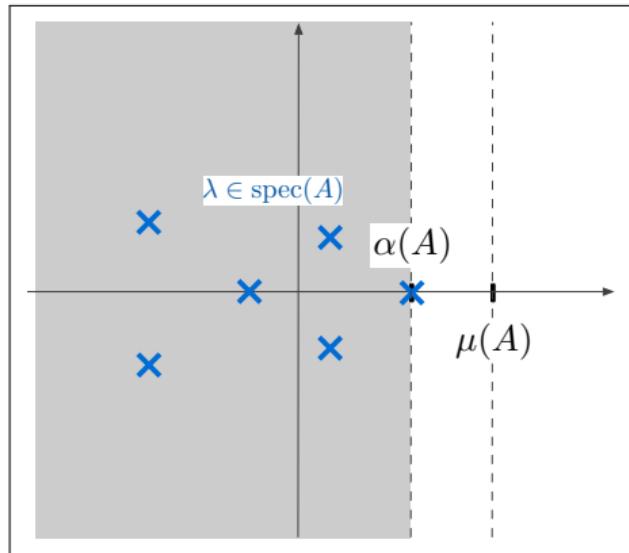
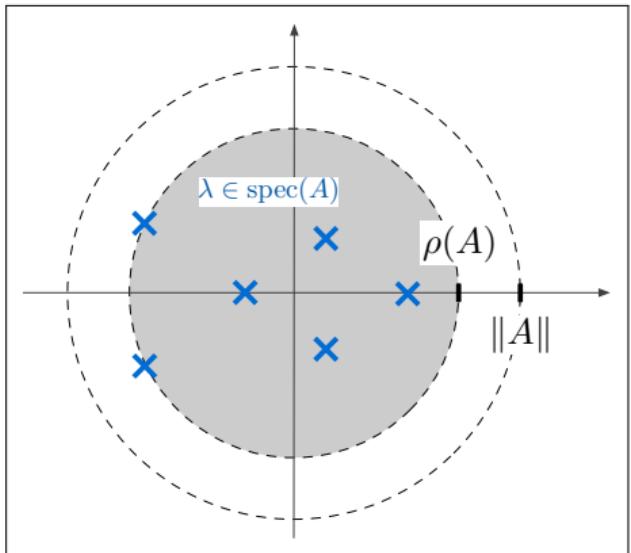
§4. Equilibrium tracking in parametric optimization

§5. Application: Online feedback optimization

§6. Conclusions

# Induced matrix norms

Vector norm	Induced matrix norm	Induced matrix log norm
$\ x\ _1 = \sum_{i=1}^n  x_i $	$\ A\ _1 = \max_{j \in \{1, \dots, n\}} \sum_{i=1}^n  a_{ij} $ = max column "absolute sum" of $A$	$\mu_1(A) = \max_{j \in \{1, \dots, n\}} \left( a_{jj} + \sum_{i=1, i \neq j}^n  a_{ij}  \right)$ absolute value only off-diagonal
$\ x\ _2 = \sqrt{\sum_{i=1}^n x_i^2}$	$\ A\ _2 = \sqrt{\lambda_{\max}(A^\top A)}$	$\mu_2(A) = \lambda_{\max}\left(\frac{A + A^\top}{2}\right)$
$\ x\ _\infty = \max_{i \in \{1, \dots, n\}}  x_i $	$\ A\ _\infty = \max_{i \in \{1, \dots, n\}} \sum_{j=1}^n  a_{ij} $ = max row "absolute sum" of $A$	$\mu_\infty(A) = \max_{i \in \{1, \dots, n\}} \left( a_{ii} + \sum_{j=1, j \neq i}^n  a_{ij}  \right)$ absolute value only off-diagonal



$\dot{x} = F(x)$       on  $\mathbb{R}^n$  with norm  $\|\cdot\|$  and induced log norm  $\mu(\cdot)$

