Contraction Theory as Guiding Principles for Design of Engineering Systems

Alexander (Sasha) Davydov

Center for Control,
Dynamical Systems & Computation
University of California, Santa Barbara
davydovalexander.github.io

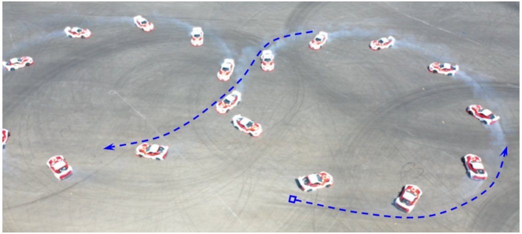
Semiautonomous Seminar @ UC Berkeley
October 18, 2024

Modern engineering systems


- Nonlinear control systems
- Complex environmental or inter-agent interactions
- Safety-critical tasks

Case study: Autonomous driving at the limits of handling

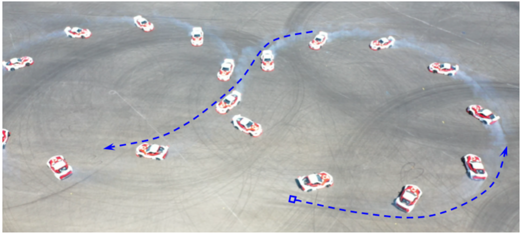
- Complex engineering system
  - Uncertain tire friction forces
  - Limited control authority
  - ...
- Unstable maneuver
  - Small errors can cause crashes
- Challenging problem for classical control methods



F. Djeumou, T. Lew, N. Ding, M. Thompson, M. Suminaka, M. Greiff, and J. Subosits. One model to drift them all: Physics-informed conditional diffusion model for driving at the limits. In *Conference on Robot Learning*, 2024. URL **https://openreview.net/forum?id=0gDbaEtVrd**

Desiderata

- **Highly-ordered** transient and asymptotic response
- **Robustness** to model error
- **Robustness** to environmental noise, time-delays, etc.
- From continuous-time to discrete-time
- Tractable methods for analysis
  - Amenable to methods based upon convex optimization
- Modularity properties
- ...

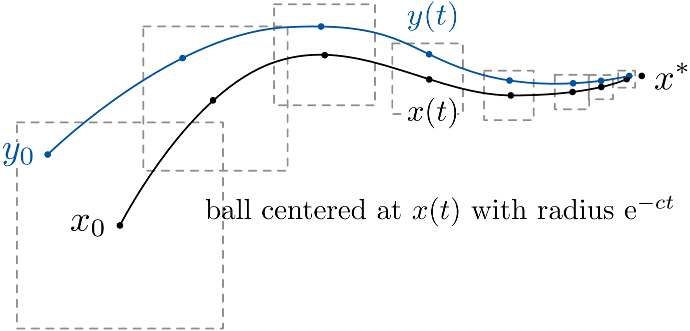


Outline

1. Contraction theory and properties
2. Optimization algorithms and optimization-based control
3. Imitation learning
4. Conclusions and opinion paper

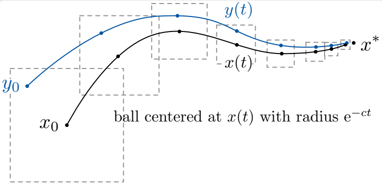
Contraction theory: Definition

Given $\dot{x} = \mathsf{F}(t, x)$, vector field F is contractive if its flow is a contraction map



W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6):683–696, 1998.

Properties of contracting dynamical systems



Highly ordered **transient** and **asymptotic** behavior:

1. time-invariant F: unique globally exponential stable equilibrium
   two natural Lyapunov functions
2. contractivity rate is natural measure/indicator of robust stability
   exponential incremental ISS
3. entrainment to periodic inputs
4. modularity and interconnection properties,
5. ...

In weighted Euclidean norms, contractivity $\iff$ existence $c > 0$ and $P = P^\top \succ 0$

$$PD\mathsf{F}(t,x) + D\mathsf{F}(t,x)^\top P \preceq -2cP, \qquad \forall x, t$$

A. Davydov and F. Bullo. Perspectives on contractivity in control, optimization and learning. *IEEE Control Systems Letters*, 8:2087–2098, 2024.

