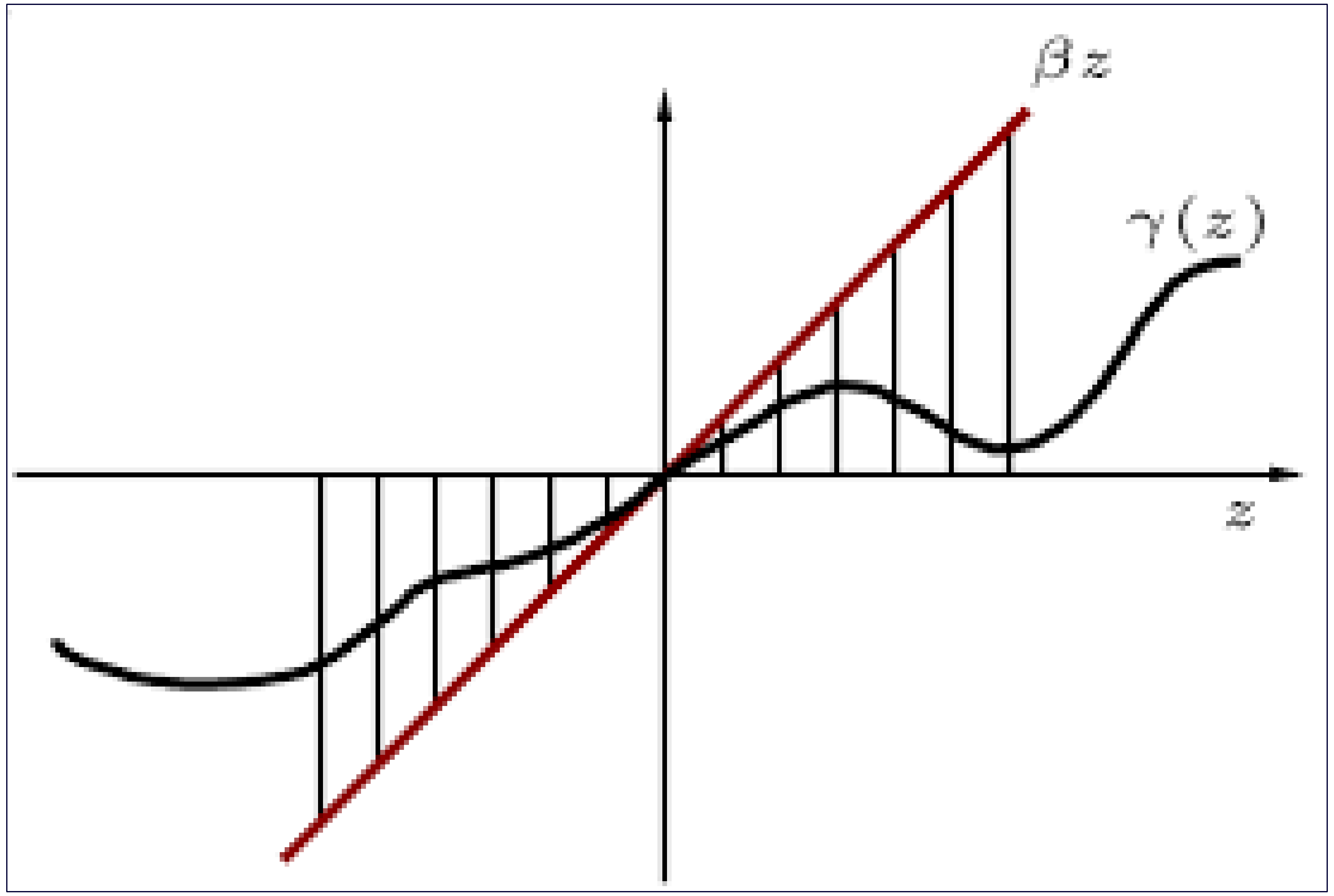


Towards Non-quadratic Absolute Stability Theory

Anton V. Proskurnikov

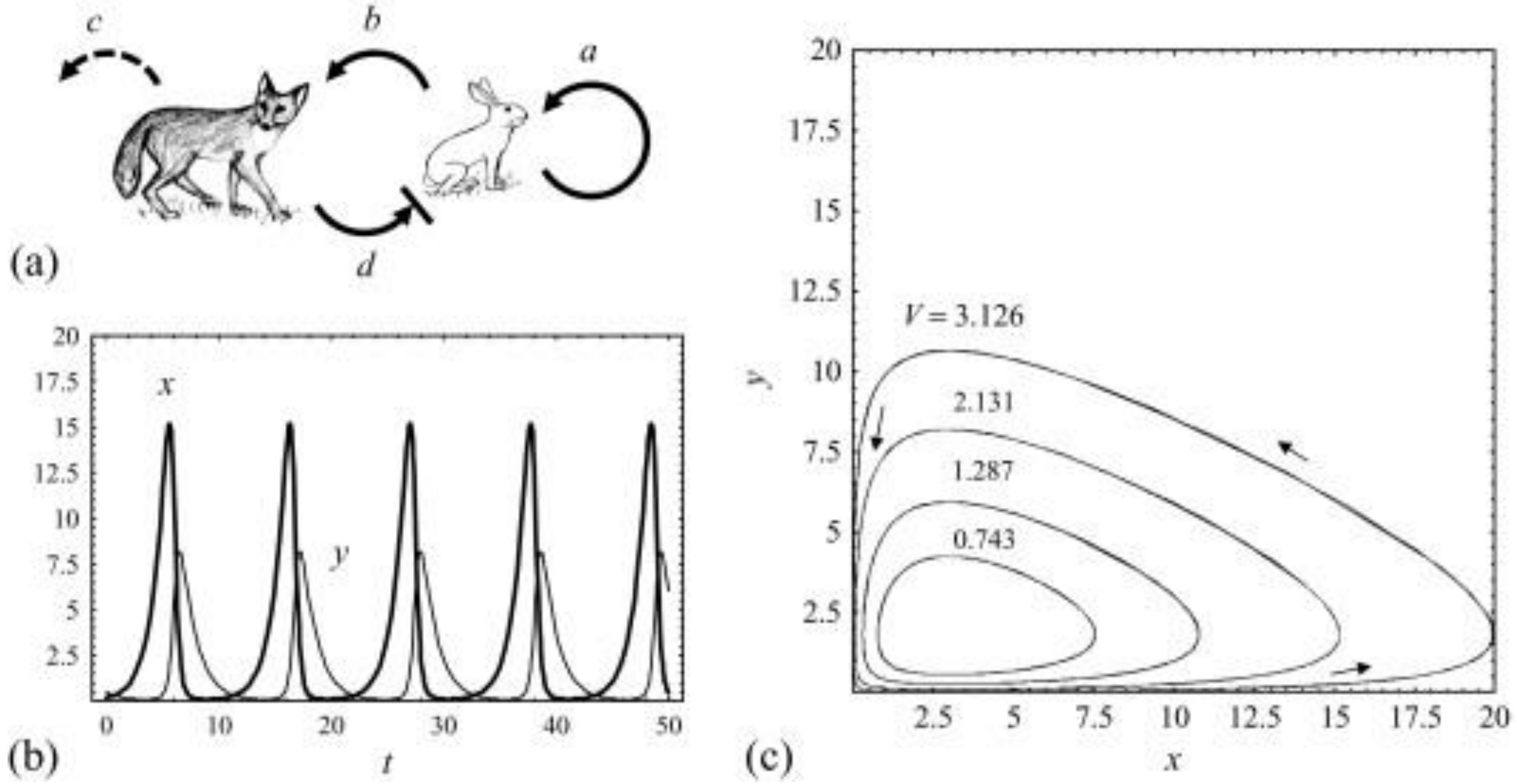
Politecnico di Torino, Italy

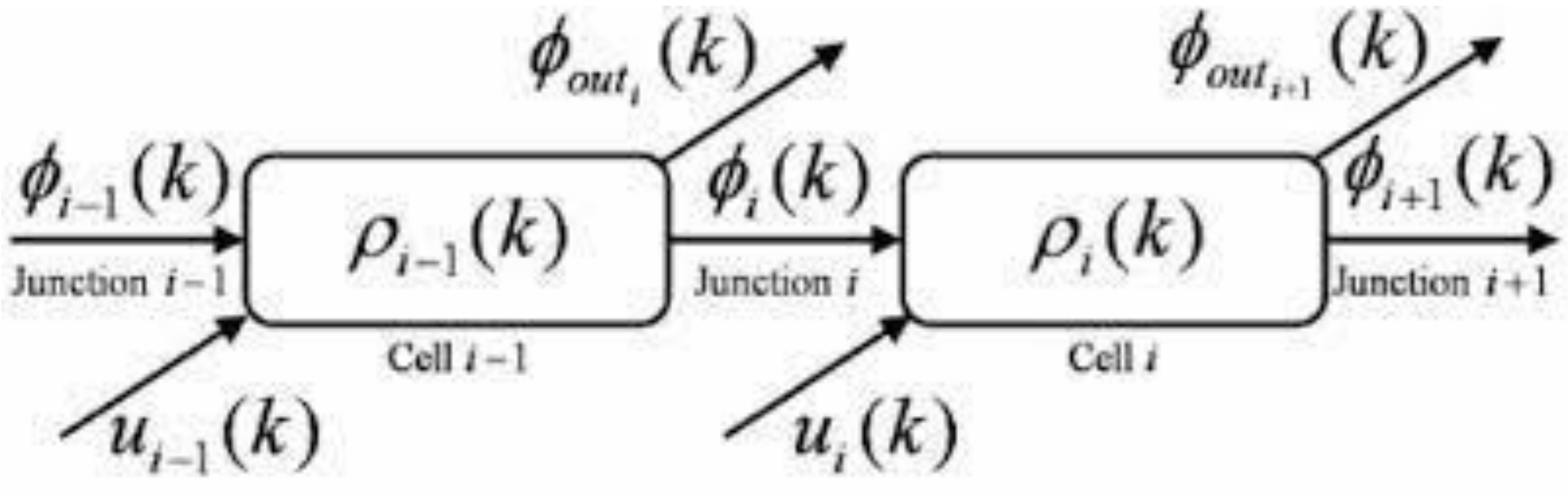
IEEE CDC 2024, Workshop on Contraction Theory for Systems, Control, Optimization, and Learning
December 15, 2024, Milan, Italy

Plan of The Talk

- **Non-quadratic Lyapunov functions: Motivation**
- **Lur’e systems and classical S-lemma**
- **General norms as Lyapunov functions: Extended S-lemma**
- **Absolute stability in non-Euclidean norms.**