## One-sided Lipschitz constant

$$\begin{aligned}\text{osLip}(F) &= \inf\{b \in \mathbb{R} \text{ such that } \|F(x) - F(y), x - y\| \leq b\|x - y\|^2 \text{ for all } x, y\} \\ &= \sup_x \mu(DF(x))\end{aligned}$$

For **scalar map**  $f$ ,  $\text{osLip}(f) = \sup_x f'(x)$

For **affine map**  $F_A(x) = Ax + a$

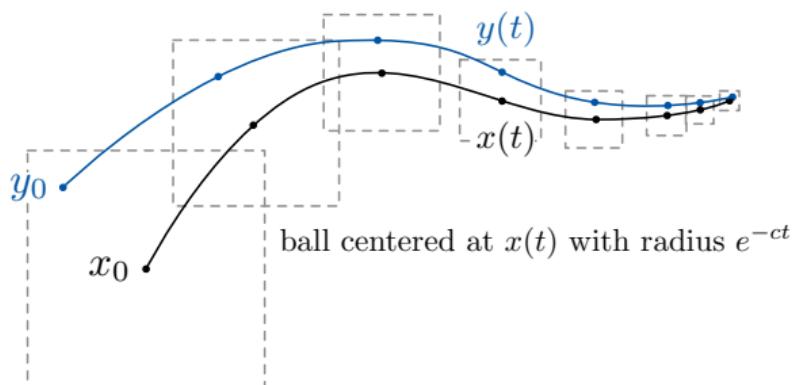
$$\text{osLip}_{2,P}(F_A) = \mu_{2,P}(A) \leq \ell \iff A^\top P + AP \preceq 2\ell P$$

$$\text{osLip}_{\infty,\eta}(F_A) = \mu_{\infty,\eta}(A) \leq \ell \iff a_{ii} + \sum_{j \neq i} |a_{ij}| \eta_i / \eta_j \leq \ell$$

## Banach contraction theorem for continuous-time dynamics:

If  $-c := \text{osLip}(F) < 0$ , then

- ①  $F$  is **infinitesimally contracting** = distance between trajectories decreases exp fast ( $e^{-ct}$ )
- ②  $F$  has a unique, glob exp stable equilibrium  $x^*$
- ③ global Lyapunov functions  $V_1(x) = \|x - x^*\|^2$  and  $V_2(x) = \|F(x)\|^2$



## Euler discretization theorem for contracting dynamics

Given arbitrary norm  $\|\cdot\|$  and differentiable  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , equivalent statements

- ①  $\dot{x} = F(x)$  is infinitesimally contracting
- ② there exists  $\alpha > 0$  such that  $x_{k+1} = x_k + \alpha F(x_k)$  is contracting

Interconnected subsystems:  $x_i \in \mathbb{R}^{N_i}$  and  $x_{-i} \in \mathbb{R}^{N-N_i}$ :

$$\dot{x}_i = F_i(x_i, x_{-i}), \quad \text{for } i \in \{1, \dots, n\}$$

**Network contraction theorem.** Assume

- **contractivity wrt  $x_i$ :**  $\text{osLip}_{x_i}(F_i) \leq -c_i < 0$ , uniformly in  $x_{-i}$
- **Lipschitz wrt  $x_j$ ,  $j \neq i$ :**  $\text{Lip}_{x_j}(F_i) \leq \ell_{ij}$ , uniformly in  $x_{-j}$

- the Lipschitz constants matrix  $\Gamma = \begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix}$  is **Hurwitz**

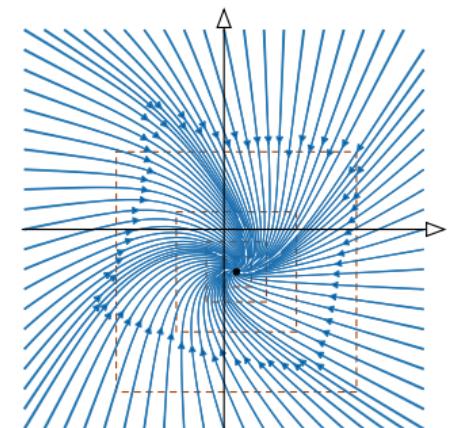
$\implies$  **interconnected system** is contracting wrt rate  $|\alpha(\Gamma)|$