Example contracting systems

1. *optimization algorithms* under strong convexity assumptions
   (primal-dual, distributed, saddle, pseudo, proximal, etc)
2. *data-driven learned models* under certain parametrizations
   (stable dynamics learning, imitation learning, etc)
3. *neural network dynamics* under assumptions on weight matrices
   (recurrent, implicit, reservoir computing, etc)

4. stable linear systems
5. Lur'e-type systems under assumptions on nonlinearity and LMI conditions
6. feedback linearizable systems with stabilizing controllers
7. incremental ISS systems
8. nonlinear systems with a locally exponentially stable equilibrium

are contracting with respect to appropriate Riemannian metric

Continuous-time dynamics and one-sided Lipschitz constants

$\dot{x} = \mathsf{F}(x)$ on $\mathbb{R}^n$ with norm $\|\cdot\|$ and induced log norm $\mu(A) = \lim_{h\to 0^+} \frac{\|I_n + hA\| - 1}{h}$

One-sided Lipschitz constant (≈ maximum expansion rate)

$$\mathsf{osLip}(\mathsf{F}) = \sup_x \mu(D\mathsf{F}(x)), \qquad \mathsf{Lip}(\mathsf{F}) = \sup_x \|D\mathsf{F}(x)\|$$

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

For **affine map** $\mathsf{F}_A(x) = Ax + a$

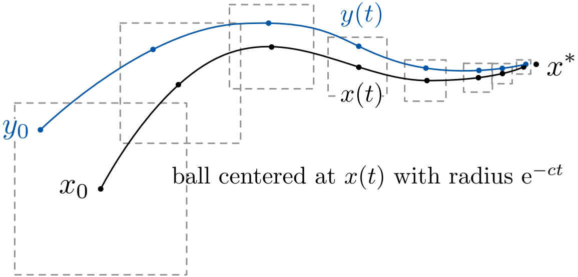
$$\mathsf{osLip}_{2,P}(\mathsf{F}_A) = \mu_{2,P}(A) \leq \ell \iff A^\top P + AP \preceq 2\ell P$$

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

Banach contraction theorem for continuous-time dynamics:

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

1. F is **infinitesimally contracting:** $\quad \|x(t) - y(t)\| \;\leq\; \mathrm{e}^{-ct}\|x_0 - y_0\|$
2. F has a unique, glob exp stable equilibrium x^*
3. global Lyapunov functions $V_1(x) = \|x - x^*\|^2$ and $V_2(x) = \|\mathsf{F}(x)\|^2$



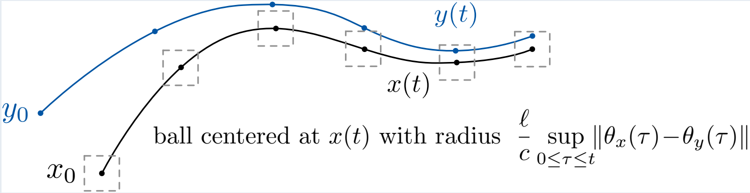
Property #1: Incremental ISS Theorem. Consider

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

- **contractivity wrt x:** $\mathrm{osLip}_x(\mathsf{F}) \leq -c < 0$
- **Lipschitz wrt θ:** $\mathrm{Lip}_\theta(\mathsf{F}) \leq \ell$

Then **incremental ISS property**:

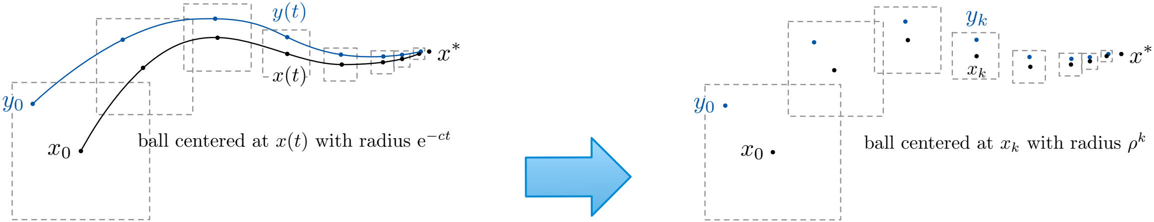
$$\|x(t) - y(t)\| \;\leq\; \mathrm{e}^{-ct}\|x_0 - y_0\| \;+\; \frac{\ell}{c} \sup_{\tau\in[0,t]} \|\theta_x(\tau) - \theta_y(\tau)\|$$