Non-Quadratic Lyapunov Functions





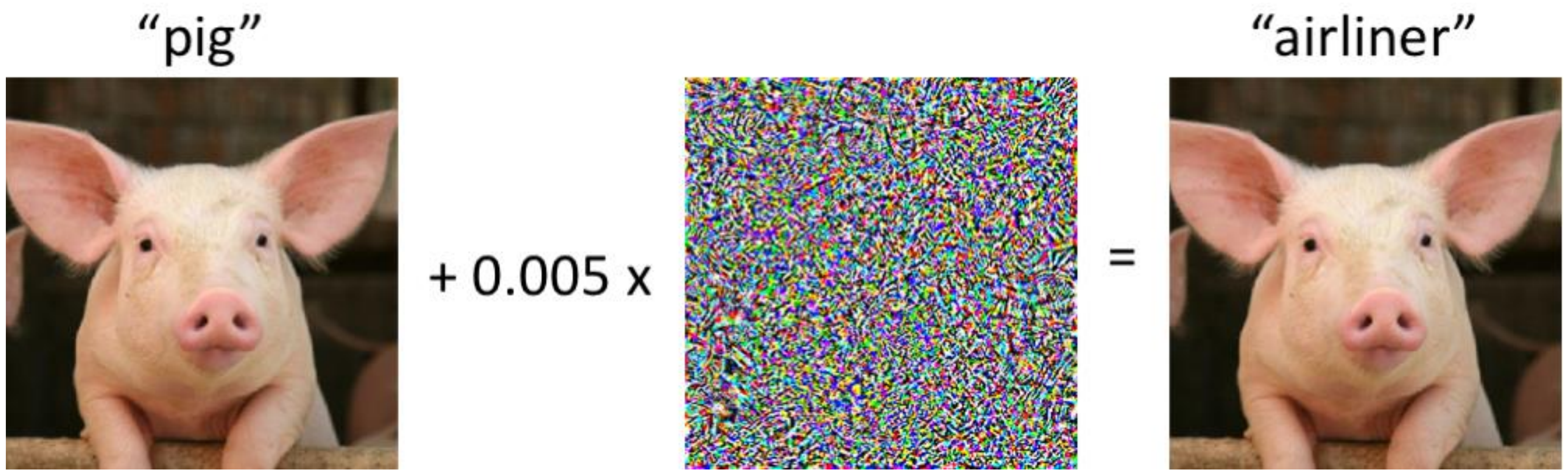
- **Non-quadratic LF** prove to be useful in many applications: Models of populations (Lotka-Volterra etc.), monotone systems, cell transmission models, compartmental systems, LPV systems etc. – **polyhedral and other LFs**
- Still underrepresented; regular design methods are still missing.
- This work: revise one of the oldest application of quadratic LF (Lur'e systems stability/contraction) and expand the design approach to some non-quadratic LF.