**contractivity = robust computationally-friendly stability**

fixed point theory + Lyapunov stability theory + geometry of metric spaces

**highly-ordered transient and asymptotic behavior, no anonymous constants/functions:**

- ① unique globally exponential stable equilibrium  
& two natural Lyapunov functions
- ② robustness properties
  - bounded input, bounded output (iss)
  - finite input-state gain
  - robustness margin wrt unmodeled dynamics
  - robustness margin wrt delayed dynamics
- ③ periodic input, periodic output
- ④ modularity and interconnection properties
- ⑤ accurate numerical integration and equilibrium point computation



**search for** contraction properties  
**design** engineering systems to be contracting  
**verify** correct/safe behavior via known Lipschitz constants

# Outline

§1. Introduction

§2. Basic contractivity concepts

§3. Gradient systems defined by strongly convex functions are contracting

§4. Equilibrium tracking in parametric optimization

§5. Application: Online feedback optimization

§6. Conclusions

**Kachurovskii's Theorem:** For differentiable  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , equivalent statements:

- ①  $f$  is **strongly convex** with parameter  $\nu$  (and minimum  $x^*$ )
- ②  $-\nabla f$  is  **$\nu$ -strongly infinitesimally contracting** (with equilibrium  $x^*$ ), that is

$$(-\nabla f(x) + \nabla f(y))^\top (x - y) \leq -\nu \|x - y\|_2^2$$

R. I. Kachurovskii. Monotone operators and convex functionals. *Uspekhi Matematicheskikh Nauk*, 15(4):213–215, 1960

## Example #1: Gradient dynamics for strongly convex function

Given differentiable, strongly convex  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  with parameter  $\nu > 0$ , **gradient dynamics**

$$\dot{x} = F_G(x) := -\nabla f(x)$$

$F_G$  is infinitesimally contracting wrt  $\|\cdot\|_2$  with rate  $\nu$

unique globally exp stable point is global minimum

## Example #2: Primal-dual gradient dynamics

strongly convex function  $f$

s.t.  $0 \prec \nu_{\min} I_n \preceq \text{Hess } f \preceq \nu_{\max} I_n$

constraint matrix  $A$

s.t.  $0 \prec a_{\min} I_m \preceq AA^\top \preceq a_{\max} I_m$

(independent rows)

**linearly constrained optimization:**

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{subj. to } Ax = b \end{aligned}$$

**primal-dual gradient dynamics:**

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = F_{\text{PDG}}(x, \lambda) := \begin{bmatrix} -\nabla f(x) - A^\top \lambda \\ Ax - b \end{bmatrix}$$

$F_{\text{PDG}}$  is infinitesimally contracting wrt  $\|\cdot\|_{2,P^{1/2}}$  with rate  $c$

$$P = \begin{bmatrix} I_n & \alpha A^\top \\ \alpha A & I_m \end{bmatrix} \text{ with } \alpha = \frac{1}{2} \min \left\{ \frac{1}{\nu_{\max}}, \frac{\nu_{\min}}{a_{\max}} \right\} \quad \text{and} \quad c = \frac{1}{4} \min \left\{ \frac{a_{\min}}{\nu_{\max}}, \frac{a_{\min}}{a_{\max}} \nu_{\min} \right\}$$