Property #2: Euler Discretization Theorem for Contracting Dynamics

Given norm $\|\cdot\|$ and Lipschitz $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$, equivalent statements

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



A. Davydov, S. Jafarpour, A. V. Proskurnikov, and F. Bullo. Non-Euclidean monotone operator theory and applications. *Journal of Machine Learning Research*, June 2023b. . To appear

Outline

Parametric and time-varying convex optimization

$$\min \mathcal{E}(x) \qquad \Longleftrightarrow \qquad \dot{x} = \mathsf{F}(x) \qquad \rightsquigarrow \qquad x^\star$$

Parametric and time-varying convex optimization

1. **parametric contracting dynamics for parametric convex optimization**

$$\min \mathcal{E}(x,\theta) \qquad \Longleftrightarrow \qquad \dot{x} = \mathsf{F}(x,\theta) \qquad \rightsquigarrow \qquad x^\star(\theta)$$

2. **contracting dynamics for time-varying strongly-convex optimization**

$$\min \mathcal{E}\big(x,\theta(t)\big) \qquad \Longleftrightarrow \qquad \dot{x} = \mathsf{F}\big(x,\theta(t)\big) \qquad \rightsquigarrow \qquad x^\star\big(\theta(t)\big)$$

A. Davydov, V. Centorrino, A. Gokhale, G. Russo, and F. Bullo. Time-varying convex optimization: A contraction and equilibrium tracking approach. *IEEE Transactions on Automatic Control*, June 2023a. doi. Conditionally Accepted

From convex optimization to contracting dynamics – Time-varying

Many convex optimization problems can be solved with contracting dynamics

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

	Convex Optimization	Contracting Dynamics
Unconstrained	$\min_{x\in\mathbb{R}^n} \quad f(x,\theta)$	$\dot{x} = -\nabla_x f(x,\theta)$
Constrained	$\min_{x\in\mathbb{R}^n} \quad f(x,\theta)$ s.t. $\quad x \in \mathcal{X}(\theta)$	$\dot{x} = -x + \mathrm{Proj}_{\mathcal{X}(\theta)}(x - \gamma\nabla_x f(x,\theta))$
Composite	$\min_{x\in\mathbb{R}^n} \quad f(x,\theta) + g(x,\theta)$	$\dot{x} = -x + \mathrm{prox}_{\gamma g_\theta}(x - \gamma\nabla_x f(x,\theta))$
Equality	$\min_{x\in\mathbb{R}^n} \quad f(x,\theta)$ s.t. $\quad Ax = b(\theta)$	$\dot{x} = -\nabla_x f(x,\theta) - A^\top\lambda,$ $\dot{\lambda} = Ax - b(\theta)$
Inequality	$\min_{x\in\mathbb{R}^n} \quad f(x,\theta)$ s.t. $\quad Ax \leq b(\theta)$	$\dot{x} = -\nabla_x f(x,\theta) - A^\top\nabla M_{\gamma,b(\theta)}(Ax+\gamma\lambda),$ $\dot{\lambda} = \gamma(-\lambda + \nabla M_{\gamma,b(\theta)}(Ax+\gamma\lambda))$

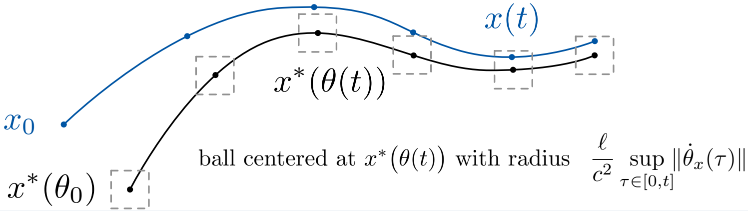
Property #3: Equilibrium Tracking Theorem. Consider

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

- **contractivity wrt x**: $\mathsf{osLip}_x(\mathsf{F}) \leq -c < 0$
- **Lipschitz wrt θ**: $\mathsf{Lip}_\theta(\mathsf{F}) \leq \ell$

Then **equilibrium tracking property**:

$$\|x(t) - x^\star(\theta(t))\| \;\leq\; \mathrm{e}^{-ct}\|x_0 - x^\star(\theta_0)\| \;+\; \frac{\ell}{c^2} \sup_{\tau\in[0,t]} \|\dot{\theta}(\tau)\|$$



