Zames–Falb Multipliers for Convergence Rate: Motivating example and convex searches

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

Zames–Falb Multipliers for Convergence Rate: Motivating example and convex searches

Jingfan.Zhang[a] Peter Seiler[b] and Joaquin Carrasco[a]

[a] Department of Electrical & Electronic Engineering, School of Engineering, University of Manchester, M13 9PL, UK. [b]Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, USA.

ABSTRACT
The bound of convergence rates for discrete-time Lur'e systems has recently attracted much attention. The contributions of this technical note are twofold. Firstly, we show an example where the asymptotic convergence rate is slower than the linear case. This example can be seen as a counterexample of the Kalman conjecture extended to the analysis of convergence rates and motivates a rigorous analysis. Secondly, we develop a result to bound the convergence rate as a convex search over the suitable subclass of noncausal FIR Zames-Falb multipliers. The convex search is demonstrated with several numerical examples.

KEYWORDS
Exponential convergence rate; Zames-Falb multipliers; integral quadratic constraint; Kalman conjecture

1. Introduction

A classic topic in control theory is the Lur'e problem (Lurie & Postnikov, 1944), which concerns the stability of a feedback interconnection between a linear time-invariant (LTI) system G and any uncertainty or nonlinearity Δ within some classes (see Figure 1[1]). Particularly, when the nonlinearity Δ is slope-restricted, the input-output stability of the closed-loop system can be tested by Zames-Falb theorem, where the class of Zames-Falb multipliers $\mathcal{M}$, defined in both continuous-time (G.Zames & P.L.Falb, 1968) and discrete-time (Willems & Brockett, 1968; Willems, 1971) (see Carrasco, Turner, and Heath (2016) for a tutorial), is often involved to preserve the positivity of the class of nonlinearities. Then the absolute stability problem is reduced to a search of $M \in \mathcal{M}$ such that

$$Re\{M(z)(-1 + KG(z))\} < 0 \quad \forall |z| = 1, \tag{1}$$

where K is the maximum slope of the nonlinearity.

This condition can be expressed equivalently in the IQC framework (Megretski & Rantzer, 1997), where the frequency domain inequality (FDI) can be converted to computable linear matrix inequalities (LMIs) by the Kalman-Yakubovich-Popov (KYP) lemma (Rantzer, 1996).

CONTACT Joaquin Carrasco Email: joaquin.carrasco@manchester.ac.uk

[1]The Lur'e system or Lur'e problem is originally defined for unforced systems, but it is relaxed to forced systems (Altshuller, 2013).

In continuous-time, there exist several search algorithms which have been summarised in Carrasco et al. (2016). In discrete-time, searches over finite impulse response (FIR) Zames-Falb multipliers have been proposed in Carrasco, Heath, Zhang, Ahmad, and Wang (2019); Wang, Heath, and Carrasco (2014).
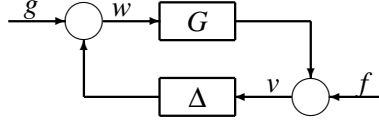

Figure 1. The Lur'e system

A direct analysis of the rate of exponential convergence has attracted much attention in recent studies, since first-order optimisation algorithms, such as gradient decent method and Nesterov method, can be rewritten as Lur'e systems (Lessard, Recht, & Packard, 2016). Therefore, the convergence rate of these optimisation algorithms can be translated into the exponential rate of the corresponding unforced Lur'e systems.

