On Weakly Contracting Dynamics for Convex Optimization

Veronica Centorrino[a], Alexander Davydov[b], Anand Gokhale[b],
Giovanni Russo[c] and Francesco Bullo[b]

Abstract

We investigate the convergence characteristics of dynamics that are *globally weakly* and *locally strongly contracting*. Such dynamics naturally arise in the context of convex optimization problems with a unique minimizer. We show that convergence to the equilibrium is *linear-exponential*, in the sense that the distance between each solution and the equilibrium is upper bounded by a function that first decreases linearly and then exponentially. The linear-exponential dependency arises naturally in certain dynamics with saturations. Additionally, we provide a sufficient condition for local input-to-state stability. Finally, we illustrate our results on, and propose a conjecture for, continuous-time dynamical systems solving linear programs.

I. Introduction

A paradigm that is becoming popular to analyze possibly time-varying optimization problems (OP) is to synthesize continuous-time dynamical systems that *converge* to an equilibrium that is also the optimal solution of the problem. A suitable tool to assess convergence is contraction theory [14], [3]. For OPs with strongly convex costs, the corresponding gradient dynamics, primal-dual dynamics (in the presence of constraints), or proximal gradient dynamics (for non-smooth costs) are strongly contracting, implying that trajectories converge to the equilibrium, which is also the optimal solution. In contrast, for OPs with only convex costs, the corresponding gradient, primal-dual, or proximal gradient dynamics are weakly contracting (or nonexpansive), and convergence depends on the existence of the minimizer.

In this context, we are interested in tackling convex OPs with a unique minimizer via continuous-time dynamical systems. This leads to the analysis of continuous-time dynamics that are globally-weakly contracting in the state space and only locally-strongly contracting (GW-LS-C). For such dynamics, we provide a comprehensive analysis, characterizing their convergence properties and local input-to-state stability (ISS). Finally, we consider linear programming (LP) to illustrate the effectiveness of our results.

Related literature: Studying optimization algorithms as continuous-time dynamical systems has been an active research area since [1], with, e.g., [17] being one of the first works to design neural networks for LPs. A neural network for solving LPs based on non-differentiable penalty functions has been proposed in [5]. Recently, there has been a renewed interest in continuous-time dynamics for optimization thanks to developments of, e.g., online and dynamic feedback optimization [2]. Additionally, OPs have been related to dynamical systems via proximal gradients, and the corresponding continuous-time proximal gradient dynamics are studied in, e.g., [10], [11]. Proximal gradients dynamics and contraction theory have been exploited in [8], [4] for tackling problems with strongly convex and only convex costs, respectively. In a broader context, there has been a growing interest in using strongly contracting dynamics to tackle OPs [15], [8]. This is mainly due to the fact that such dynamics enjoy highly ordered

[a]Veronica Centorrino is with Scuola Superiore Meridionale, Italy. veronica.centorrino@unina.it.

[b]Alexander Davydov, Anand Gokhale, and Francesco Bullo are with the Center for Control, Dynamical Systems, and Computation, UC Santa Barbara, Santa Barbara, CA 93106 USA. davydov@ucsb.edu, anand_gokhale@ucsb.edu, bullo@ucsb.edu.

[c]Giovanni Russo is with the Department of Information and Electric Engineering and Applied Mathematics, University of Salerno, Italy. giovarusso@unisa.it.

transient and asymptotic behaviors. Specifically, (i) initial conditions are exponentially forgotten and the distance between any two trajectories decays exponentially quickly [14], (ii) unique globally exponential stable equilibrium for time-invariant dynamics [14] (iii) entrainment to periodic inputs [16] (iv) highly robust behavior, such as ISS [18]. The asymptotic behavior of weakly contracting dynamics is instead characterized in, e.g., [7] for monotone systems and in [12] for primal-dynamics with a locally stable equilibrium.

Contributions: We analyze convergence of GW-LS-C dynamics, showing that this is *linear-exponential*, in the sense that the distance between each solution of the system and the equilibrium is upper bounded by a *linear-exponential function*, introduced in this paper. Through a novel technical result, we characterize the evolution of certain dynamics with saturation in terms of the linear-exponential function. This technical lemma is exploited for our convergence analysis, which is carried out considering two cases that require distinct mathematical approaches. First, we consider systems that are GW-LS-C with respect to (w.r.t.) the same norm. Then, we consider the case where the dynamics is GW-LS-C w.r.t. two different norms. Specifically, we give a convergence bound that, as discussed below, is sharper than the one in [4]. Additionally, we characterize local ISS for input-dependent dynamics that are GW-LS-C w.r.t. the same norm. Finally, we show the effectiveness of our results by considering a continuous-time dynamics tackling LPs and propose a general conjecture. The code to replicate our numerical example is given at https://shorturl.at/vGNY1.

While the treatment in this paper is inspired by the results in [4], we extend those results in several ways. First, when the dynamics is GW-LS-C w.r.t. two different norms, the bound we give here is continuous and always sharper than one given in [4], see Remark III.7. Second, when the dynamics is GW-LS-C w.r.t. the same norm, the technical lemma used to establish linear-exponential convergence is novel. Third and finally, local ISS was not considered in [4].

II. Mathematical Preliminaries

We denote by $\mathbb{0}_n \in \mathbb{R}^n$ the all-zeros vector of size n. Vector inequalities of the form $x \leq (\geq) y$ are entrywise. We let I_n be the $n \times n$ identity matrix. Given $A, B \in \mathbb{R}^{n\times n}$ symmetric, we write $A \preceq B$ (resp. $A \prec B$) if $B - A$ is positive semidefinite (resp. definite). We denote by $\lambda_{\max}(A)$ the maximum eigenvalue of A. We say that A is *Hurwitz* if $\alpha(A) := \max\{\mathrm{Re}(\lambda) \mid \lambda \text{ eigenvalue of } A\} < 0$, where $\mathrm{Re}(\lambda)$ denotes the real part of λ.

a) Norms, Logarithmic Norms and Weak Pairings

We let $\|\cdot\|$ denote both a norm on $\mathbb{R}^n$ and its corresponding induced matrix norm on $\mathbb{R}^{n\times n}$. Given $x \in \mathbb{R}^n$ and $r > 0$, we let $B_p(x, r) := \{z \in \mathbb{R}^n \mid \|z - x\|_p \leq r\}$ be the *ball of radius* r *centered at* x computed with respect to the norm p. Given two norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ on $\mathbb{R}^n$ there exist positive *equivalence coefficients* k_α^β and k_β^α satisfying $\|x\|_\alpha \leq k_\alpha^\beta \|x\|_\beta$, $\|x\|_\beta \leq k_\beta^\alpha \|x\|_\alpha$, $\forall x \in \mathbb{R}^n$. The *equivalence ratio between* $\|\cdot\|_\alpha$ *and* $\|\cdot\|_\beta$ is $k_{\alpha,\beta} := k_\alpha^\beta k_\beta^\alpha$, with k_α^β and k_β^α minimal equivalence coefficients.

Given $A \in \mathbb{R}^{n\times n}$ the *logarithmic norm* (log-norm) induced by $\|\cdot\|$ is

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

For an ℓ_p norm, $p \in [1, \infty]$, and for an invertible $Q \in \mathbb{R}^{n\times n}$, the Q-weighted ℓ_p norm is $\|x\|_{p,Q} := \|Qx\|_p$. The corresponding log-norm is $\mu_{p,Q}(A) = \mu_p(QAQ^{-1})$.

We let $[\![\cdot,\cdot]\!]$ denote a weak pairing on $\mathbb{R}^n$ compatible with the norm $\|\cdot\|$. We recall some of the main standing assumption on weak paring useful for our analysis.

Definition 1: A weak pairing $[\![\cdot,\cdot]\!]$, compatible with the norm $\|\cdot\|$, satisfies:

(i) *sub-additivity of first argument:* $[\![x + z, y]\!] \leq [\![x, y]\!] + [\![z, y]\!]$, for all $x, y, z \in \mathbb{R}^n$;

(ii) *curve norm derivative formula:* $\|y(t)\|D^+\|y(t)\| = [\![\dot{y}(t), y(t)]\!]$, for every differentiable curve $y\colon]a,b[\to \mathbb{R}^n$ and a.e. $t \in]a,b]$;

(iii) *Cauchy-Schwartz inequality:* $|[\![x,y]\!]| \le \|x\|\|y\|$, for all $x, y \in \mathbb{R}^n$;

(iv) *Lumer's equality:* $\mu(A) = \sup_{z\in\mathbb{R}^n, z\neq \mathbb{0}_n} \dfrac{[\![Az, z]\!]}{[\![z, z]\!]}$, for every $A \in \mathbb{R}^{n\times n}$.

We refer to [3] for a recent review of those tools.

b) Mathematical Operators

Given two normed spaces $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$, a map $T\colon \mathcal{X} \to \mathcal{Y}$ is Lipschitz with constant $L \geq 0$ if $\|T(x_1) - T(x_2)\|_{\mathcal{Y}} \le L\|x_1 - x_2\|_{\mathcal{X}}$, for all $x_1, x_2 \in \mathcal{X}$. The *upper-right Dini derivative* of a function $f\colon \mathbb{R} \to \mathbb{R}$ is $D^+f := \limsup_{h\to 0^+}\big(f(t+h) - f(t)\big)/h$. The *ceiling function*, $\lceil\ \rceil\colon \mathbb{R} \to \mathbb{Z}$, is defined by $\lceil x \rceil = \min\{y \in \mathbb{Z} \mid x \le y\}$. Given $d > 0$, the *saturation function*, $\mathrm{sat_d}\colon \mathbb{R} \to [-d, d]$, is defined by $\mathrm{sat}_d(x) = x$ if $|x| \le d$, $\mathrm{sat}_d(x) = d$ if $x > d$, and $\mathrm{sat}_d(x) = -d$ if $x < -d$. Given a set $\mathcal{C}$, the function $\iota_{\mathcal{C}}\colon \mathbb{R}^n \to [0, +\infty]$ is the *zero-infinity indicator function on* $\mathcal{C}$ and is defined by $\iota_{\mathcal{C}}(x) = 0$ if $x \in \mathcal{C}$ and $\iota_{\mathcal{C}}(x) = +\infty$ otherwise. The function $\mathrm{ReLU}\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$, is defined by $\mathrm{ReLU}(x) = \max\{0, x\}$.