Property #4: Exact Tracking Theorem. Consider

$$\dot{x}(t) = \mathsf{F}(x(t), \theta(t)) - (D_x\mathsf{F}(x(t), \theta(t)))^{-1} D_\theta \mathsf{F}(x(t), \theta(t))\dot{\theta}(t)$$

- **contractivity wrt x**: $\mathrm{osLip}_x(\mathsf{F}) \leq -c < 0$
- **smoothness:** $D_x\mathsf{F}$ and $D_\theta\mathsf{F}$ exist

Then **exact tracking property**:

$$\|\mathsf{F}(x(t), \theta(t))\| \leq e^{-ct}\|\mathsf{F}(x(0), \theta(0))\| \quad \text{and} \quad \|x(t) - x^\star(\theta(t))\| \leq \frac{1}{c}e^{-ct}\|\mathsf{F}(x(0), \theta(0))\|$$

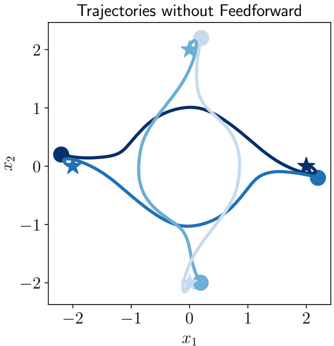
General way to design feedforward control for tracking in contracting dynamics

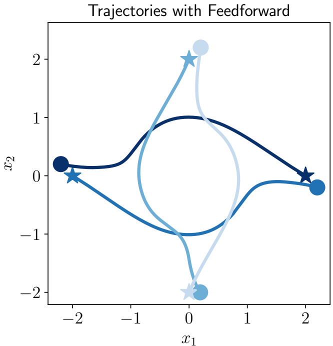
Application to safety filters

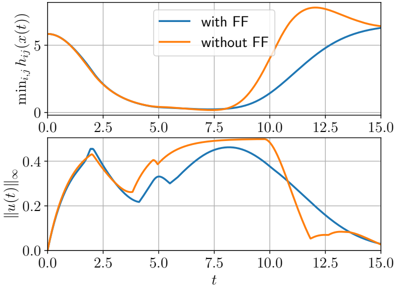
$$\dot{x} = F(x) + G(x)u^\star(x), \qquad \textbf{(Dynamics)}$$

$$\begin{aligned} u^\star(x) = \underset{u\in\mathbb{R}^m}{\operatorname{argmin}} \quad & \frac{1}{2}\|u - u_{\text{nom}}(x)\|_2^2, \qquad \textbf{(Safety filter)} \\ \text{s.t.} \quad & a_i(x)^\top u \le b_i(x), \quad i \in \{1,\dots,k\} \\ & \|u\|_\infty \le \overline{u}, \end{aligned}$$

Use contracting dynamics to solve **(Safety filter)** online







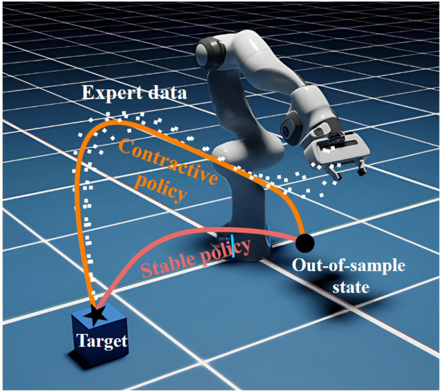
Outline

Imitation learning – Global performance guarantees

- Collection of expert demonstrations – $\mathcal{D} = \{(x_i, \dot{x}_i)_{i=1}^{N}\}$
- Learn dynamics $\dot{x} = \mathsf{F}(x)$
- Use low-level controller to track

Challenges:

- Ensure trajectories converge
- Robustness to uncertainty



Enforce contraction globally!