F. Bullo**, Contraction Theory for Dynamical Systems,** Kindle Direct Publishing, Edition 1.1, Mar 1, 2023

F. Bullo**, Lectures on Network Systems,** Kindle Direct Publishing, Edition 1.6, Jan 1, 2022

Contraction in ”Non-Standard” Norms: Resilience

Resilience and regularization of CNNs, RNNs, GNNs to adversarial attacks, efficient Lipschitz constants

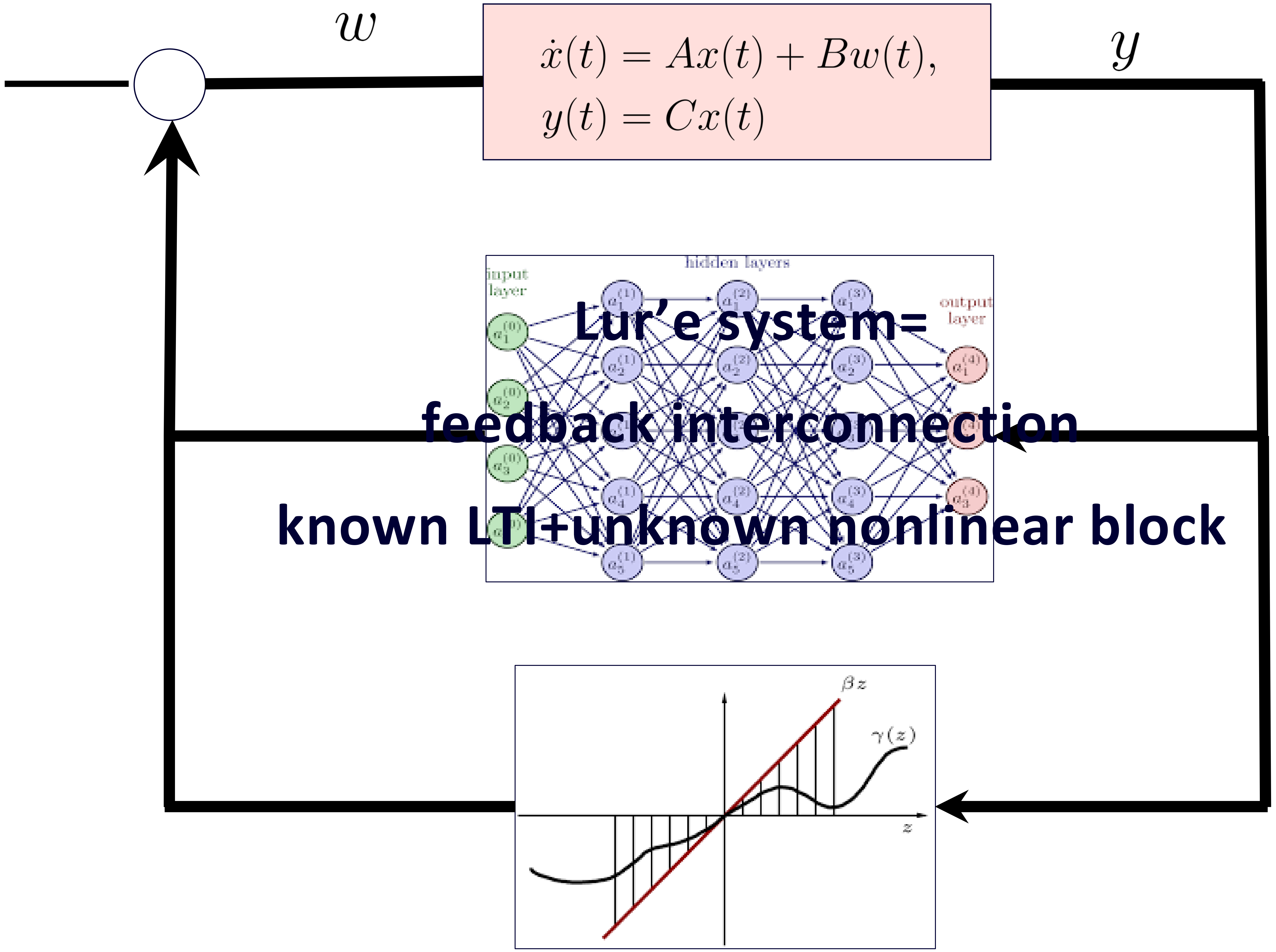


Slight perturbation of an image leads to incorrect classification: “adversarial examples”

- **Noises are naturally characterized in non-Euclidean norms (standard or weighted l_∞)**
- **Lipschitz constants are also mainly interesting in l_∞**
- **Computation of Lipschitz constants can be much more efficient than for the Euclidean norm by (LP vs. SDP). Algorithms to compute fixed-points of neural nets, tight convergence bounds in l_1 and l_∞**
- **Topological equivalence of norms on a space of a huge dimension is misleading: comparison factors are too large.**

Liang, Huang**, Large Norms of CNN Layers Do Not Hurt Adversarial Robustness//***2021,AAAI Conference on Artificial Intelligence,35(10),8565-8573.*
Davydov et al., **Non-Euclidean Contraction Analysis of Continuous-Time Neural Networks, IEEE TAC 2025 (online)**
Jafarpour et al., **Robust Implicit Networks via Non-Euclidean Contractions**//Adv. Neural Information Processing Systems 34, 2021

Lur'e Problem (1957): The Launchpad for Robust Nonlinear Control



Anatolii I. Lur'e (1901-1980)