We recall the following [6], [9]:

Theorem II.1 (Mean value theorem for locally Lipschitz function): Let $\mathcal{C} \subseteq \mathbb{R}^n$ be open and convex, $f\colon \mathcal{C} \to \mathbb{R}^m$ locally Lipschitz. Then, a.e. $x, y \in \mathcal{C}$ it holds:

$$f(x) - f(y) = \left(\int_0^1 Df(y + s(x-y))ds\right)(x-y),$$

where the integral of a matrix is to be understood component wise.

Whenever it is clear from the context, we omit to specify the dependence of functions on time t.

A. Proximal Operator

Given $g\colon \mathbb{R}^n \to \overline{\mathbb{R}} := [-\infty, +\infty]$, the *epigraph* of g is the set $\mathrm{epi}(g) = \{(x,y) \in \mathbb{R}^{n+1} \mid g(x) \le y\}$. The map g is (i) *convex* if its epigraph is a convex set, (ii) *proper* if its value is never $-\infty$ and is finite somewhere, and (iii) *closed* if it is proper and its epigraph is a closed set.

The *proximal operator of* g *with parameter* $\gamma > 0$, $\mathrm{prox}_{\gamma g}\colon \mathbb{R}^n \to \mathbb{R}^n$, is defined by

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

the associated *Moreau envelope*, $M_{\gamma g}\colon \mathbb{R}^n \to \mathbb{R}$, and its gradient are given by:

$$M_{\gamma g}(x) = g(\mathrm{prox}_{\gamma g}(x)) + \frac{1}{2\gamma}\|x - \mathrm{prox}_{\gamma g}(x)\|_2^2, \tag{2}$$

$$\nabla M_{\gamma g}(x) = \frac{1}{\gamma}(x - \mathrm{prox}_{\gamma g}(x)). \tag{3}$$

The gradient of the Moreau envelope always exists and is Lipschitz on $(\mathbb{R}^n, \|\cdot\|_2)$ with constant $1/\gamma$.

Finally, we recall that given a convex set $\mathcal{C}$, the proximal operator of the zero-infinity indicator function on $\mathcal{C}$ is the Euclidean projection onto $\mathcal{C}$, that is $\mathrm{prox}_{\gamma\iota_{\mathcal{C}}}(x) = \mathbb{P}_{\mathcal{C}}(x) := \operatorname*{argmin}_{z\in\mathcal{C}} \|x - z\|_2 \in \mathcal{C}$.

B. Contraction Theory for Dynamical Systems

Consider a dynamical system

$$\dot{x}(t) = f\big(t, x(t)\big), \tag{4}$$

where $f\colon \mathbb{R}_{\geq 0} \times \mathcal{C} \to \mathbb{R}^n$, is a smooth nonlinear function with $\mathcal{C} \subseteq \mathbb{R}^n$ forward invariant set for the dynamics. We let $t \mapsto \phi_t(x_0)$ be the flow map of (4) at time t starting from initial condition $x(0) := x_0$. Then, we give the following:

Definition 2 (Contracting systems): Given a norm $\|\cdot\|$ with associated log-norm μ, a smooth function $f\colon \mathbb{R}_{\geq 0} \times \mathcal{C} \to \mathbb{R}^n$, with $\mathcal{C} \subseteq \mathbb{R}^n$ f-invariant, open and convex, and a constant $c > 0$ ($c = 0$) referred as *contraction rate*, f is c-strongly (weakly) infinitesimally contracting on $\mathcal{C}$ if

$$\mu\big(Df(t, x)\big) \leq -c, \quad \text{for all } x \in \mathcal{C} \text{ and } t \in \mathbb{R}_{\geq 0}, \tag{5}$$

where $Df(t, x) := \partial f(t, x)/\partial x$.

If f is contracting, then for any two trajectories $x(\cdot)$ and $y(\cdot)$ of (4) it holds that

$$\|\phi_t(x_0) - \phi_t(y_0)\| \leq \mathrm{e}^{-ct}\|x_0 - y_0\|, \qquad \text{for all } t \geq 0,$$

i.e., the distance between the two trajectories converges exponentially with rate c if f is c-strongly infinitesimally contracting, and never increases if f is weakly infinitesimally contracting.

In [9, Theorem 16] condition (5) is generalized for locally Lipschitz function, for which, by Rademacher's theorem, the Jacobian exists almost everywhere (a.e.) in $\mathcal{C}$. Specifically, if f is locally Lipschitz, then f is infinitesimally contracting on $\mathcal{C}$ if condition (5) holds a.e. $x \in \mathcal{C}$ and $t \in \mathbb{R}_{\geq 0}$.

III. Linear-Exponential Decay of Globally-Weakly and Locally-Strongly Contracting Systems

In this section, we conduct a comprehensive analysis of the convergence of GW-LS-C systems. First, we define the *linear-exponential function*, which plays a pivotal role in bounding the convergence behavior of such dynamics.

Definition 3 (Linear-exponential function): Given a *linear decay rate* $c_{\text{lin}} > 0$, an *intercept* $q > 0$, an *exponential decay rate* $c_{\text{exp}} > 0$, and a *linear-exponential crossing time* $t_{\text{c}} < q/c_{\text{lin}}$, the *linear-exponential function* $\text{lin-exp}(\cdot)\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is defined by

$$\text{lin-exp}(t) = \begin{cases} q - c_{\text{lin}}t & \text{if } t \leq t_{\text{c}}, \\ \big(q - c_{\text{lin}}t_{\text{c}}\big)\mathrm{e}^{-c_{\text{exp}}(t - t_{\text{c}})} & \text{if } t > t_{\text{c}}. \end{cases} \tag{6}$$

We write $\text{lin-exp}(t\ ;\ q, c_{\text{lin}}, c_{\text{exp}}, t_{\text{c}})$ when we want to highlight the parameters in (6). See Figure 1 for an illustration of (3) for some parameters.

Before giving the main convergence results of this section, we prove the following:

Lemma III.1 (Property of the linear-exponential function): Let c_{exp} and d be positive scalars. Consider the dynamics

$$\dot{x}(t) = -c_{\text{exp}}\,\text{sat}_d(x(t)), \quad x_0 = q > d. \tag{7}$$

Then, $x(t) = \text{lin-exp}(t\ ;\ q, c_{\text{lin}}, c_{\text{exp}}, t_{\text{c}})$, with $c_{\text{lin}} = dc_{\text{exp}}$ and $t_{\text{c}} := qc_{\text{lin}}^{-1} - c_{\text{exp}}^{-1} > 0$, is a solution of (7).

Proof: First, we note that being the right end side of (7) locally Lipschitz continuous, the ODE (7) admits a unique continuous solution at least within a certain neighborhood of the initial condition.

Using the definition of saturation function, for all $t \in \mathbb{R}_{\geq 0}$ we can write the ODE (7) as

$$\dot{x}(t) = \begin{cases} dc_{\text{exp}} & \text{if } x(t) < -d, \\ -c_{\text{exp}}x(t) & \text{if } x(t) \in [-d, d], \\ -dc_{\text{exp}} & \text{if } x(t) > d, \end{cases} \tag{8}$$

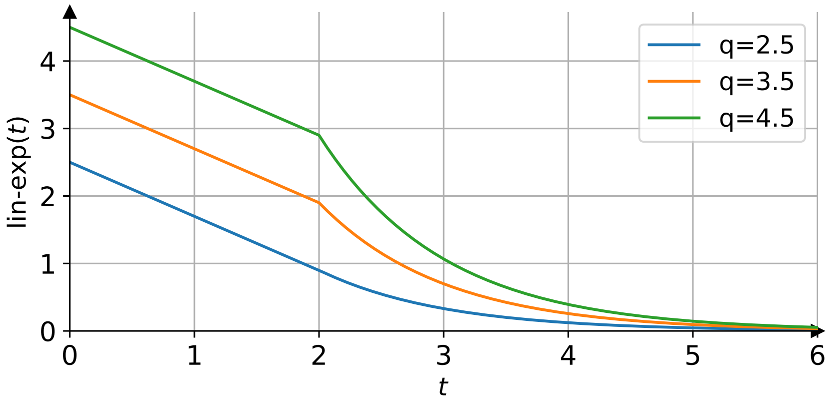

Figure 1: Plot of the linear-exponential function (6) with linear decay rate $c_{\text{lin}} = 0.8$, intercepts $q = \{2.5, 3.5, 4.5\}$, exponential decay rate $c_{\text{exp}} = 1$, and linear-exponential crossing time $t_{\text{c}} = 2$.

which, in each interval $[t_0, t_1] \subseteq \mathbb{R}_{\geq 0}$ where the solution is continuous and does not change regime, has general solution

$$x(t) = \begin{cases} dc_{\text{exp}}t + x(t_0) & \text{if } x(t) < -d, \\ x(t_0)e^{-c_{\text{exp}}(t-t_0)} & \text{if } x(t) \in [-d, d], \\ -dc_{\text{exp}}t + x(t_0) & \text{if } x(t) > d. \end{cases} \tag{9}$$