## Example #3: Laplacian-based distributed gradient

Given  $\Pi_n = I_n - \mathbb{1}_n \mathbb{1}_n^\top / n = \text{orthogonal projection onto } \text{span}\{\mathbb{1}_n\}^\perp$ ,

$$0 \prec \lambda_2 \Pi_n \preceq L \preceq \lambda_n I_n$$

**decomposable cost:**  $\min_{x \in \mathbb{R}^n} \sum_{i=1}^n f_i(x)$  where each  $f_i$  is  $\nu_i$ -strongly convex

$$\begin{cases} \min_{x_{[i]} \in \mathbb{R}} & \sum_{i=1}^n f_i(x_{[i]}) \\ \text{subj. to} & \sum_{j=1}^n a_{ij}(x_i - x_j) = 0 \end{cases}$$

**Laplacian-based distributed gradient** (primal-dual gradient,  $2n$  vars):

$$\begin{cases} \dot{x}_{[i]} = -\nabla f_i(x_{[i]}) - \sum_{j=1}^n a_{ij}(\lambda_i - \lambda_j) & \text{for each node } i \\ \dot{\lambda}_i = \sum_{j=1}^n a_{ij}(x_i - x_j) & \text{for each node } i \end{cases}$$

$F_{\text{Laplacian-DistributedG}}$  is infinitesimally contracting<sup>†</sup> with  $c = \frac{1}{4} \left( \frac{\lambda_2}{\lambda_n} \right)^2 \min_i \nu_i$

## Detour: Composite optimization and the proximal operator

**composite minimization** (cost = sum of terms with structurally different properties):

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} f(x, u) + g(x)$$

$f(x, u)$  is convex and differentiable in  $x$        $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is convex, closed, and proper (ccp)

**proximal operator:**

$$\operatorname{prox}_{\gamma g}(z) := \operatorname{argmin}_{x \in \mathbb{R}^n} g(x) + \frac{1}{2\gamma} \|x - z\|_2^2$$

generalized form of projection for nonsmooth/constrained/large-scale/distributed optimization

## Detour: Composite optimization and the proximal operator

**composite minimization** (cost = sum of terms with structurally different properties):

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} f(x, u) + g(x)$$

$f(x, u)$  is convex and differentiable in  $x$        $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is convex, closed, and proper (ccp)

**proximal operator:**

$$\operatorname{prox}_{\gamma g}(z) := \operatorname{argmin}_{x \in \mathbb{R}^n} g(x) + \frac{1}{2\gamma} \|x - z\|_2^2$$

generalized form of projection for nonsmooth/constrained/large-scale/distributed optimization

### Equivalence property:

①  $x^*$  is minimizer for:

$$\min_{x \in \mathbb{R}^n} f(x, u) + g(x)$$

②  $x^*$  is fixed point for:

$$x = \operatorname{prox}_{\gamma g}(x - \gamma \nabla f(x, u)) \quad \text{for all } \gamma$$

## Example #4: Proximal gradient dynamics

Equivalence property motivates:

**proximal gradient dynamics:**

$$\dot{x} = \mathsf{F}_{\text{ProxG}}(x) := -x + \text{prox}_{\gamma g}(x - \gamma \nabla f(x))$$

projected gradient descent is special case

$\mathsf{F}_{\text{ProxG}}$  is infinitesimally contracting wrt  $\|\cdot\|_2$

$$\text{for } 0 < \gamma < \frac{2}{\ell}, \quad \text{with rate} \quad c = 1 - \max\{|1 - \gamma\nu|, |1 - \gamma\ell|\},$$

$$\text{for } \gamma^* = \frac{2}{\nu + \ell}, \quad \text{with maximal rate} \quad c^* = \frac{2\nu}{\nu + \ell}$$

## Example #5: Saddle dynamics

Assume  $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$

- $x \mapsto f(x, y)$  is  $\nu_x$ -strongly convex, uniformly in  $y$
- $y \mapsto f(x, y)$  is  $\nu_y$ -strongly concave, uniformly in  $x$

**saddle dynamics (primal-descent / dual-ascent):**

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \mathsf{F}_S(x, y) := \begin{bmatrix} -\nabla_x f(x, y) \\ \nabla_y f(x, y) \end{bmatrix}$$