- **Absolute stability** – **GAS of $x = 0$ for any nonlinear element from a given class**
- **Absolute contractivity – contractivity for any nonlinear element from a given class**

  Class – usually, very broad (sector, slope or other quadratic constraint)

Lur'e, Postnikov, Some Nonlinear Problems in the Theory of Automatic Control**, Her Majesty Stationary Office, 1957**

Absolute Stability: Two Main Tracks in 1960-1990s

Common Lyapunov stability certificates for the state space models

Lur'e, Barbashin, Persidskii, Kalman, Rosenwasser, Yakubovich, Barabanov, Narendra,…

$$V(x) = x^\top H x \qquad \left[+\int_0^y \varphi(s)ds \quad \text{(additional term, optional)}\right]$$

- **Lyapunov condition $\dot{V} < 0$:** quadratic constraints **+ S-Procedure**
- **Existence of the LF --** frequency-domain condition via KYP lemma (old) or **LMIs** (modern).
- **Depending on a problem, entails** GAS, contractivity, output stability, instability, oscillations.

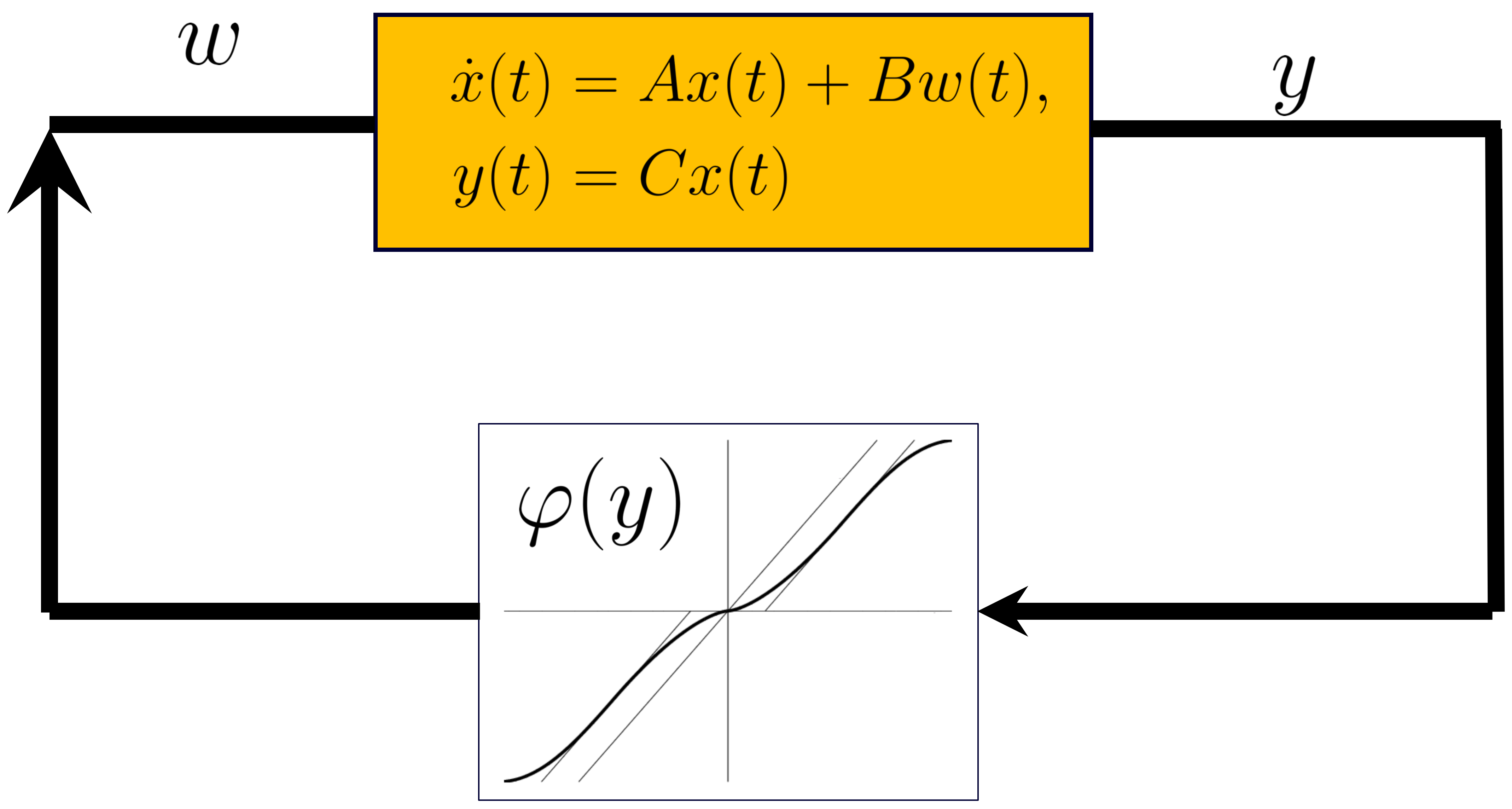
Integral operators in L2 space

Popov, Desoer, Vidyasagar, Yakubovich, Zames, Falb, Rasvan, Megretskii, Rantzer,…

General Volterra equations (also, infinite-dimensional)
Stability in integral L_2 norm, little information about transient process
Conditions in frequency domain: hard to verify in MIMO case

Vasile Mihai Popov
(born 1928)

S-Procedure: A Simple Example



$$0 \leq \frac{\varphi(y_1) - \varphi(y_2)}{y_1 - y_2} \leq \varkappa$$

S-procedure: implication (?) holds if

$$2(\Delta x)^\top H(A\Delta x + B\Delta w) + \varepsilon\|x\|_2^2 \leq \tau \Delta w(\varkappa^{-1}\Delta w - C\Delta x)$$

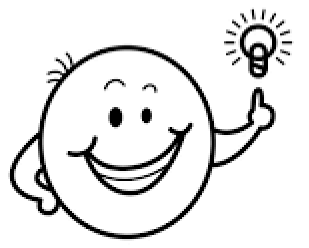
for some $\tau \geq 0$ and all variables

$$\Delta w \in \mathbb{R}, \Delta x \in \mathbb{R}^n.$$

It is an LMI on $H \succ 0, \tau \geq 0$

Try to prove contractivity using a quadratic (incremental) LF

$$V(\Delta x) = (\Delta x)^\top H(\Delta x) \Longrightarrow \dot{V} = 2(\Delta x)^\top H(A\Delta x + B\Delta w)$$



Lyapunov condition for contractivity: how to get it?