At time $t = 0$, we have $x_0 = q > d$. For continuity of the solution, there exists $t^\star$ such that $x(t) > d$ for all $t \in [0, t^\star]$. Thus from (9) and being $x(t_0 = 0) = q$, it is $x(t) = -dc_{\text{exp}}t + q$, for all $t \in [0, t^\star]$. Moreover being $x(t)$ a decreasing function, the time value $t^\star$ is finite and there exists a time, say it $\bar{t}$ such that $x(\bar{t}) = d$. Let $c_{\text{lin}} := dc_{\text{exp}}$, we have

$$x(\bar{t}) = d \iff -dc_{\text{exp}}\bar{t} + q = d \iff \bar{t} = qc_{\text{lin}}^{-1} - c_{\text{exp}}^{-1} := t_{\text{c}}.$$

In summary, we have shown that the solution of (8) is equal to $x(t) = q - c_{\text{lin}}t$ for all $t \in [0, t_{\text{c}}]$ and is equal to d at time t_{c}. Therefore from (9) and being $x(t_{\text{c}}) = q - c_{\text{lin}}t_{\text{c}}$, for all time $t > t_{\text{c}}$ we have $x(t) = \big(q - c_{\text{lin}}t_{\text{c}}\big)e^{-c_{\text{exp}}(t-t_{\text{c}})}$. Specifically, $x(t) > 0$ for all $t > t_{\text{c}}$, thus it can never be the case $x(t) < -d$. This concludes the proof. ■

Next, we study the convergence behavior of GW-LS-C systems of the form of (4), where the function $f\colon \mathbb{R}_{\geq 0} \times \mathcal{C} \to \mathbb{R}^n$ is locally Lipschitz and with $\mathcal{C} \subseteq \mathbb{R}^n$ being f-invariant, open and convex. In what follows, we make the following:

Assumption 1: There exist $\|\cdot\|_{\text{G}}, \|\cdot\|_{\text{L}}$ on $\mathbb{R}^n$ such that
(*A*1) f is weakly infinitesimally contracting on $\mathbb{R}^n$ w.r.t. $\|\cdot\|_{\text{G}}$;
(*A*2) f is c_{exp}-strongly infinitesimally contracting on a forward-invariant set $\mathcal{S}$ w.r.t. $\|\cdot\|_{\text{L}}$;
(*A*3) $x^\star \in \mathcal{S}$ is an equilibrium point, i.e., $f(t, x^\star) = \mathbb{0}_n$, for all $t \geq 0$.

Remark III.2: Assumptions (*A*2), (*A*3) can be equivalently replaced by assuming the existence of a locally exponentially stable equilibrium. ■

First, we consider GW-LS-C systems with respect to the same norm. Then, dynamics that are GW-LS-C with respect to different norms. In both scenarios, we show that the convergence is (globally) *linear-exponential*. That is, given a trajectory $x(t)$ of the dynamics, the distance $\|x(t) - x^\star\|_{\text{G}}$ is upper bounded by a linear-exponential function (6).

A. *Convergence of Globally-Weakly and Locally-Strongly Contracting Dynamics with Respect to the Same Norm*

We start by giving a bound on the upper right Dini derivative of the distance of any solution of (4) with respect to the equilibrium $x^\star$.

Lemma III.3 (Saturated error dynamics): Consider system (4) and let Assumptions (A1) – (A3) hold with $\|\cdot\|_{\mathrm{G}} = \|\cdot\|_{\mathrm{L}} := \|\cdot\|$. Let r be the largest radius such that $B\big(x^\star, r\big) \subseteq \mathcal{S}$. Then, for every trajectory $x(t)$ starting from $x_0 \notin \mathcal{S}$, a.e. $t \geq 0$, we have

$$D^+\|x(t) - x^\star\| \leq -c_{\exp}\,\mathrm{sat}_r(\|x(t) - x^\star\|). \tag{10}$$

Proof: Consider an arbitrary trajectory $x(t)$ starting from $x_0 \notin \mathcal{S}$ and a second trajectory equal to the equilibrium $x^\star$. Let μ be the log-norm associated to $\|\cdot\|$. For a.e. $t \geq 0$ it holds ([6], [9]):

$$D^+\|x(t) - x^\star\| \leq \int_0^1 \mu\Big(Df\big(t, x^\star + \alpha(x(t) - x^\star)\big)\Big) d\alpha \cdot \|x(t) - x^\star\| := RHS.$$

where $\alpha \in [0, 1]$, and $x^\star + \alpha(x(t) - x^\star)$ is the segment from $x^\star$ to $x(t)$.

For each $t \geq 0$, if $\|x(t) - x^\star\| \leq r$, then Assumption (A2) implies

$$RHS \leq \int_0^1 (-c_{\exp}) d\alpha \cdot \|x(t) - x^\star\| = -c_{\exp}\|x(t) - x^\star\| = -c_{\exp}\,\mathrm{sat}_r(\|x(t) - x^\star\|),$$

where in the last equality we have used the definition of saturation function. If $\|x(t) - x^\star\| \geq r$, define $\alpha^* = r/\|x(t) - x^\star\|$ and note that, a.e. $t \geq 0$, Assumptions (A2) and (A1) imply

$$\begin{aligned} \alpha \leq \alpha^* \quad &\Longrightarrow \quad \mu\big(Df(t, x^\star + \alpha(x(t) - x^\star))\big) \leq -c_{\exp}, \\ \alpha > \alpha^* \quad &\Longrightarrow \quad \mu\big(Df(t, x^\star + \alpha(x(t) - x^\star))\big) \leq 0. \end{aligned}$$

Therefore, a.e. $t \geq 0$, it holds

$$\begin{aligned} RHS &\leq \Big(\int_0^{\alpha^*} \mu\big(Df(t, x^\star + \alpha(x(t) - x^\star))\big) d\alpha + \int_{\alpha^*}^1 \mu\big(Df\big(t, x^\star + \alpha(x(t) - x^\star)\big)\big) d\alpha\Big) \cdot \|x(t) - x^\star\| \\ &\leq (-c_{\exp}\alpha^* + 0)\|x(t) - x^\star\| = -c_{\exp} r = -c_{\exp}\,\mathrm{sat}_r(\|x(t) - x^\star\|), \end{aligned} \tag{11}$$

where in the last equality we used the definition of saturation function. We illustrate this result about the average of the log-norm of the Jacobian in Figure 2. This concludes the proof. ■

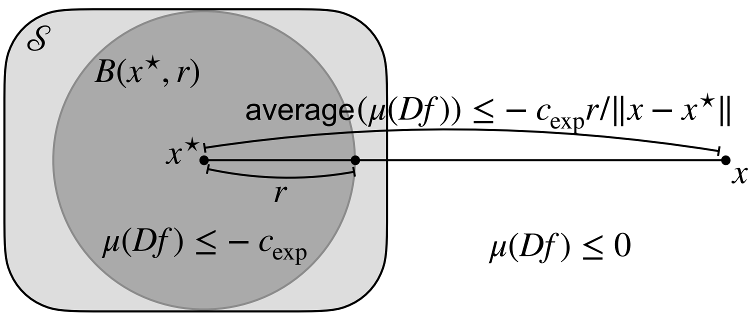

Figure 2: Illustration of the inequality (11).

We can now give our convergence result for GW-LS-C systems with respect to the same norm.

Theorem III.4 (Linear-exponential convergence of GW-LS-C systems w.r.t. the same norm): Consider system (4) and let Assumptions (A1) – (A3) hold with $\|\cdot\|_{\mathrm{G}} = \|\cdot\|_{\mathrm{L}} := \|\cdot\|$. Also, let r be the largest radius such that $B\big(x^\star, r\big) \subseteq \mathcal{S}$. For each trajectory $x(t)$ starting from x_0, it holds that

(i) if $x_0 \in \mathcal{S}$, then, a.e. $t \geq 0$,

$$\|x(t) - x^\star\| \leq \mathrm{e}^{-c_{\exp} t}\|x_0 - x^\star\|;$$

(ii) if $x_0 \notin \mathcal{S}$, then, a.e. $t \geq 0$,

$$\|x(t) - x^\star\| \leq \text{lin-exp}(t\ ;\ q, c_{\text{lin}}, c_{\exp}, t_{\text{c}}), \tag{12}$$

with

- exponential decay rate $c_{\exp} > 0$;
- linear decay rate $c_{\text{lin}} = c_{\exp}\, r$;
- intercept $q = \|x_0 - x^\star\|$;
- linear-exponential crossing time $t_{\text{c}} = (q - r)/c_{\text{lin}}$.

Proof: Item (i) follows from Assumption *(A2)*. Item (ii) follows by using the Comparison Lemma [13, pp. 102-103] to upper bound the solution to the differential inequality (10). Additionally, the upper bound obeys precisely the initial value (7) in Lemma III.1, for parameter values $d = r$, $c_{\text{lin}} = c_{\exp} r$, $q = \|x_0 - x^\star\|$, and $t_{\text{c}} = (q - r)/c_{\text{lin}}$. This concludes the proof. ■

B. *Convergence of Globally-Weakly and Locally-Strongly Contracting Dynamics with Respect to Different Norms*

We begin by introducing the *ρ-contraction time*, where $0 < \rho < 1$.

Definition 4 (ρ-contraction time): Let system (4) be strongly infinitesimally contracting with respect to a norm $\|\cdot\|_\alpha$. Consider the *contraction factor* $0 < \rho < 1$, a norm $\|\cdot\|_\beta$, and a vector $x \in \mathbb{R}^n$.

- The *ρ-contraction time* is the time required for each trajectory starting in $B_\alpha\big(x, r\big)$, for some $r > 0$, to be inside $B_\alpha\big(x, \rho r\big)$;
- The *ρ-contraction time with respect to* $\|\cdot\|_\beta$ is the time required for each trajectory starting in $B_\beta\big(x, r\big)$, for some $r > 0$, to be inside $B_\beta\big(x, \rho r\big)$.