$\mathsf{F}_S$  is infinitesimally contracting wrt  $\|\cdot\|_2$  with rate  $\min\{\nu_x, \nu_y\}$

unique globally exp stable point is saddle point (min in  $x$ , max in  $y$ )

## Example #6: Pseudogradient play

Each player  $i$  aims to minimize its own cost function  $J_i(x_i, x_{-i})$  (not a potential game)

**pseudogradient dynamics (aka gradient play in game theory):**

$$\begin{aligned}\dot{x} &= \mathsf{F}_{\text{PseudoG}}(x) = -(\nabla_1 J_1(x_1, x_{-1}), \dots, \nabla_n J_n(x_n, x_{-n})) \quad (\text{stacked vector}) \\ \iff \dot{x}_i &= -\nabla_i J_i(x_i, x_{-i})\end{aligned}$$

• **strong convexity wrt  $x_i$ :**  $J_i$  is  $\mu_i$  strongly convex wrt  $x_i$ , uniformly in  $x_{-i}$

• **Lipschitz wrt  $x_{-i}$ :**  $\text{Lip}_{x_j}(\nabla_i J_i) \leq \ell_{ij}$ , uniformly in  $x_{-j}$

•  $\mathsf{F}_{\text{PseudoG}}$  gain matrix is Hurwitz

⇒  $\mathsf{F}_{\text{PseudoG}}$  is infinitesimally contracting wrt appropriate diag-weighted  $\|\cdot\|_2$

## Example #7: Best response play

Each player  $i$  aims to minimize its own cost function  $J_i(x_i, x_{-i})$

$\text{BR}_i : x_{-i} \rightarrow \operatorname{argmin}_{x_i} J_i(x_i, x_{-i})$  best response of player  $i$  wrt other decisions  $x_{-i}$

**best response dynamics:**

$$\begin{aligned}\dot{x} &= F_{\text{BR}}(x) := \text{BR}(x) - x \\ \iff \dot{x}_i &= \text{BR}_i(x_{-i}) - x_i\end{aligned}$$

- **strong convexity wrt  $x_i$ :**  $J_i$  is  $\mu_i$  strongly convex wrt  $x_i$ , uniformly in  $x_{-i}$
- **Lipschitz wrt  $x_{-i}$ :**  $\text{Lip}_{x_j}(\nabla_i J_i) \leq \ell_{ij}$ , uniformly in  $x_{-j}$   
 $\implies \text{BR}_i$  is Lipschitz wrt  $x_j$  with constant  $\ell_{ij}/\mu_i$
- $F_{\text{BR}}$  gain matrix is Hurwitz  $\iff$  BR is a discrete-time contraction  
 $\implies \text{BR} - \text{Id}$  is infinitesimally contracting wrt appropriate diag-weighted  $\|\cdot\|_2$

## Equivalent statements:

①  $F_{\text{PseudoG}}$  gain matrix:

$$\begin{bmatrix} -\mu_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -\mu_n \end{bmatrix} \text{ is Hurwitz}$$

②  $F_{\text{BR}}$  gain matrix:

$$\begin{bmatrix} -1 & \dots & \ell_{1n}/\mu_1 \\ \vdots & & \vdots \\ \ell_{n1}/\mu_n & \dots & -1 \end{bmatrix} \text{ is Hurwitz}$$

③ discrete-time  $F_{\text{BR}}$  gain matrix:

$$\begin{bmatrix} 0 & \dots & \ell_{1n}/\mu_1 \\ \vdots & & \vdots \\ \ell_{n1}/\mu_n & \dots & 0 \end{bmatrix} \text{ is Schur}$$

**Aggregative games:**  $J_i(x_i, x_{-i}) = f_i(x_i, \frac{1}{n} \sum_{j=1}^n x_j)$

assume  $f_i$  is  $\mu_i$ -strongly convex wrt  $x_i$  and  $\ell_i = \text{Lip}_y(\nabla_{x_i} f_i(x_i, y))$