Slope inequality!

$$\dot{V} + \varepsilon\|\Delta x\|_2^2 \leq 0 \qquad \text{(A)}$$

(?) ⟵

$$\Delta w(\varkappa^{-1}\Delta w - \Delta y) \leq 0, \quad \Delta y = C\Delta x \qquad \text{(B)}$$

S-Lemma: S-Procedure is "Lossless".

Vladimir A. Yakubovich 1926-2012

The Yakubovich S-Lemma:

(?) holds for all variables

$$\Delta w \in \mathbb{R}, \Delta x \in \mathbb{R}^n$$

if and only if $\tau \geq 0$ exists satisfying the inequality

S-procedure finds **all** quadratic LFs helping to derive contraction from the single slope constraint.

S-procedure:

implication (?) holds if

$$2(\Delta x)^\top H(A\Delta x + B\Delta w) + \varepsilon\|x\|_2^2 \leq \tau \Delta w(\varkappa^{-1}\Delta w - C\Delta x)$$

for some $\tau \geq 0$ and all variables

$$\Delta w \in \mathbb{R}, \Delta x \in \mathbb{R}^n.$$

It is an LMI on $H \succ 0, \tau \geq 0.$

The idea of S-procedure is inspired by Lagrange multipliers.
There is an explicit relation between S-lemma and the Lagrange duality.

$$\dot{V} + \varepsilon\|\Delta x\|_2^2 \leq 0 \quad \overset{?}{\Longleftarrow} \quad \begin{aligned} \Delta w(\varkappa^{-1}\Delta w - \Delta y) &\leq 0, \\ \Delta y &= C\Delta x \end{aligned}$$

Further Developments of S-Lemma

- S.V. Gusev, A.L. Likhtarnikov, **Kalman-Popov-Yakubovich lemma and the S-procedure: A historical essay**, Automation and Remote Control, 67 (2006), pp. 1768–1810

- I. Polik and T. Terlaky, **A survey of the S-lemma**, SIAM Review, 49 (2007), pp. 371–418

$$x^\top Q_1 x \leq 0, \ldots, x^\top Q_n x \leq 0 \Longrightarrow x^\top P x \leq 0$$

- **Holds if there exist multipliers $\tau_1 \geq 0, \tau_2 \geq 0, \ldots, \tau_n \geq 0$ such that** $P \leq \sum_j \tau_j Q_j$

- **For usual quadratic forms, it is SDP feasibility. Can also be validated for integral functionals on L_2.**

- **The "only if" part is a rather exceptional property:**

a) n=1 (Yakubovich);
b) n=2 (Hermitian forms on a complex space (Yakubovich, Fradkov);
c) integral functionals on L_2 (Megretsky, Treil, Yakubovich, Matveev etc.) – n arbitrary.

Non-Quadratic Lyapunov Functions in Absolute Stability

- Special class of Non-quadratic LFs – squared norms $V(x) = \|x\|^2$
- General results on absolute stability of differential inclusions and switching systems (E.S. Pyatnitskii et al., Barabanov N.E.): existence of the non-Euclidean LF of this structure, approximations by polyhedral norms.

- Davydov et al. work - New insights about use of such LFs for contraction and stability

✓ **Derive the Lyapunov condition $\dot{V} < 0$ from "pseudo-quadratic" constraint + S-Lemma**

✓ **Existence of LF: conditions involving the log-norm ("matrix measure"). For each specific norm, convex optimization problem.**

✓ **Efficient stability/contractivity conditions in non-Euclidean norm, e.g., l_1 and l_∞ (important in machine learning, e.g., robustness certificates for NN, also for systems with conserved quantities, distributed iterations)**

Alexander Davydov, Saber Jafarpour, Francesco Bullo, Non-Euclidean Contraction Theory for Robust Nonlinear Stability, IEEE TAC 2022

Technical Prelimianries - I

Norm on a finite-dimensional space $\|\cdot\|$ → **Operator norm on matrices**

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

Compatible "weak pairing" (WP) -------------- **Log-norm ("measure") on matrices**

$$[x,x] = \|x\|^2, \forall x \qquad [\cdot,\cdot]$$

$$\mu(A) = \lim_{h\to 0+} \frac{\|I + hA\| - 1}{h}$$

a *weak pairing* on $\mathbb{R}^n$ is a map $[\![\cdot,\cdot]\!] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying

(i) (Subadditivity and continuity of first argument) $[\![x_1 + x_2, y]\!] \leq [\![x_1, y]\!] + [\![x_2, y]\!]$, for all $x_1, x_2, y \in \mathbb{R}^n$ and $[\![\cdot,\cdot]\!]$ is continuous in its first argument,

(ii) (Weak homogeneity) $[\![\alpha x, y]\!] = [\![x, \alpha y]\!] = \alpha\, [\![x, y]\!]$ and $[\![-x, -y]\!] = [\![x, y]\!]$, for all $x, y \in \mathbb{R}^n$, $\alpha \geq 0$,

(iii) (Positive definiteness) $[\![x, x]\!] > 0$, for all $x \neq 0_n$,

(iv) (Cauchy-Schwarz inequality) $|[\![x, y]\!]| \leq [\![x, x]\!]^{1/2} [\![y, y]\!]^{1/2}$, for all $x, y \in \mathbb{R}^n$.