Remark III.5: It is implicit in Definition 4 that the ρ-contraction time for a specific trajectory depends on the initial condition and the center of the ball. ■

We can now give our convergence result for GW-LS-C systems with respect to the different norms.

Theorem III.6 (Linear-exponential convergence of GW-LS-C systems): Let $\|\cdot\|_{\text{L}}$ and $\|\cdot\|_{\text{G}}$ be two norms on $\mathbb{R}^n$ with equivalence ratio $k_{\text{L,G}}$. Consider system (4) satisfying Assumptions *(A1)* – *(A3)*. Let r be the largest radius such that $B_{\text{G}}\big(x^\star, r\big) \subseteq \mathcal{S}$. For each trajectory $x(t)$ starting from x_0, it holds that

(i) if $x_0 \in \mathcal{S}$, then, a.e. $t \geq 0$,

$$\|x(t) - x^\star\|_{\text{G}} \leq k_{\text{L,G}} \mathrm{e}^{-c_{\exp} t}\|x_0 - x^\star\|_{\text{G}}; \tag{13}$$

(ii) if $x_0 \notin \mathcal{S}$, then for any *contractor factor* $0 < \rho < 1$ and, a.e. $t \geq 0$,

$$\|x(t) - x^\star\|_{\text{G}} \leq \text{lin-exp}(t\ ;\ q, c_{\text{lin}}, c_{\exp}, t_{\text{c}}), \tag{14}$$

with

- exponential decay rate $c_{\exp} > 0$;
- linear decay rate $c_{\text{lin}} = c_{\exp} r(1-\rho)/\ln(k_{\text{L,G}}\rho^{-1})$;
- intercept $q = \|x_0 - x^\star\|_{\text{G}} + r(1-\rho)\dfrac{\ln(k_{\text{L,G}})}{\ln(k_{\text{L,G}}\rho^{-1})}$;
- linear-exponential crossing time $t_{\text{c}} = \left\lceil \frac{\|x_0 - x^\star\|_{\text{G}} - r}{(1-\rho)r} \right\rceil \ln(k_{\text{L,G}}\rho^{-1})/c_{\exp} + \ln(k_{\text{L,G}})/c_{\exp}$.

Proof: Consider a trajectory $x(t)$ starting from initial condition x_0. If $x_0 \in \mathcal{S}$, then item (i) follows from assumption *(A2)* and the equivalence of norms.

Indeed, Assumption (A2) implies that for every $x_0 \in \mathcal{S}$ and a.e. $t \geq 0$, it holds

$$\|\phi_t(x_0) - x^\star\|_{\mathrm{L}} \leq \mathrm{e}^{-c_{\exp} t}\|x_0 - x^\star\|_{\mathrm{L}}.$$

Applying the equivalence of norms to the above inequality, we get

$$\|\phi_t(x_0) - x^\star\|_{\mathrm{G}} \leq k_{\mathrm{L,G}}\mathrm{e}^{-c_{\exp} t}\|x_0 - x^\star\|_{\mathrm{G}}. \tag{15}$$

If $x_0 \notin \mathcal{S}$, define the point $y_0 := x^\star + r\frac{x_0 - x^\star}{\|x_0 - x^\star\|_{\mathrm{G}}} \in \partial B_{\mathrm{G}}(x^\star, r)$[1]. The norm $\|y_0 - x^\star\|_{\mathrm{G}} = r$, therefore y_0 is a point on the boundary of $B_{\mathrm{G}}(x^\star, r)$. Moreover, the points $x^\star$, y_0, and x_0 lie on the same line segment, thus

$$\|x_0 - x^\star\|_{\mathrm{G}} = \|x_0 - y_0\|_{\mathrm{G}} + r. \tag{16}$$

By Lemma I.2(ii) and because each trajectory originating in $B_{\mathrm{G}}(x^\star, r)$ remains in $\mathcal{S}$, the ρ-contraction with respect to $\|\cdot\|_{\mathrm{G}}$ for the $c_{\exp}$-strongly contracting vector field f is

$$t_\rho^{\mathrm{L,G}} = \frac{\ln(k_{\mathrm{L,G}}\rho^{-1})}{c_{\exp}}. \tag{17}$$

Then, a.e. $t \in [0, t_\rho^{\mathrm{L,G}}]$, we have

$$\begin{aligned}
\|\phi_t(x_0) - x^\star\|_{\mathrm{G}} &\leq \|\phi_t(x_0) - \phi_t(y_0)\|_{\mathrm{G}} + \|\phi_t(y_0) - x^\star\|_{\mathrm{G}} && (18)\\
&\leq \|x_0 - y_0\|_{\mathrm{G}} + k_{\mathrm{L,G}}\mathrm{e}^{-c_{\exp} t}\|y_0 - x^\star\|_{\mathrm{G}} && (19)\\
&\overset{(16)}{=} \|x_0 - x^\star\|_{\mathrm{G}} - \|x^\star - y_0\|_{\mathrm{G}} + k_{\mathrm{L,G}}\mathrm{e}^{-c_{\exp} t} r \\
&\overset{t = t_\rho^{\mathrm{L,G}}}{\leq} \|x_0 - x^\star\|_{\mathrm{G}} - r(1 - k_{\mathrm{L,G}}\mathrm{e}^{-c_{\exp} t_\rho^{\mathrm{L,G}}}) \\
&\overset{(17)}{=} \|x_0 - x^\star\|_{\mathrm{G}} - r(1 - \rho), && (20)
\end{aligned}$$

where in (18) we added and subtracted $\phi_t(y_0)$ and applied the triangle inequality, while inequality (19) follows from Assumption (A1) and inequality (15). Now, (20) implies that $\|\phi_{t_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} \leq \|x_0 - x^\star\|_{\mathrm{G}} - r(1-\rho)$. If $\|x_0 - x^\star\|_{\mathrm{G}} - r(1-\rho) \leq r$, then by Assumption (A2), a.e. in $t \geq t_\rho^{\mathrm{L,G}}$, we have

$$\|\phi_t(x_0) - x^\star\|_{\mathrm{G}} \leq k_{\mathrm{L,G}}\mathrm{e}^{-c_{\exp}(t - t_\rho^{\mathrm{L,G}})}\big(\|x_0 - x^\star\|_{\mathrm{G}} - r(1-\rho)\big).$$

If $\|x_0 - x^\star\|_{\mathrm{G}} - r(1-\rho) > r$, we iterate the process. Specifically, let $x_\rho := \phi_{t_\rho^{\mathrm{L,G}}}(x_0)$, and define $y_\rho := x^\star + r\dfrac{x_\rho - x^\star}{\|x_\rho - x^\star\|_{\mathrm{G}}} \in \partial B_{\mathrm{G}}(x^\star, r)$. Consider the solution to $\dot{y} = f(t, y)$ with initial condition $y(t_\rho^{\mathrm{L,G}}) = y_\rho$ and note that $\phi_t(x_\rho) = \phi_{t + t_\rho^{\mathrm{L,G}}}(x_0)$. For a.e. $t \in [t_\rho^{\mathrm{L,G}}, 2t_\rho^{\mathrm{L,G}}]$, we compute

$$\begin{aligned}
\|\phi_{t + t_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} &\leq \|\phi_t(x_\rho) - \phi_t(y_\rho)\|_{\mathrm{G}} + \|\phi_t(y_\rho) - x^\star\|_{\mathrm{G}} && (21)\\
&\leq \|x_\rho - y_\rho\|_{\mathrm{G}} + k_{\mathrm{L,G}}\mathrm{e}^{-c(t - t_\rho^{\mathrm{L,G}})}\|y_0 - x^\star\|_{\mathrm{G}} && (22)\\
&\overset{(16)}{=} \|x_\rho - x^\star\|_{\mathrm{G}} - \|x^\star - y_\rho\|_{\mathrm{G}} + k_{\mathrm{L,G}}\mathrm{e}^{-c(t - t_\rho^{\mathrm{L,G}})} r \\
&\leq \|\phi_{t_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} - r(1 - k_{\mathrm{L,G}}\mathrm{e}^{-c(t - t_\rho^{\mathrm{L,G}})}) \\
&\overset{(20)}{\leq} \|x_0 - x^\star\|_{\mathrm{G}} - r(1-\rho) - r(1 - k_{\mathrm{L,G}}\mathrm{e}^{-c(t - t_\rho^{\mathrm{L,G}})}) \\
&\overset{t = 2t_\rho^{\mathrm{L,G}}}{\leq} \|x_0 - x^\star\|_{\mathrm{G}} - 2r(1-\rho),
\end{aligned}$$

[1]Note that $\partial B_{\mathrm{G}}(x^\star, r)$ means the boundary of $B_{\mathrm{G}}(x^\star, r)$.

where in (21) we added and subtracted $\phi_t(y_0)$ and applied the triangle inequality, while (19) follows from Assumption *(A1)* and (15). We now reason as done in $[0, t_\rho^{\mathrm{L,G}}]$. If $\|x_0 - x^\star\|_{\mathrm{G}} - 2r(1-\rho) \le r$, then Assumption *(A2)* implies

$$\|\phi_{t+t_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} \le k_{\mathrm{L,G}}\big(\|x_0 - x^\star\|_{\mathrm{G}} - 2r(1-\rho)\big)\mathrm{e}^{-c(t-2t_\rho^{\mathrm{L,G}})}, \quad \forall t \ge 2t_\rho^{\mathrm{L,G}}.$$