$\mu_i > \ell_i$  for each agent  $i$   $\implies$  gain matrix is Hurwitz

# Outline

§1. Introduction

§2. Basic contractivity concepts

§3. Gradient systems defined by strongly convex functions are contracting

§4. Equilibrium tracking in parametric optimization

§5. Application: Online feedback optimization

§6. Conclusions

Many convex optimization problems can be solved with contracting dynamics

$$\dot{x} = F(x)$$

Many convex optimization problems can be solved with contracting dynamics

$$\dot{x} = F(x)$$

**contracting dynamics for parametric optimization**

$$\dot{x} = F(x, \theta)$$

**contracting dynamics for time-varying optimization**

$$\dot{x} = F(x, \theta(t))$$

# Equilibrium tracking

For parameter-dependent vector field  $\mathbf{F} : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$  and differentiable  $\theta : \mathbb{R}_{\geq 0} \rightarrow \Theta \subset \mathbb{R}^d$

$$\dot{x}(t) = \mathbf{F}(x(t), \theta(t))$$

Assume there exist norms  $\|\cdot\|_{\mathcal{X}}$  and  $\|\cdot\|_{\Theta}$  s.t.

- **contractivity wrt  $x$ :**  $\text{osLip}_x(\mathbf{F}) \leq -c < 0$ , uniformly in  $\theta$
- **Lipschitz wrt  $\theta$ :**  $\text{Lip}_{\theta}(\mathbf{F}) \leq \ell$ , uniformly in  $x$

## Theorem: Equilibrium tracking for contracting dynamics

- ① for each fixed  $\theta$ , there exists a unique equilibrium  $x^*(\theta)$
- ② the equilibrium map  $x^*(\cdot)$  is Lipschitz with constant  $\frac{\ell}{c}$

# Equilibrium tracking

For parameter-dependent vector field  $\mathbf{F} : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$  and differentiable  $\theta : \mathbb{R}_{\geq 0} \rightarrow \Theta \subset \mathbb{R}^d$

$$\dot{x}(t) = \mathbf{F}(x(t), \theta(t))$$

Assume there exist norms  $\|\cdot\|_{\mathcal{X}}$  and  $\|\cdot\|_{\Theta}$  s.t.

- **contractivity wrt  $x$ :**  $\text{osLip}_x(\mathbf{F}) \leq -c < 0$ , uniformly in  $\theta$
- **Lipschitz wrt  $\theta$ :**  $\text{Lip}_{\theta}(\mathbf{F}) \leq \ell$ , uniformly in  $x$

## Theorem: Equilibrium tracking for contracting dynamics

- ① for each fixed  $\theta$ , there exists a unique equilibrium  $x^*(\theta)$
- ② the equilibrium map  $x^*(\cdot)$  is Lipschitz with constant  $\frac{\ell}{c}$
- ③  $D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq -c \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \|\dot{\theta}(t)\|_{\Theta}$

## Consequences for tracking error

$$D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq -c \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \|\dot{\theta}(t)\|_{\Theta}$$

## Consequences for tracking error

$$D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq -c \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \|\dot{\theta}(t)\|_{\Theta}$$

- bounded input, bounded error  
with asymptotic bound:

$$\limsup_{t \rightarrow \infty} \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq \frac{\ell}{c^2} \limsup_{t \rightarrow \infty} \|\dot{\theta}(t)\|_{\Theta}$$

## Consequences for tracking error

$$D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq -c \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \|\dot{\theta}(t)\|_{\Theta}$$

- bounded input, bounded error  
with asymptotic bound:

$$\limsup_{t \rightarrow \infty} \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq \frac{\ell}{c^2} \limsup_{t \rightarrow \infty} \|\dot{\theta}(t)\|_{\Theta}$$