Compatible WP is non-unique

Can always (and henceforth will be) chosen in such a way that

$$\mu(A) = \sup_{|x|=1} [Ax, x]$$

Derivative of the (differentiable) curve's norm

$$\frac{d}{dt}\|x(t)\|^2 = [\dot{x}(t), x(t)] \quad \text{a.e.}$$

Technical Prelimianries - II

Norm	Weak pairing	Log norms and Lumer's equality
$\|x\|_2 = \sqrt{x^\top x}$	$[\![x,y]\!]_2 = x^\top y$	$\mu_2(A) = \frac{1}{2}\lambda_{\max}(A + A^\top) = \max_{\|x\|_2=1} x^\top A x$
$\|x\|_p = \left(\sum_i \|x_i\|^p\right)^{1/p}$, $1 < p < \infty$	$[\![x,y]\!]_p = \|y\|_p^{2-p}(y \circ \|y\|^{p-2})^\top x$	$\mu_p(A) = \max_{\|x\|_p=1} (x \circ \|x\|^{p-2})^\top A x$
$\|x\|_1 = \sum_i \|x_i\|$	$[\![x,y]\!]_1 = \|y\|_1 \operatorname{sign}(y)^\top x$	$\mu_1(A) = \max_{j\in\{1,\dots,n\}} \Big(a_{jj} + \sum_{i\neq j} \|a_{ij}\|\Big) = \sup_{\|x\|_1=1} \operatorname{sign}(x)^\top A x$
$\|x\|_\infty = \max_i \|x_i\|$	$[\![x,y]\!]_\infty = \max_{i\in I_\infty(y)} x_i y_i$	$\mu_\infty(A) = \max_{i\in\{1,\dots,n\}} \Big(a_{ii} + \sum_{j\neq i} \|a_{ij}\|\Big) = \sup_{\|x\|_\infty=1} \max_{i\in I_\infty(x)} (Ax)_i x_i$

TABLE 1

Table of norms, weak pairings, and log norms for ℓ_2, ℓ_p for $p \in (1,\infty)$, ℓ_1, and ℓ_∞ norms. We adopt the shorthand $I_\infty(x) = \{i \in \{1,\dots,n\} \mid |x_i| = \|x\|_\infty\}$.

Easy to find: p=1,2, ∞

Quite difficult to find for a general p, leads to the non-convex optimization problem.

The operator p-norm computation is known to be NP hard (seems to be an open problem for log-norms).

“Pseudo-Quadratic” S-Lemma.

Suppose that one wishes to prove the following implication with “pseudo-quadratic” forms:

$$[Px, x] \leq 0 \quad \overset{?}{\Longleftarrow} \quad [Qx, x] \leq 0$$

<u>Lemma.</u> Let the WP [,] and the log-norm μ correspond to the same vector norm $\|\cdot\|$.

The implication (?) is valid for each vector *x* if a parameter $\tau \geq 0$ exists such that $\mu(P - \tau Q) \leq 0.$

<u>Proof</u>: Without loss of generality (since both inequalities are homogeneous), $\|x\| = 1.$

$$[Px, x] \leq [Px - \tau Qx, x] + [\tau Qx, x] \leq \mu(P - \tau Q) + \tau [Qx, x] \leq \tau [Qx, x].$$

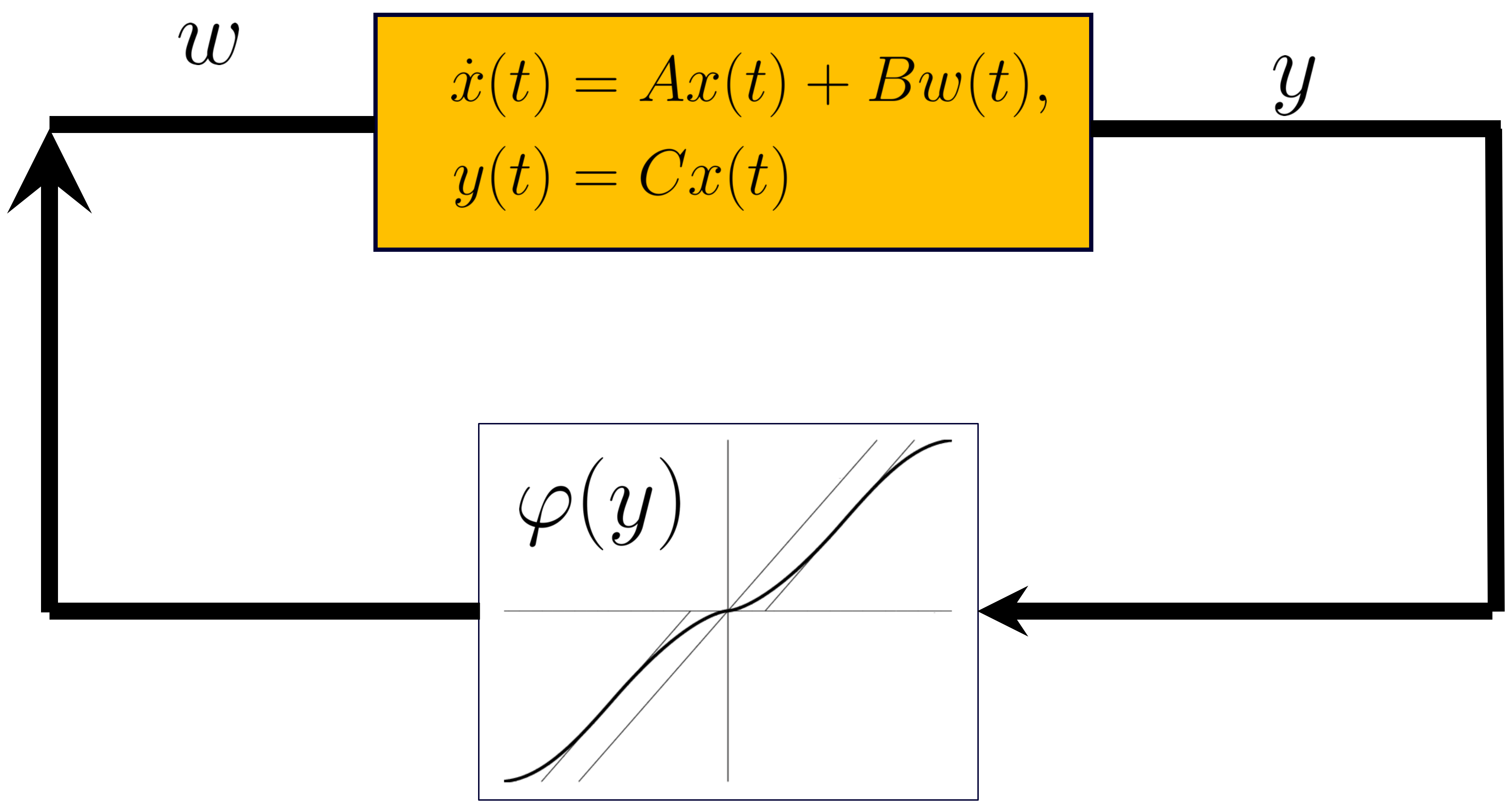
"Pseudo-Quadratic" S-Lemma: Some Remarks

$$x^\top Px = [Px, x] \leq 0 \quad \Longleftarrow \quad x^\top Qx = [Qx, x] \leq 0$$

$$\mu(P - \tau Q) \to \inf \quad \text{s.t. } \tau \geq 0$$

- **S-lemma in the general case reduces to optimization problem (the implication holds if inf <= 0)**
- **The problem is convex, because the log-norm is a convex function**
- **The oracle computing the function (and gradient, if needed) is of high complexity, except for special norms, e.g. l_1 , l_2, l_∞ .**
- **Can be trivially generalized to the case of multiple "pseudo-quadratic constraints" (several multipliers arise)**
- **Leads to new absolute stability and absolute contraction criteria in non-Euclidean norms.**