If $\|x_0 - x^\star\|_{\mathrm{G}} - 2r(1-\rho) > r$, we proceed analogously until $\|x_0 - x^\star\|_{\mathrm{G}} - 2r(1-\rho) > r$, which happens after $T =: \left\lceil \frac{\|x_0 - x^\star\|_{\mathrm{G}} - r}{(1-\rho)r} \right\rceil$ steps. Iterating the previous process at step T, a.e. $t \in [(T-1)t_\rho^{\mathrm{L,G}}, Tt_\rho^{\mathrm{L,G}}]$, we get

$$\begin{aligned}\|\phi_{t+(T-1)t_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} &\le \|x_0 - x^\star\|_{\mathrm{G}} - (T-1)r(1-\rho) - r(1 - k_{\mathrm{L,G}}\mathrm{e}^{-c(t-(T-1)t_\rho^{\mathrm{L,G}})}),\\ &\overset{t=kt_\rho^{\mathrm{L,G}}}{\le} \|x_0 - x^\star\|_{\mathrm{G}} - Tr(1-\rho) \le r,\end{aligned}$$

where the last inequality follows from the definition of T. Local strong contractivity then implies

$$\|\phi_{t+Tt_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} \le k_{\mathrm{L,G}}\big(\|x_0 - x^\star\|_{\mathrm{G}} - Tr(1-\rho)\big)\mathrm{e}^{-c(t-Tt_\rho^{\mathrm{L,G}})}, \qquad \text{a.e. } t \ge Tt_\rho^{\mathrm{L,G}}.$$

The above reasoning together with Assumption *(A1)* implies that a.e. $t \in [it_\rho^{\mathrm{L,G}}, (i+1)t_\rho^{\mathrm{L,G}}]$, $i \in \{0, \dots, T-1\}$, we have

$$\|\phi_{t+it_\rho^{\mathrm{L,G}}}(x_0) - x^\star\|_{\mathrm{G}} \le \min\Big\{\|x_0 - x^\star\|_{\mathrm{G}} - ir(1-\rho), \|x_0 - x^\star\|_{\mathrm{G}} - ir(1-\rho) - r\big(1 - k_{\mathrm{L,G}}\mathrm{e}^{-c(t-it_\rho^{\mathrm{L,G}})}\big)\Big\}.$$

By partitioning the time interval $[0, +\infty[$ as $[0, t_\rho^{\mathrm{L,G}}[\cup \cdots \cup [(T-1)t_\rho^{\mathrm{L,G}}, Tt_\rho^{\mathrm{L,G}}[\cup [Tt_\rho^{\mathrm{L,G}}, +\infty[$ and summing up the above inequalities we obtain the bound:

$$\begin{aligned}\|\phi_t(x_0) - x^\star\|_{\mathrm{G}} \le \sum_{i=0}^{T-1} \mathbf{1}_{\{it_\rho^{\mathrm{L,G}} \le t < (i+1)t_\rho^{\mathrm{L,G}}\}} \min\Big\{&\|x_0 - x^\star\|_{\mathrm{G}} - ir(1-\rho), \|x_0 - x^\star\|_{\mathrm{G}} - ir(1-\rho)\\ &- r\big(1 - k_{\mathrm{L,G}}\mathrm{e}^{-c(t-it_\rho^{\mathrm{L,G}})}\big)\Big\} + \mathbf{1}_{\{t \ge Tt_\rho^{\mathrm{L,G}}\}} \min\Big\{\|x_0 - x^\star\|_{\mathrm{G}} - Tr(1-\rho),\\ &k_{\mathrm{L,G}}\left(\|x_0 - x^\star\|_{\mathrm{G}} - Tr(1-\rho)\right)\mathrm{e}^{-c(t-Tt_\rho^{\mathrm{L,G}})}\Big\} := g_{\mathrm{B}}(t). \end{aligned} \tag{23}$$

Finally, item (ii) follows by noticing that $g_{\mathrm{B}}(t) \le \text{lin-exp}(t\,;\, q, c_{\mathrm{lin}}, c_{\mathrm{exp}}, t_{\mathrm{c}})$, $t \ge 0$, for the values $t_{\mathrm{c}} = T\ln(k_{\mathrm{L,G}}\rho^{-1})/c_{\mathrm{exp}} + \ln(k_{\mathrm{L,G}})/c_{\mathrm{exp}}$, $c_{\mathrm{lin}} = r\, c_{\mathrm{exp}}(1-\rho)/\ln(k_{\mathrm{L,G}}\rho^{-1})$, and $q = \|x_0 - x^\star\|_{\mathrm{G}} + r(1-\rho)\dfrac{\ln(k_{\mathrm{L,G}})}{\ln(k_{\mathrm{L,G}}\rho^{-1})}$. This concludes the proof. ■

Remark III.7: Theorem III.6 sharpens the convergence bound established in [4]. Namely, it provides a more accurate intercept and linear-exponential crossing time. An illustration between the bound in equation (14) and the one in [4] is given in Figure 3. ■

Remark III.8: The bound in Theorem III.6 generalizes the result for equal norms in Theorem III.4. In fact, the factor $(1-\rho)/\ln(k_{\mathrm{L,G}}\rho^{-1})$ is always less than 1 for $k_{\mathrm{L,G}} > 1$. Moreover, when $k_{\mathrm{L,G}} = 1$ it results $\lim_{\rho \to 1}(1-\rho)/\ln(k_{\mathrm{L,G}}\rho^{-1}) = 1$, thereby exactly recovering the equal-norm result. ■

IV. Local Stability in the Presence of External Inputs

We now characterize local ISS for GW-LS-C systems w.r.t. the same norm. Specifically, we consider the system

$$\dot{x}(t) = f\big(t, x(t), u(t)\big). \tag{24}$$

where, $f\colon \mathbb{R}_{\ge 0} \times \mathcal{C} \times \mathcal{U} \to \mathbb{R}^n$, the map $x \mapsto f(t, x, u)$ is locally Lipschitz, for all t, u, with $\mathcal{C} \subseteq \mathbb{R}^n$ f-invariant, open and convex, and $\mathcal{U} \subset \mathbb{R}^m$. Given $\bar{u} \in \mathbb{R}^m$, we define the set of bounded inputs $\bar{\mathcal{U}} := \{u\colon \mathbb{R}_{\ge 0} \to \mathcal{U} \mid \|u(t)\|_{\mathcal{U}} \le \bar{u}, \forall t \ge 0\}$. We make the following:

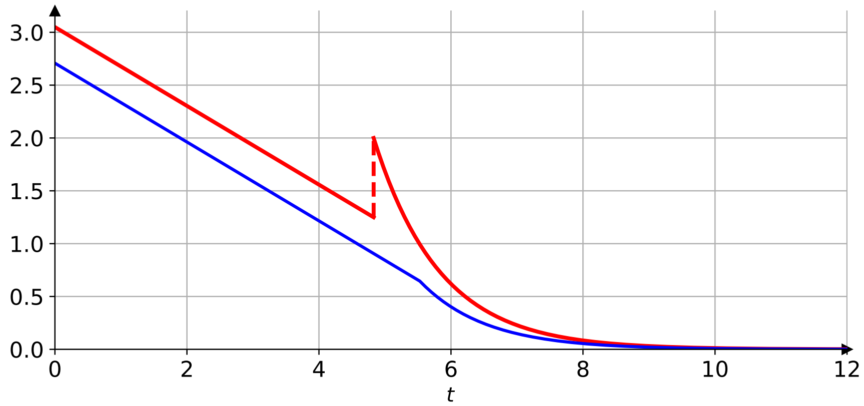

Figure 3: Linear-exponential bound in equation (14) (solid blue curve) and the decay bound presented in [4] (red curve) for $\|x_0 - x^\star\|_G = 2.4$, $r = 1$, $c_{\exp} = 1$, $k_{L,G} = 2$, $\rho = 0.4$.

Assumption 2: there exist norms $\|\cdot\|, \|\cdot\|_{\mathcal{U}}$ on $\mathcal{C}$ and $\mathcal{U}$, respectively, such that

(A1') for all t, u, the map $x \mapsto f(t,x,u)$ is weakly infinitesimally contracting on $\mathbb{R}^n$ w.r.t. $\|\cdot\|$;

(A2') for all t, x, the map $u \mapsto f(t,x,u)$ is Lipschitz with constant $L_u \geq 0$;

(A3') there exist a forward-invariant set $\mathcal{S}$ and $c_{\exp} > 0$ such that, for all t, for each $u \in \bar{\mathcal{U}}$, the map $x \mapsto f(t,x,u(t))$ is $c_{\exp}$-strongly infinitesimally contracting on $\mathcal{S}$ w.r.t. $\|\cdot\|$;

(A4') at $u(t) = \mathbb{0}_m$, for all t, there exists an equilibrium point $x^\star \in \mathcal{S}$.

We begin by giving two technical lemmas, needed to prove the main result of this section.

Lemma IV.1 (Error dynamics for input-dependent systems): Consider system (24) satisfying Assumption (A2'). Then any two solutions $x(t)$ and $y(t)$ with input $u_x, u_y \colon \mathbb{R}_{\geq 0} \to \mathbb{R}^m$, satisfy a.e. $t \geq 0$,

$$D^+\|x(t) - y(t)\| \leq \int_0^1 \mu\Big(Df(y + \alpha(x-y), u_y)\Big) d\alpha \|x(t) - y(t)\| + L_u \|u_x(t) - u_y(t)\|_{\mathcal{U}}. \tag{25}$$

Proof: Let $x(t)$ and $y(t)$ be two trajectories of (24) with input signals u_x, u_y, respectively. Let $[\![\cdot,\cdot]\!]$ be a weak pairing compatible with $\|\cdot\|$. We compute

$$\begin{aligned}
\|x(t) - y(t)\| D^+\|x(t) - y(t)\| &= [\![f(t,x,u_x) - f(t,y,u_y), x - y]\!] \qquad (26)\\
&\leq [\![f(t,x,u_y) - f(t,y,u_y), x - y]\!] + \|f(t,x,u_x) - f(t,x,u_y)\| \|x - y\| \qquad (27)\\
&\leq [\![f(t,x,u_y) - f(t,y,u_y), x - y]\!] + L_u \|u_x - u_y\|_{\mathcal{U}} \|x - y\|, \qquad (28)
\end{aligned}$$

where in (26) we used the curve norm derivative formula (ii), in (27) we added and subtracted $f(t,x,u_y)$ and used the sub-additivity (i) and the Cauchy-Schwartz inequality (iii), and in (28) we used Assumption (A2').