- bounded energy input, bounded energy error
- vanishing input, vanishing error
- exponentially vanishing input  $\sim e^{-ht}$ , exponentially vanishing error  $\sim e^{-\min\{c,h\}t}$
- periodic input, periodic error

# Outline

§1. Introduction

§2. Basic contractivity concepts

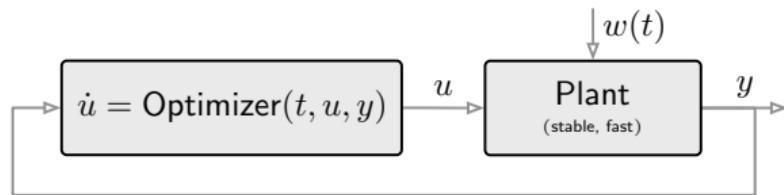
§3. Gradient systems defined by strongly convex functions are contracting

§4. Equilibrium tracking in parametric optimization

§5. Application: Online feedback optimization

§6. Conclusions

# Application: Online feedback optimization



## online feedback optimization

online optimization, optimization-based feedback, input/output regulation ...

$$\begin{cases} \min & \text{cost}_1(u) + \text{cost}_2(y) \\ \text{subj. to} & y = \text{Plant}(u, w(t)) \end{cases} \implies \begin{cases} \dot{u} = \text{Optimizer}(t, u, y) \\ y = \text{Plant}(u, w(t)) \end{cases}$$

A. Hauswirth, S. Bolognani, G. Hug, and F. Dorfler. Timescale separation in autonomous optimization. *IEEE Transactions on Automatic Control*, 66(2):611–624, 2021. [doi](#)

G. Bianchin, J. Cortés, J. I. Poveda, and E. Dall'Anese. Time-varying optimization of LTI systems via projected primal-dual gradient flows. *IEEE Transactions on Control of Network Systems*, 9(1):474–486, 2022. [doi](#)

## Example #8: Gradient controller

- fast/stable LTI plant with control input  $u$  and state/measurement disturbance  $w(t)$ :

$$\begin{aligned}\epsilon \dot{x} &= Ax + Bu + Ew(t) && A \text{ Hurwitz} \\ y &= Cx + Dw(t)\end{aligned}$$

- in singular perturbation limit as  $\epsilon \rightarrow 0^+$ , **steady state map (Y<sub>u</sub> and Y<sub>w</sub>)**

$$y = \underbrace{-CA^{-1}B u}_{=: Y_u} + \underbrace{(D - CA^{-1}E) w}_{=: Y_w}$$

## Example #8: Gradient controller

- fast/stable LTI plant with control input  $u$  and state/measurement disturbance  $w(t)$ :

$$\begin{aligned}\epsilon \dot{x} &= Ax + Bu + Ew(t) && A \text{ Hurwitz} \\ y &= Cx + Dw(t)\end{aligned}$$

- in singular perturbation limit as  $\epsilon \rightarrow 0^+$ , **steady state map ( $Y_u$  and  $Y_w$ )**

$$y = \underbrace{-CA^{-1}B u}_{=: Y_u} + \underbrace{(D - CA^{-1}E) w}_{=: Y_w}$$

- define **cost function  $\mathcal{E}$  on  $u$  and  $y$** :

$$\mathcal{E}(u, w) = \phi(u) + \psi(Y_u u + Y_w w), \quad (\phi \text{ is } \nu\text{-strongly convex and } \psi \text{ is convex})$$

and note

$$\begin{aligned}\nabla_u \mathcal{E}(u, w) &= \nabla \phi(u) + Y_u^\top \nabla \psi(Y_u u + Y_w w) \\ &= \nabla \phi(u) + Y_u^\top \nabla \psi(y) \quad (\text{no need to measure } w(t))\end{aligned}$$