S-procedure with a Non-Euclidean Norm



$$0 \le \frac{\varphi(y_1) - \varphi(y_2)}{y_1 - y_2} \le \varkappa$$

$$V(\Delta x) = \|\Delta x\|_{p,R}^2 = \|R\Delta x\|_p^2, \ \det R \ne 0.$$

Program: repeat the trick with S-lemma.

- **Write the slope restriction as a "pseudo-quadratic" constraint**
- **Write the Lyapunov condition as a "pseudo-quadratic" constraint**
- **Apply S-lemma**
- **Principal difference: extended state-control vector needs to be used!**
- **For simplicity, let *R=I*.**

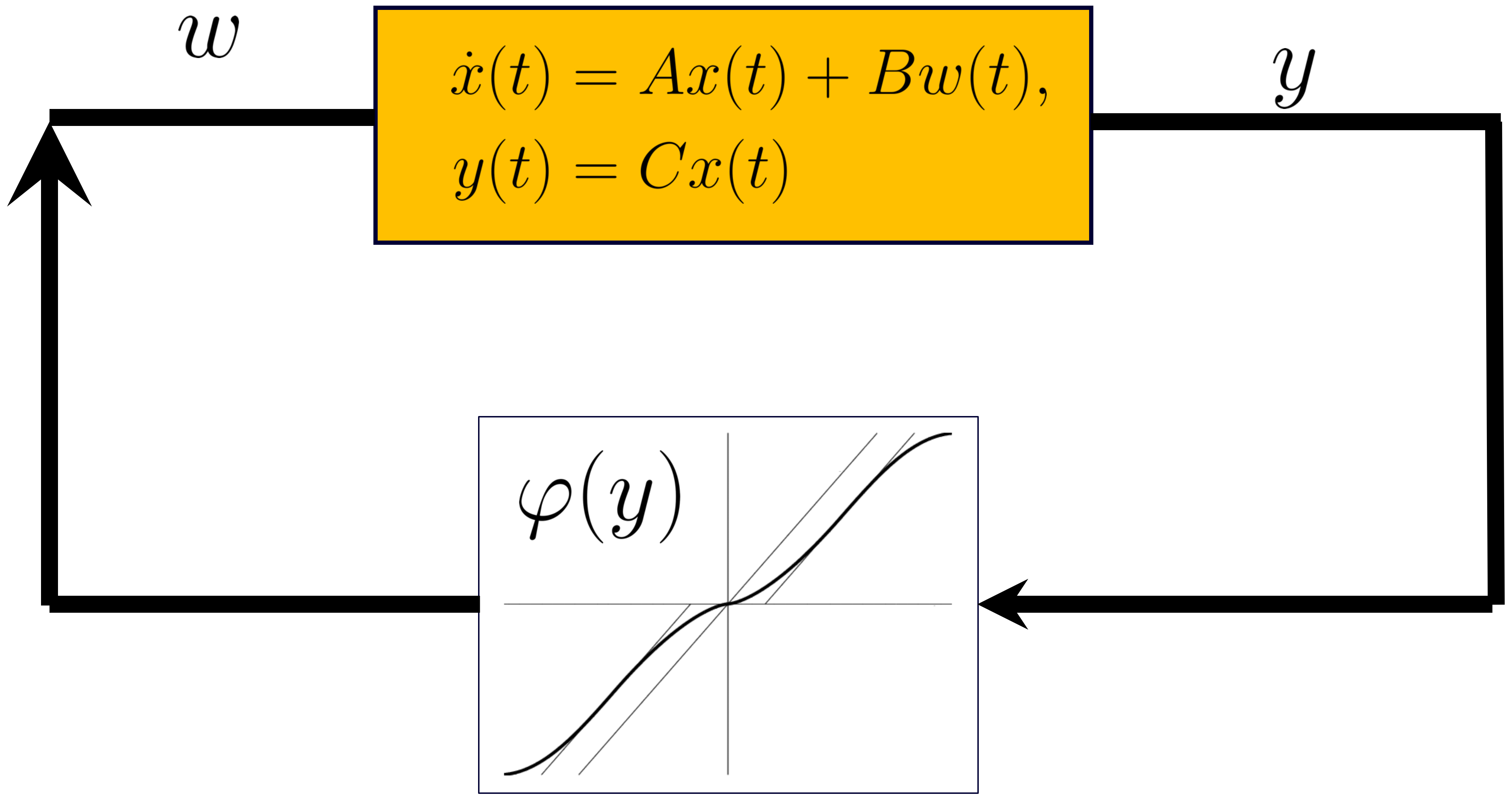
It remains to apply S-lemma!

$$\dot{V} \le -c\|\Delta x\|_p^2 \iff$$
$$\iff [A\Delta x + c\Delta x + B\Delta w, \Delta x]_p \le 0$$
$$\iff \left[\begin{pmatrix} A + cI & B \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix}, \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix}\right]_p \le 0$$

?

$$0 \le \frac{\Delta w}{\Delta y} \le \varkappa \iff$$
$$\left[\begin{pmatrix} 0 & 0 \\ -C & \varkappa^{-1} \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix}, \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix}\right]_p \le 0$$

Applying S-Lemma: Contraction in l_p Norm



$$0 \le \frac{\varphi(y_1) - \varphi(y_2)}{y_1 - y_2} \le \varkappa$$

Theorem.