Next, by dividing both side for $\|x(t) - y(t)\|$ we get

$$\begin{aligned}
D^+\|x(t) - y(t)\| &\leq \frac{[\![f(t,x,u_y) - f(t,y,u_y), x - y]\!]}{\|x - y\|} + L_u \|u_x - u_y\|_{\mathcal{U}}\\
&= \frac{[\![f(t,x,u_y) - f(t,y,u_y), x - y]\!]}{\|x - y\|^2} \|x - y\| + L_u \|u_x - u_y\|_{\mathcal{U}}. \qquad (29)
\end{aligned}$$

By applying the mean-value Theorem II.1 to (29), a.e., we get

$$\begin{aligned}
D^+\|x(t) - y(t)\| &\leq \frac{\Big[\!\!\Big[\int_0^1 Df(y + s(x-y), u_y) ds (x-y), x - y\Big]\!\!\Big]}{\|x - y\|} \frac{\|x - y\|}{\|x - y\|} + L_u \|u_x - u_y\|_{\mathcal{U}} \qquad (30)\\
&\leq \int_0^1 \frac{[\![Df(y + s(x-y), u_y) ds (x-y), x - y]\!]}{\|x - y\|^2} ds \|x - y\| + L_u \|u_x - u_y\|_{\mathcal{U}}, \qquad (31)
\end{aligned}$$

where in (30) we have used the weak pairing sub-additivity (i). Next, recall that Lumer's equality (iv) implies, $\frac{[\![Az,z]\!]}{[\![z,z]\!]} \leq \mu(A)$ for every $A \in \mathbb{R}^{n\times n}$ and $z \neq \mathbb{0}_n$. By applying this equality to (31) (with $A = Df(y+s(x-y),u_y)$ and $z = x-y$) we get inequality (25). This concludes the proof. ■

The next result gives a linear-exponential bound for the solution of dynamics with saturations and additive inputs.

Lemma IV.2 (Solution of dynamics with saturations and additive inputs): Let c_{exp} and d be positive scalars, and $u\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ satisfying $\|u(t)\|_\infty = u_{\max} < dc_{\text{exp}}$, for all t. Consider the dynamics

$$\dot{x}(t) = -c_{\text{exp}}\,\text{sat}_d(x(t)) + u(t), \quad x_0 = q > d. \tag{32}$$

Then, a solution of (32) satisfies

$$x(t) \leq \text{lin-exp}(t\ ;\ q, c_{\text{lin}}, c_{\text{exp}}, t_c) + \iota_{[t_c,+\infty[}(t)(1-\mathrm{e}^{-c_{\text{exp}}(t-t_c)})\frac{u_{\max}}{c_{\text{exp}}},$$

with $c_{\text{lin}} := dc_{\text{exp}} - u_{\max} > 0$ and $t_c := \frac{q-d}{c_{\text{lin}}} > 0$.

Proof: Using the definition of saturation function, for all $t \in \mathbb{R}_{\geq 0}$ we can upper bound the ODE (32) as

$$\dot{x}(t) \leq \dot{y}(t) := \begin{cases} -dc_{\text{exp}} + u_{\max} & \text{if } y(t) > d, \\ -c_{\text{exp}}x(t) + u_{\max} & \text{if } y(t) \in [-d,d], \\ dc_{\text{exp}} + u_{\max} & \text{if } y(t) < -d, \end{cases} \tag{33}$$

which, in each interval $[t_0,t_1] \subseteq \mathbb{R}_{\geq 0}$ where the solution is continuous and does not change regime, has general solution

$$y(t) = \begin{cases} (-dc_{\text{exp}} + u_{\max})t + y(t_0) & \text{if } y(t) > d, \\ \Big(y(t_0) - \dfrac{\bar{u}}{c_{\text{exp}}}\Big)\mathrm{e}^{-c_{\text{exp}}(t-t_0)} + \dfrac{u_{\max}}{c_{\text{exp}}} & \text{if } y(t) \in [-d,d], \\ (dc_{\text{exp}} + u_{\max})t + y(t_0) & \text{if } y(t) < -d. \end{cases} \tag{34}$$

At time $t = 0$, we have $x_0 = q > d$. For continuity of the solution, there exists $t^\star$ such that $y(t) > d$ for all $t \in [0,t^\star]$. Thus from (34) and being $x(t_0 = 0) = q$, it is

$$y(t) = (-dc_{\text{exp}} + u_{\max})t + q,$$

for all $t \in [0,t^\star]$. Moreover being $u_{\max} < dc_{\text{exp}}$, the function $y(t)$ is decreasing, the time value $t^\star$ is finite and there exists a time, say it $\bar{t}$ such that $y(\bar{t}) = d$. Let $c_{\text{lin}} := dc_{\text{exp}} - u_{\max}$, we have

$$y(\bar{t}) = d \iff -c_{\text{lin}}\bar{t} + q = d \iff \bar{t} = \frac{q-d}{c_{\text{lin}}} := t_c.$$

In summary, we have shown that $y(t) = q - c_{\text{lin}}t$ for all $t \in [0,t_c]$ and is equal to d at time t_c. Therefore from (34) and being $y(t_c) = q - c_{\text{lin}}t_c$, for all time $t > t_c$ we have

$$y(t) = \big(q - c_{\text{lin}}t_c\big)\mathrm{e}^{-c_{\text{exp}}(t-t_c)} + \big(1-\mathrm{e}^{-c_{\text{exp}}(t-t_c)}\big)\frac{u_{\max}}{c_{\text{exp}}}.$$

Specifically, $y(t) > 0$ for all $t > t_c$, thus it can never be $y(t) < -d$. This concludes the proof. ■

We are now ready to state the following:

Theorem IV.3 (Local ISS for input-dependent GW-LS-C systems): Consider system (24) satisfying Assumptions (A1') – (A4'). Let r be the largest radius such that $B\big(x^\star,r\big) \subseteq \mathcal{S}$, $\bar{u} < rc_{\text{exp}}$, and $u_{\max} := \sup_{\tau\in[0,t]}\|u_x(\tau)\|_{\mathcal{U}}$. For each trajectory $x(t)$ with input $u_x \in \bar{\mathcal{U}}$ starting from $x_0 \notin \mathcal{S}$, a.e. $t \geq 0$, we have:

(i) $D^+\|x(t)-x^\star\| \leq -c_{\mathrm{exp}}\,\mathrm{sat}_r(\|x(t)-x^\star\|)+L_u\|u_x(t)\|_{\mathcal{U}}$.

(ii) $\|x(t)-x^\star\| \leq \text{lin-exp}(t\ ;\ q, c_{\mathrm{lin}}, c_{\mathrm{exp}}, t_{\mathrm{c}}) + \iota_{[t_{\mathrm{c}},+\infty[}(t)\dfrac{L_u}{c_{\mathrm{exp}}}(1-\mathrm{e}^{-c_{\mathrm{exp}}t})u_{\max}$,

with

- exponential decay rate $c_{\mathrm{exp}} > 0$;
- intercept $q = \|x_0 - x^\star\|$;
- linear decay rate $c_{\mathrm{lin}} = rc_{\mathrm{exp}} - u_{\max}$;
- linear-exponential crossing time $t_{\mathrm{c}} = (q-r)/c_{\mathrm{lin}}$.

Proof: Consider an arbitrary trajectory $x(t)$ starting from $x_0 \notin \mathcal{S}$ with input u_x and a second trajectory equal to the equilibrium $x^\star$ with input $u = \mathbb{0}_m$. To prove statement (i), let μ be the log-norm associated to $\|\cdot\|$. By applying inequality (25) to those trajectories, a.e. $t \geq 0$, we have

$$D^+\|x(t)-x^\star\| \leq \int_0^1 \mu\Big(Df(x^\star + \alpha(x(t)-x^\star), 0)\Big)d\alpha\|x(t)-x^\star\| + L_u\|u_x\|_{\mathcal{U}}. \tag{35}$$

The proof follows by using similar reasoning as the one in the proof of Lemma III.3. Item (ii) follows by using the Comparison Lemma [13, pp. 102-103] and Lemma IV.2 to upper bound the solution to the differential inequality (i). ■

V. Tackling Linear Programs

We now show the efficacy of the previous results by applying them to a dynamical system solving the LP problem. Given $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, we consider the *linear program*:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^\top x, \\ \text{s.t.} \quad & Ax \leq b, \end{aligned} \tag{36}$$

and its equivalent unconstrained formulation

$$\min_{x\in\mathbb{R}^n} c^\top x + \iota_{\mathcal{I}_b}(Ax), \tag{37}$$

where $\mathcal{I}_b = \{y \in \mathbb{R}^m \mid y - b \leq \mathbb{0}_m\}$. We assume that (37) admits a unique equilibrium. Note that (37) is a particular composite minimization problem:

$$\min_{x\in\mathbb{R}^n} f(x) + g(Ax), \tag{38}$$

with $f(x) = c^\top x$ and $g(Ax) = \iota_{\mathcal{I}_b}(Ax)$. To solve (37), we leverage the proximal augmented Lagrangian approach proposed in [10] and consider the *proximal augmented Lagrangian*, $\tilde{L}_\gamma\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, defined by

$$\tilde{L}_\gamma(x,\lambda) = f(x) + M_{\gamma g}(Ax + \gamma\lambda) - \frac{\gamma}{2}\|\lambda\|_2^2, \tag{39}$$

where $\lambda \in \mathbb{R}^m$ is the Lagrange multiplier, $\gamma > 0$ is a parameter, and $M_{\gamma g}$ is Moreau envelope of g.