## Example #8: Gradient controller

**equilibrium trajectory** let  $u^*(t)$  be solution to

$$\begin{aligned} \min_u \quad & \phi(u) + \psi(y(t)) && (\nu\text{-strongly convex } \phi, \text{ convex } \psi) \\ \text{subj to} \quad & y(t) = Y_u u + Y_w w(t) \end{aligned}$$

**gradient controller**

$$\dot{u} = F_{\text{GradCtrl}}(u, w) := -\nabla \mathcal{E}_u(u, w) = -\nabla \phi(u) - Y_u^\top \nabla \psi(Y_u u + Y_w w)$$

## Example #8: Gradient controller

**equilibrium trajectory** let  $u^*(t)$  be solution to

$$\begin{aligned} \min_u \quad & \phi(u) + \psi(y(t)) && (\nu\text{-strongly convex } \phi, \text{ convex } \psi) \\ \text{subj to} \quad & y(t) = Y_u u + Y_w w(t) \end{aligned}$$

**gradient controller**

$$\dot{u} = F_{\text{GradCtrl}}(u, w) := -\nabla \mathcal{E}_u(u, w) = -\nabla \phi(u) - Y_u^\top \nabla \psi(Y_u u + Y_w w)$$

### Equilibrium tracking for the gradient controller

- ①  $\text{osLip}_u(F_{\text{GradCtrl}}) \leq -\nu$  (gradient of  $\nu$ -strongly convex function)
- ②  $\text{Lip}_w(F_{\text{GradCtrl}}) = \ell_w := \|Y_u^\top\| \text{Lip}(\nabla \psi) \|Y_w\|$

$$\limsup_{t \rightarrow \infty} \|u(t) - u^*(t)\| \leq \frac{\ell_w}{\nu^2} \limsup_{t \rightarrow \infty} \|\dot{w}(t)\|$$

## Example #9: Projected gradient controller

### Constrained feedback optimization:

$$\begin{aligned} \min_u \quad & \mathcal{E}(u, w) = \phi(u) + \psi(Y_u u + Y_w w) \quad (\nu \text{ strongly convex}, \ell_u \text{ strongly smooth}, \ell_w) \\ \text{subj. to} \quad & u \in \mathcal{U} \quad (\text{nonempty, closed, convex. } P_{\mathcal{U}} = \text{orthogonal projection}) \end{aligned}$$

**Projected gradient controller** (example of proximal gradient dynamics):

$$\dot{u} = F_{\text{PGC}}(u, w) := -u + P_{\mathcal{U}}(u - \gamma \nabla_u \mathcal{E}(u, w))$$

**Equilibrium tracking for projected gradient controller** At  $\gamma = \frac{2}{\nu + \ell_u}$ ,

①  $\text{osLip}_u(F_{\text{PGC}}) \leq -c_{\text{PGC}} := -\frac{2\nu}{\nu + \ell_u}$  (contractivity prox gradient)

②  $\text{Lip}_w(F_{\text{PGC}}) = \ell_{\text{PGC}} := \frac{2}{\nu + \ell_u} \ell_w$

$$\limsup_{t \rightarrow \infty} \|u(t) - u^*(t)\| \leq \frac{\ell_{\text{PGC}}}{c_{\text{PGC}}^2} \limsup_{t \rightarrow \infty} \|\dot{w}(t)\| \quad (\text{eq tracking})$$

§1. Introduction

§2. Basic contractivity concepts

§3. Gradient systems defined by strongly convex functions are contracting

§4. Equilibrium tracking in parametric optimization

§5. Application: Online feedback optimization

§6. Conclusions

**contractivity = robust computationally-friendly stability**

fixed point theory + Lyapunov stability theory + geometry of metric spaces

- theory
- examples
- control application

## Ongoing work

- ① applications to ML and biologically-inspired neural networks
- ② applications to optimization-based control designs:  
model predictive control, control barrier functions, low-gain integral control
- ③ equilibrium tracking with noise  
applications to optimization-based control