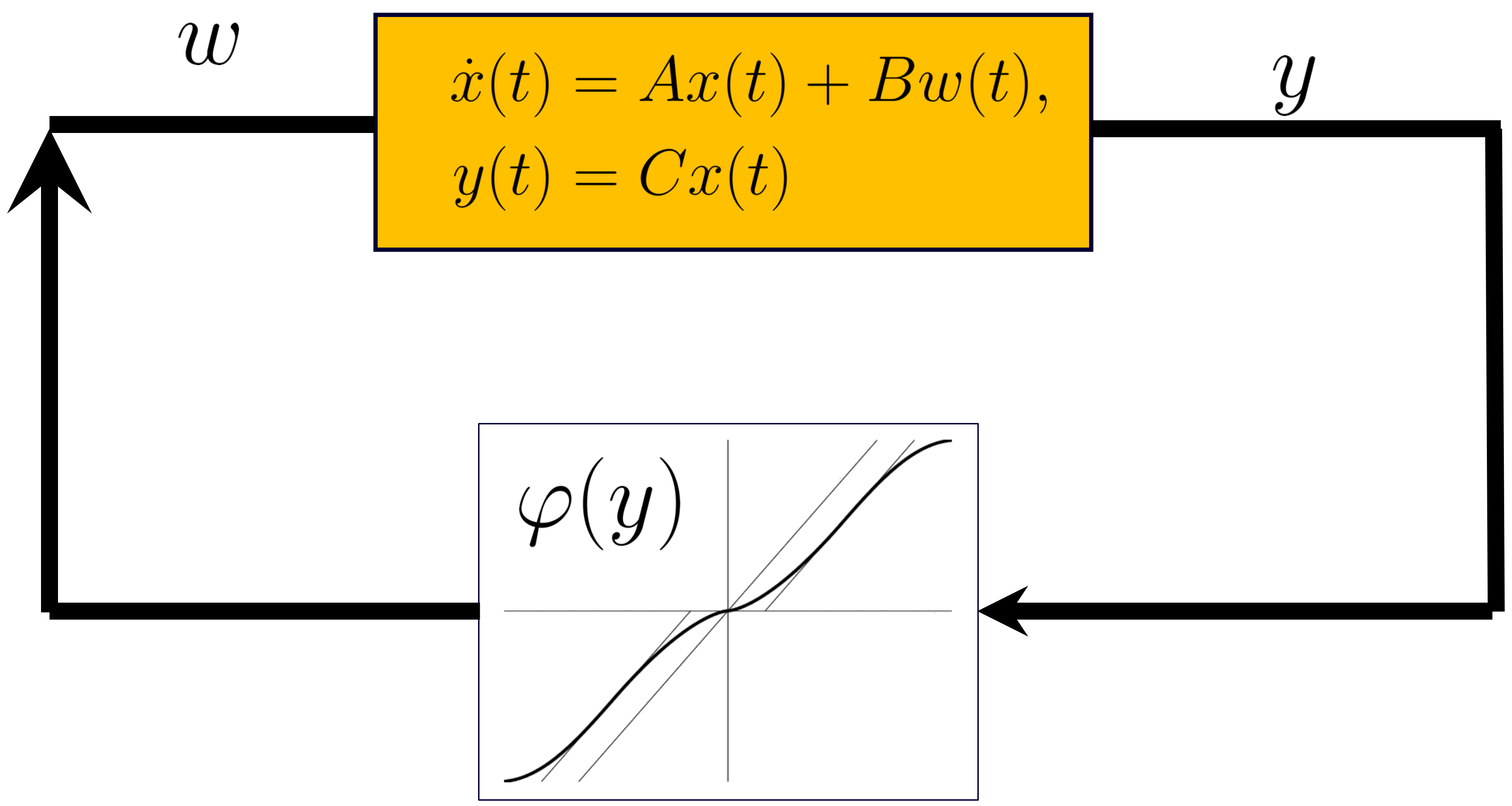
The system is contractive with respect to the norm $\|\Delta x\|_p$ and rate *c* if $\tau \geq 0$ exists such that

$$\mu_p \begin{pmatrix} A + cI & RB \\ \tau C & -\tau \varkappa^{-1} \end{pmatrix} \le 0$$

Remarks:

- This constraint is **convex!**
- However, the computation of the log-norm is troublesome (optimization with a **computationally expensive oracle**).
- Easily computable for:
  - **p=1 or ∞**
  - **p=2: classical LMIs.**

More General Result: Contraction in Weighted l_p Norm



$$0 \leq \frac{\varphi(y_1) - \varphi(y_2)}{y_1 - y_2} \leq \varkappa$$

Theorem. The system is contractive with respect to the (non-Euclidean) norm

$$\|\Delta x\|_{p,R} = \|R\Delta x\|_p$$

and rate *c* if *$\tau \geq 0$* exists such that

$$\mu_p \begin{pmatrix} RAR^{-1} + cI & RB \\ \tau CR^{-1} & -\tau\varkappa^{-1} \end{pmatrix} \leq 0$$

Remarks:

- For a **fixed** weight matrix, the condition is convex. For a variable weight, the constraint is non-convex!
- Sometimes, the problem reduces to **quasi-convex optimization (**Davydov, Proskurnikov, Bullo, IEEE TAC 2025**)**
  - p=1 or ∞ + diagonal weight matrix *R*
  - p=2: classical LMIs
- Example (in the journal paper): strong Kalman conjecture for positive systems. If $A, A - \varkappa BC$ are Hurwitz and Metzler, then the *R= R(p)* exists for all p ***(contraction is any norm is guaranteed).***

Conclusions and Future Works

Journal Paper: A.V. Proskurnikov, A. Davydov, F. Bullo**,** *The Yakubovich S-Lemma Revisited: Stability and Contractivity in Non-Euclidean Norms*, SICON 2023

- Many details omitted in this talk
- Some examples

Journal Paper: A.V. Proskurnikov, A. Davydov, F. Bullo**,** Non-Euclidean contraction analysis of continuous-time neural networks, TAC 2025

- Applications to contractivity and Lipschitz constant computation for RNNs
- Special norms: Weighted l_1 and l_∞

Topics for future research:

- **Efficient computational algorithms: how to validate log-norm criteria?**
- **Complicated nonlinearities:**
- **NNs, dynamical blocks, optimization solvers (control allocation logic, MPC etc.)**
- **Does non-quadratic norms work better?**

  Can we outperform classical stability criteria with non-quadratic norms?


thank you
MILANO
PANORAMICA
AGENZIA FOTOGRAFICA