Remark V.1: For f continuously differentiable, convex, and with a Lipschitz continuous gradient, and g convex, closed and proper, solving the composite minimization problem (38) corresponds to finding saddle points of (39), simultaneously updating the primal and dual variables [10, Theorem 2]. ■

Next, consider the *continuous-time augmented primal-dual dynamics* [10] (that can be interpreted as a continuous-time neural network) associated to the proximal augmented Lagrangian of problem (37)

$$\begin{aligned} \dot{x} &= -\nabla_x\tilde{L}_\gamma(x,\lambda) = -c - A^\top\nabla M_{\gamma\iota_{\mathcal{I}_b}}(Ax+\gamma\lambda) = -c - \frac{1}{\gamma}A^\top\,\mathrm{ReLU}\big(Ax+\gamma\lambda-b\big), \\ \dot{\lambda} &= \nabla_\lambda\tilde{L}_\gamma(x,\lambda) = -\gamma\lambda + \gamma\nabla M_{\gamma\iota_{\mathcal{I}_b}}(Ax+\gamma\lambda) = -\gamma\lambda + \mathrm{ReLU}\big(Ax+\gamma\lambda-b\big). \end{aligned} \tag{40}$$

We let $F_{\text{LP}}\colon \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ denote the vector field for (40).

Remark V.2: Equation (40) follows directly after noticing that for a.e. $y \in \mathbb{R}^m$ it results

$$\nabla M_{\gamma\iota_{\mathcal{I}}}(y) = \frac{1}{\gamma}\big(y - \mathbb{P}_{\iota_{\mathcal{I}}}(y)\big) = \frac{1}{\gamma}\big(y - \min\{y, b\}\big) = \frac{1}{\gamma}\operatorname{ReLU}\big(y - b\big).$$

■

The next result characterizes the convergence of (40).

Theorem V.3 (Convergence of the linear program): Consider the dynamics (40) and let $(x^\star, \lambda^\star) \in \mathbb{R}^{n+m}$ be an equilibrium point. If $DF_{\text{LP}}\big(x^\star, \lambda^\star\big)$ is Hurwitz, then any solution of (40) linear-exponentially converges towards $(x^\star, \lambda^\star)$.

Proof: To prove the statement we show that (40) satisfies the assumptions of Theorem III.6. First, we prove that the system is globally-weakly contracting. To this purpose, let $z := (x, \lambda) \in \mathbb{R}^{n+m}$, $y := Ax + \gamma\lambda - b$ and define $G(y) := D\operatorname{ReLU}\big(y\big)$, a.e. $y \in \mathbb{R}^m$. The Jacobian of (40) is

$$DF_{\text{LP}}(z) = \begin{bmatrix} -\frac{1}{\gamma}A^\top G(y)A & -A^\top G(y) \\ G(y)A & -\gamma(I_m - G(y)) \end{bmatrix}.$$

Being $0 \preceq G(y) \preceq I_m$[2], a.e. $y \in \mathbb{R}^m$, we have

$$\sup_z \mu_2\big(DF_{\text{LP}}(z)\big) \leq \max_{0 \preceq G \preceq I_m} \mu_2\left(\begin{bmatrix} -\gamma^{-1}A^\top GA & -A^\top G \\ GA & \gamma\big(G - I_m\big) \end{bmatrix}\right),$$

By definition of μ_2, we have that

$$\begin{aligned}\mu_2\left(\begin{bmatrix} -\gamma^{-1}A^\top GA & -A^\top G \\ GA & \gamma\big(G - I_m\big) \end{bmatrix}\right) &= \lambda_{\max}\left(\begin{bmatrix} -\gamma^{-1}A^\top GA & 0 \\ 0 & \gamma\big(G - I_m\big) \end{bmatrix}\right) \\ &= \max\{\lambda_{\max}\big(-\gamma^{-1}A^\top GA\big), \lambda_{\max}\big(\gamma\big(G - I_m\big)\big)\} = 0.\end{aligned}$$

The last equality follows from the fact that $\lambda_{\max}\big(-\gamma^{-1}A^\top GA\big) = \lambda_{\max}\big(\gamma\big(G - I_m\big)\big) = 0$. In particular, the equality $\lambda_{\max}\big(-\gamma\big(G - I_m\big)\big) = 0$ follows directly from $0 \preceq G \preceq I_m$; while $\lambda_{\max}\big(-\gamma^{-1}A^\top GA\big) = 0$, follows noticing that $A^\top GA \succeq 0$[3]. This implies that (40) is weakly contracting on $\mathbb{R}^{n+m}$ w.r.t. $\|\cdot\|_2$. Next, we prove that the system is locally-strongly contracting. To do so, we first note that for any equilibrium point $z^\star := (x^\star, \lambda^\star)$ of (40), the KKT conditions implies that $y^\star := Ax^\star + \gamma\lambda^\star - b \neq \mathbb{0}_m$. This in turn implies that both $D\operatorname{ReLU}(y^\star)$ and $DF_{\text{LP}}(z^\star)$ are well defined. Now, being by assumption $DF_{\text{LP}}(z^\star)$ Hurwitz, there exists Q invertible such that $\mu_{2,Q}\big(DF_{\text{LP}}(z^\star)\big) < 0$ [3, Corollary 2.33]. Let $\mathcal{K}$ be the set of differentiable points in a neighborhood of $z^\star$. Then, by the continuity property of the log-norm, there exists $B_{2,Q}\big(z^\star, p\big)$, with $p := \sup\{p > 0 \mid B_{2,Q}\big(z^\star, p\big) \subset \mathcal{K}\}$, where $DF_{\text{LP}}(z)$ exists and $\mu_{2,Q}\big(DF_{\text{LP}}(z)\big) < -c_{\text{exp}}$ for all $z \in B_{2,Q}\big(z^\star, p\big)$, for some $c_{\text{exp}} > 0$. Therefore (40) is strongly infinitesimally contracting w.r.t. $\|\cdot\|_{2,Q}$ in $B_{2,Q}\big(z^\star, p\big)$. This concludes the proof. ■

A key hypothesis in Theorem V.3 is that $DF_{\text{LP}}\big(x^\star, \lambda^\star\big)$ is Hurwitz. To this aim, we make the following:

Conjecture V.4: Let $(x^\star, \lambda^\star)$ be the equilibrium of (40). The LP (36) has a unique solution, $x^\star$, if and only if $DF_{\text{LP}}\big(x^\star, \lambda^\star\big)$ is Hurwitz.

[2]For every $\gamma > 0$, $0 \preceq \nabla^2 M_{\gamma g}(y) \preceq \frac{1}{\gamma}I_n$, a.e. $y \in \mathbb{R}^m$ [8, Lemma 18].

[3]$A^\top GA \succeq 0 \iff x^\top A^\top GAx \geq 0 \quad \forall x \in \mathbb{R}^n \iff y^\top Gy \geq 0, \forall y \in \mathbb{R}^m \iff G \succeq 0.$

a) *Numerical Experiments*

Consider the following LP

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & x_1 + x_2 + x_3, \\ \text{s.t.} \quad & -1 \le x_1 \le 1, -1 \le x_2 \le 1, -1 \le x_3 \le 1. \end{aligned} \tag{41}$$

for which the unique optimal solution is $x^\star = (-1, -1, -1)$.

Next, consider the corresponding continuous-time augmented primal-dual dynamics (40)

$$\begin{aligned} \dot{x}_1 &= -1 - \frac{1}{\gamma}\Big(\mathrm{ReLU}\left(x_1 + \gamma\lambda_1 - 1\right) - \mathrm{ReLU}\left(-x_1 + \gamma\lambda_4 - 1\right)\Big), \\ \dot{x}_2 &= -1 - \frac{1}{\gamma}\Big(\mathrm{ReLU}\left(x_2 + \gamma\lambda_2 - 1\right) - \mathrm{ReLU}\left(-x_2 + \gamma\lambda_5 - 1\right)\Big), \\ \dot{x}_3 &= -1 - \frac{1}{\gamma}\Big(\mathrm{ReLU}\left(x_3 + \gamma\lambda_3 - 1\right) - \mathrm{ReLU}\left(-x_3 + \gamma\lambda_6 - 1\right)\Big), \\ \dot{\lambda}_1 &= -\gamma\lambda_1 + \mathrm{ReLU}\left(x_1 + \gamma\lambda_1 - 1\right) \\ \dot{\lambda}_2 &= -\gamma\lambda_2 + \mathrm{ReLU}\left(x_2 + \gamma\lambda_2 - 1\right), \\ \dot{\lambda}_3 &= -\gamma\lambda_3 + \mathrm{ReLU}\left(x_3 + \gamma\lambda_3 - 1\right), \\ \dot{\lambda}_4 &= -\gamma\lambda_4 + \mathrm{ReLU}\left(-x_1 + \gamma\lambda_4 - 1\right), \\ \dot{\lambda}_5 &= -\gamma\lambda_5 + \mathrm{ReLU}\left(-x_2 + \gamma\lambda_5 - 1\right), \\ \dot{\lambda}_6 &= -\gamma\lambda_6 + \mathrm{ReLU}\left(-x_3 + \gamma\lambda_6 - 1\right). \end{aligned} \tag{42}$$

We set $\gamma = 0.5$ and simulate the dynamics (40) over the time interval $t \in [0, 40]$ with a forward Euler discretization with step-size $\Delta t = 0.001$, starting from 150 initial conditions generated as follows: we first randomly generate an initial condition and then define the remaining 149 initial conditions by adding, to the first initial condition, random noise generated from a normal distribution with mean 0 and standard deviation 2. The simulation results are that each resulting trajectory converges to the point $z^\star = (-1, -1, -1, 0, 0, 0, 1, 1, 1)$. Next, we numerically found that $DF_{LP}(z^\star)$ is Hurwitz (in alignment with our conjecture). Figure 4 illustrates the mean and standard deviation of the lognorm of the ℓ_2 distance of the 150 simulated trajectories of (40) w.r.t. $z^\star$. In agreement with Theorem V.3 the convergence is linearly-exponentially bounded.