Anonymous. Contractive dynamical imitation policies for efficient out-of-sample recovery. *International Conference on Learning Representations (ICLR)*, 2025. URL **https://openreview.net/forum?id=lILEtkWOXD**. Under review

Learning contracting dynamics

Parametrization of contracting dynamics. Consider

$$\dot{x} = \mathsf{F}(x) = A(x, x^\star)(x - x^\star), \qquad \textbf{(ELCD)}$$

- **parametrization**:

$$A(x, x^\star) = -P_s(x, x^\star)^\top P_s(x, x^\star) + P_a(x, x^\star) - P_a(x, x^\star)^\top - cI_d$$

- P_s, P_a neural networks, $x^\star$ learnable equilibrium

Then **exp. convergence in ℓ_2 and contraction wrt Riemannian metric**:

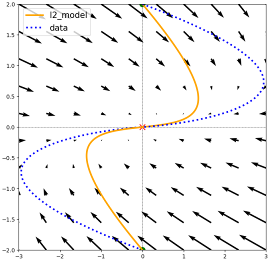
$$\|x(t) - x^\star\| \;\leq\; \mathrm{e}^{-ct}\|x(0) - x^\star\| \quad \text{and} \quad \|x_1(t) - x_2(t)\| \leq \kappa \mathrm{e}^{-ct}\|x_1(0) - x_2(0)\|$$

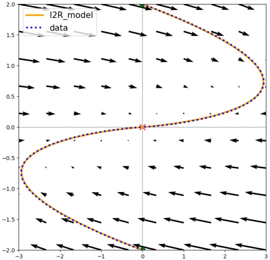
S. Jaffe, A. Davydov, D. Lapsekili, A. K. Singh, and F. Bullo. Learning neural contracting dynamics: Extended linearization and global guarantees. In *Advances in Neural Information Processing Systems*, 2024. doi. To appear

Bijective layers for more expressive dynamics

Parametrization alone cannot express all contracting dynamics

Idea: use parametrization in latent space and use diffeomorphism to map to data space





$$\dot{x} = A(x, x^\star)(x - x^\star), \qquad z = g(x)$$

Contraction is preserved under diffeomorphism

L. Dinh, J. Sohl-Dickstein, and S. Bengio. Density estimation using Real NVP. *International Conference on Learning Representations (ICLR)*, 2017. URL **https://arxiv.org/pdf/1605.08803.pdf**. arXiv preprint arXiv:1605.08803

Numerical results

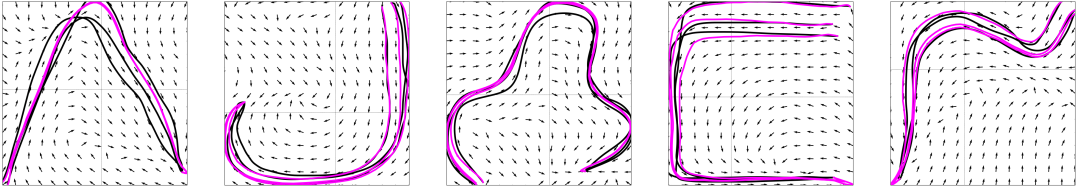


Table: DTWD on LASA, Pendulum, and Rosenbrock datasets

	SDD	EFlow	NCDS	ELCD
LASA-2D	0.37 ± 0.32	1.05 ± 0.25	0.59 ± 0.61	$\mathbf{0.12 \pm 0.11}$
LASA-4D	2.49 ± 2.4	2.24 ± 0.12	2.19 ± 1.23	$\mathbf{0.80 \pm 0.54}$
LASA-8D	5.26 ± 0.50	2.66 ± 0.63	5.04 ± 0.77	$\mathbf{1.52 \pm 0.61}$
Pendulum-4D	0.49 ± 0.11	0.17 ± 0.01	1.35 ± 2.26	$\mathbf{0.03 \pm 0.01}$
Pendulum-8D	0.75 ± 0.08	0.33 ± 0.01	2.88 ± 0.69	$\mathbf{0.14 \pm 0.03}$
Pendulum-16D	1.86 ± 0.14	0.45 ± 0.01	1.65 ± 0.31	$\mathbf{0.44 \pm 0.09}$
Rosenbrock-8D	NaN	1.90 ± 0.16	2.74 ± 0.15	$\mathbf{1.22 \pm 0.01}$
Rosenbrock-16D	NaN	3.57 ± 0.66	3.68 ± 0.12	$\mathbf{2.57 \pm 0.09}$

Outline

Conclusions

Summary:

1. motivation and overview of contraction theory
2. tracking-bounds for time-varying contracting systems
3. applications to imitation learning