VI. Conclusion

We analyzed the convergence characteristics of GW-LS-C dynamics, which naturally arise from convex optimization problems with a unique minimizer. For such dynamics, we showed linear-exponential convergence to the equilibrium. Specifically, we demonstrated that linear-exponential dependency arises naturally in certain dynamics with saturations and used this result for our convergence analysis. Depending on the norms where the system is GW-LS-C, we considered two different scenarios that required two distinct mathematical approaches, yielding convergence bounds that are sharper than those in [4]. Finally, after giving a sufficient condition for local ISS, we illustrated our results on the continuous-time augmented primal-dual dynamics solving LPs. Our results motivated a conjecture relating the optimal solution of LPs to the local stability properties of the equilibrium of the resulting dynamics. Our future work will include proving this conjecture, extending our ISS analysis to the case of different norms and further developing our results to design biologically plausible neural networks.

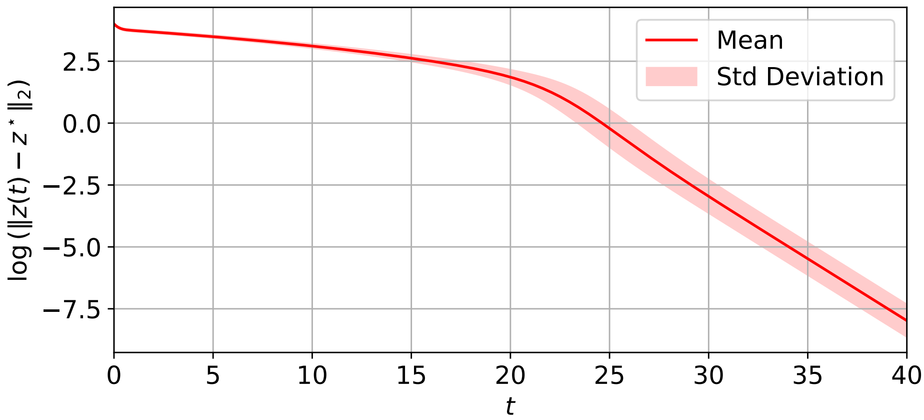

Figure 4: Mean (red curve) and standard deviation (shadow curve) of the lognorm of the Euclidean distance of 150 simulated trajectories of (40) with respect to the equilibrium point $z^\star$. In agreement with Theorem V.3 the convergence is linearly-exponentially bounded.

Acknowledgement

The authors wish to thank Eduardo Sontag for stimulating conversations about contraction theory.

This work was in part supported by AFOSR project FA9550-21-1-0203 and NSF Graduate Research Fellowship under Grant No. 2139319.

Giovanni Russo wishes to acknowledge financial support by the European Union - Next Generation EU, under PRIN 2022 PNRR, Project "Control of Smart Microbial Communities for Wastewater Treatment".

Appendix I
Contraction times with respect to distinct norms

First we recall the following [4, Lemma V.1]

Lemma I.1 (Inclusion between balls computed with respect to different norms): Given two norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ on $\mathbb{R}^n$, for all $x \in \mathbb{R}^n$ and $r > 0$, it holds

$$B\big(x, r/k_\alpha^\beta; \beta\big) \subseteq B\big(x, r; \alpha\big) \subseteq B\big(x, r k_\beta^\alpha; \beta\big). \tag{43}$$

The following Lemma is inspired by [4, Theorem V.2]. For completeness, we here provide a self-contained proof.

Lemma I.2 (Contraction times with respect to distinct norms): Given $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ norms on $\mathbb{R}^n$ with equivalence ratio $k_{\alpha,\beta}$, consider system (4) satisfying Assumptions (*A2*), (*A3*) with $\|\cdot\|_{\mathrm{L}} = \|\cdot\|_\alpha$. Then, for each $0 < \rho < 1$,

(i) the ρ-contraction time is $t_\rho = \ln(\rho^{-1})/c$;
(ii) the ρ-contraction time with respect to the norm $\|\cdot\|_\beta$ is $t_\rho^{\alpha,\beta} = \ln(k_{\alpha,\beta}\,\rho^{-1})/c$.

Proof: Consider a trajectory $x(t)$ of system (4) such that $\|x_0\|_\alpha \leq r$. To prove (i) we need to find the first time t_ρ such that $\|x(t_\rho) - x^\star\|_\alpha \leq \rho r$. Clearly the worst-case time is achieved when $\|x_0 - x^\star\|_\alpha = r$. But c-strongly infinitesimal contractivity with respect to $\|\cdot\|_\alpha$ implies $\|x(t) - x^\star\|_\alpha \leq \mathrm{e}^{-ct}\|x_0 - x^\star\|_\alpha$ and so t_ρ is determined by the equality $\mathrm{e}^{-ct_\rho} r = \rho r$, from which item (i) follows.

Regarding item (ii), we need to find the first time $t_\rho^{\alpha,\beta}$ such that $\|x(t_\rho) - x^\star\|_\beta \leq \rho r$. We note that

$$x_0 \in B_\beta\big(x^\star, r\big) \overset{(43),\ 2^{\text{nd}}\text{ inequality}}{\Longrightarrow} x_0 \in B_\alpha\big(x^\star, k_\alpha^\beta r\big),$$
$$x(t_\rho) \in B_\beta\big(x^\star, \rho r\big) \overset{(43),\ 1^{\text{st}}\text{ inequality}}{\Longleftarrow} x(t_\rho) \in B_\alpha\big(x^\star, \rho r/k_\beta^\alpha\big).$$

Thus, the contraction time from $B_\beta(x^\star, r)$ to $B_\beta(x^\star, \rho r)$ is upper bounded by the contraction time from $B_\alpha(x^\star, k_\alpha^\beta r)$ to $B_\alpha(x^\star, \rho r/k_\beta^\alpha)$. Therefore, the contraction factor with respect to the $\|\cdot\|_\alpha$ norm is $(\rho r/k_\beta^\alpha)/(k_\alpha^\beta r) = \rho/k_{\alpha,\beta}$. Item (ii) then follows from item (i). ■

References

[1] K. J. Arrow, L. Hurwicz, and H. Uzawa, editors. *Studies in Linear and Nonlinear Programming*. Stanford University Press, 1958.

[2] 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:10.1109/TCNS.2021.3112762.

[3] F. Bullo. *Contraction Theory for Dynamical Systems*. Kindle Direct Publishing, 1.1 edition, 2023, ISBN 979-8836646806. URL: https://fbullo.github.io/ctds.

[4] V. Centorrino, A. Gokhale, A. Davydov, G. Russo, and F. Bullo. Positive competitive networks for sparse reconstruction. *Neural Computation*, January 2024. To appear. doi:10.48550/arXiv.2311.03821.

[5] E. K. P. Chong, S. Hui, and S. H. Zak. An analysis of a class of neural networks for solving linear programming problems. *IEEE Transactions on Automatic Control*, 44(11):1995–2006, 1999. doi:10.1109/9.802909.

[6] F. H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley & Sons, 1983, ISBN 047187504X.

[7] S. Coogan. A contractive approach to separable Lyapunov functions for monotone systems. *Automatica*, 106:349–357, 2019. doi:10.1016/j.automatica.2019.05.001.

[8] 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. doi:10.48550/arXiv.2305.15595.

[9] A. Davydov, A. V. Proskurnikov, and F. Bullo. Non-Euclidean contraction analysis of continuous-time neural networks. *IEEE Transactions on Automatic Control*, August 2023. Submitted. doi:10.48550/arXiv.2110.08298.

[10] N. K. Dhingra, S. Z. Khong, and M. R. Jovanović. The proximal augmented Lagrangian method for nonsmooth composite optimization. *IEEE Transactions on Automatic Control*, 64(7):2861–2868, 2019. doi:10.1109/TAC.2018.2867589.

[11] S. Hassan-Moghaddam and M. R. Jovanović. Proximal gradient flow and Douglas-Rachford splitting dynamics: Global exponential stability via integral quadratic constraints. *Automatica*, 123:109311, 2021. doi:10.1016/j.automatica.2020.109311.

[12] S. Jafarpour, P. Cisneros-Velarde, and F. Bullo. Weak and semi-contraction for network systems and diffusively-coupled oscillators. *IEEE Transactions on Automatic Control*, 67(3):1285–1300, 2022. doi:10.1109/TAC.2021.3073096.

[13] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002, ISBN 0130673897.

[14] W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6):683–696, 1998. doi:10.1016/S0005-1098(98)00019-3.

[15] H. D. Nguyen, T. L. Vu, K. Turitsyn, and J.-J. E. Slotine. Contraction and robustness of continuous time primal-dual dynamics. *IEEE Control Systems Letters*, 2(4):755–760, 2018. doi:10.1109/LCSYS.2018.2847408.

[16] G. Russo, M. Di Bernardo, and E. D. Sontag. Global entrainment of transcriptional systems to periodic inputs. *PLoS Computational Biology*, 6(4):e1000739, 2010. doi:10.1371/journal.pcbi.1000739.

[17] D. W. Tank and J. J. Hopfield. Simple "neural" optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit. *IEEE Transactions on Circuits and Systems*, 33(5):533–541, 1986. doi:10.1109/TCS.1986.1085953.

[18] S. Xie, G. Russo, and R. H. Middleton. Scalability in nonlinear network systems affected by delays and disturbances. *IEEE Transactions on Control of Network Systems*, 8(3):1128–1138, 2021. doi:10.1109/TCNS.2021.3058934.