Ongoing work and open problems:

1. applications on hardware
2. real-time closed-loop contraction controllers
3. learning closed-loop contracting systems
4. lots of flexibility in imitation learning

Opinion paper

- Overview of theory
- Recent applications
- Extensions
- Open problems

IEEE CONTROL SYSTEMS LETTERS, VOL. 8, 2024 2087

Perspectives on Contractivity in Control, Optimization, and Learning

Alexander Davydov, *Graduate Student Member, IEEE*, and Francesco Bullo, *Fellow, IEEE*

(Opinion Paper)

***Abstract*—Contraction theory is a mathematical framework for studying the convergence, robustness, and modularity properties of dynamical systems and algorithms. In this opinion paper, we provide five main opinions on the virtues of contraction theory. These opinions are (i) contraction theory is a unifying framework emerging from classical and modern works, (ii) contractivity is computationally-friendly, robust, and modular stability, (iii) numerous dynamical systems are contracting, (iv) contraction theory is relevant to modern applications, and (v) contraction theory can be vastly extended in numerous directions. We survey recent theoretical and applied research in each of these five directions.**

***Index Terms*—Contraction theory, incremental input-to-state stability, dynamical systems, neural networks.**

Introduction

A. Problem Description and Motivation

A DISCRETE-TIME dynamical system is *contracting* if its update map is a contraction in some metric. Analogously, a continuous-time system is contracting if its flow map is a contraction. Contraction theory for dynamical systems is a set of concepts and tools for the study and design of continuous and discrete-time dynamical systems. In this letter, we expose a few comprehensive opinions on this field and, by doing so, we review the basic theoretical foundations, the main computational and modularity properties, three main example dynamical systems, some modern applications, and extensions to local, weak, and Riemannian contraction.

Opinion #1: Contraction theory is a unifying framework emerging from classical and modern works. Contraction theory originates from the seminal work of Stefan Banach[1] in

systems can be traced back to the work of Lewis [62], Demidovich [35], and Krasovskiĭ [61]. Logarithmic norms and contraction for numerical integration of differential equations was studied in the seminal works [25], [64]. Later, logarithmic norms were applied to control problems by Desoer and Haneda [36] and Desoer and Vidyasagar [37]. The term "contraction analysis" was coined in the seminal work by Lohmiller and Slotine where they studied contraction with respect to Riemannian metrics [63].

It is now known that establishing contraction with respect to any norm has equivalent differential tests (i.e., conditions on the Jacobian of the vector field) and integral tests (i.e., conditions on the vector field itself) [30]. Before this unifying treatment, differential and integral conditions for contractivity have been discovered and rediscovered under different names. For example, focusing on integral conditions the following eight notions are either identical or very closely related:

1) one-sided Lipschitz maps in: [26] and [50] (Section I.10, Exercise 6)
2) uniformly decreasing maps in: [19]
3) no-name in: [43] (Chapter 1, page 5)
4) maps with negative nonlinear measure in: [75]
5) dissipative Lipschitz maps in: [16]
6) maps with negative lub log Lipschitz constant in: [84]
7) QUAD maps in: [65]
8) incremental quadratically stable maps in: [27]

In other words, despite its deep historical roots, contraction theory is still awaiting uniformization and broader appreciation. After decades of disparate work, a comprehensive framework is now emerging that clarifies the relationship among different strands of theoretical research.

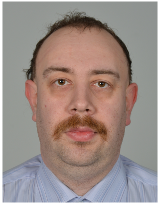
Acknowledgments

Veronica Centorrino
Scuola Sup Meridionale

Anand Gokhale
UCSB

Sean Jaffe
UCSB

Giovanni Russo
Univ Salerno



Anton Proskurnikov
Politecnico di Torino

Francesco Bullo
UCSB

A. Davydov, V. Centorrino, A. Gokhale, G. Russo, and F. Bullo. Time-varying convex optimization: A contraction and equilibrium tracking approach. *IEEE Transactions on Automatic Control*, June 2023a. doi. Conditionally Accepted

S. Jaffe, A. Davydov, D. Lapsekili, A. K. Singh, and F. Bullo. Learning neural contracting dynamics: Extended linearization and global guarantees. In *Advances in Neural Information Processing Systems*, 2024. doi. To appear