One of the cornerstone to understand the development of the absolute stability theory is the Kalman conjecture (Kalman, 1957). It states that the Lur'e system is stable for all slope-restricted nonlinearities if and only if the system is stable for all the corresponding linear gains (Kalman, 1957). This conjecture is false in general, but provides useful insights. A first contribution of this note is to extended the Kalman conjecture to the analysis of convergence rate and to show that it is false in Section 3. To the best of our knowledge, we show the first counterexample to this conjecture for the convergence rate. **The examples in the literature have shown systems with limit cycles (Lessard et al., 2016; Scoy, Freeman, & Lynch, 2018) so they are equivalent to the counterexamples to the Kalman conjecture given in (Carrasco, Heath, & de la Sen, 2015; Heath, Carrasco, & de la Sen, 2015). For all these examples, the classical analysis would not be able to show stability (Carrasco et al., 2019; Willems & Brockett, 1968). Here, we show that the convergence rate of an asymptotically stable system could be slower than the linearised systems, hence the classical analysis is insufficient. The fundamental difference between both classes of examples is the required nonlinearity as all trajectories are asymptotically convergence, so a discontinuity in the derivative around zero is required.**

In Boczar, Lessard, Packard, and Recht (2017); Boczar, Lessard, and Recht (2015); Freeman (2018); Hu and Seiler (2016), the convergence rate ρ of a Lur'e system has been estimated in two frameworks, as summarised below. By these techniques, less-conservative results are obtained in optimisation (Badithela & Seiler, 2019; Cyrus, Hu, Scoy, & Lessard, 2018; Fazlyab, Ribeiro, Morari, & Preciado, 2018; Scoy et al., 2018) and control (Nelson & Mallada, 2018).

- On the one hand, the conventional IQC analysis is applied in Hu and Seiler (2016), where exponential stability (with rate $0 < \rho < 1$) of the original system in Figure 1 is tested by linearly stability ($\rho = 1$) of the scaled system in Figure 2. The IQCs are constructed for the scaled uncertainty Δ_ρ, and the corresponding exponential stability condition is to search for a multiplier that belongs to a suitable subset of $\mathscr{M}$, such that

$$Re\{M(z)(-1 + KG_\rho(z))\} < 0 \quad \forall |z| = 1, \tag{2}$$

  where the multiplier is in the same form of original FIR Zames-Falb multipliers, while the ℓ_1-norm condition is penalised with ρ.
- On the other hand, in Boczar et al. (2017, 2015); Freeman (2018), the so-called ρ-IQC
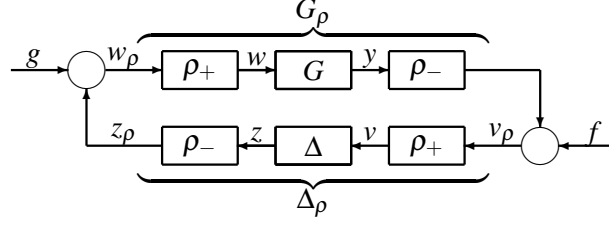
Figure 2. The scaled system in Hu and Seiler (2016). A mathematical definition of the operators ρ_+ and ρ_- is given in Definition 2.8

framework is proposed, where the exponential stability of the original system in Figure 1 can be tested directly. Similarly, ρ-IQCs are constructed for the original uncertainty Δ, and the exponential stability condition is to search for a multiplier that also belongs to a subset of $\mathcal{M}$, such that

$$Re\{M(\rho z)(-1 + KG(\rho z))\} < 0 \quad \forall |z| = 1, \tag{3}$$

where the multiplier is constructed from the original FIR Zames-Falb multiplier $M(z)$ by replacing z by ρz, and the ℓ_1-norm condition is also penalised with ρ.

In both approaches, sound analysis on the causal FIR Zames-Falb multipliers are provided in the literature above. In the IQC approach, noncausal multipliers are not used in Hu and Seiler (2016). In the ρ-IQC approach, the noncausal FIR Zames-Falb multiplier in the form $M(\rho z)$ is studied in Freeman (2018), where its modified ℓ_1-norm condition is deliberated. **In contrast, the techniques to search over the class of noncausal multipliers have not been considered in the literature above.** In addition, due to the extra complexity to deal with the asymmetric variables ρz and $\frac{\rho}{z}$ (instead of $\frac{1}{\rho z}$), the IQC factorisations (Scherer & Weiland, 2011) are restricted in the ρ-IQC framework. The relation is shown in Figure 3.
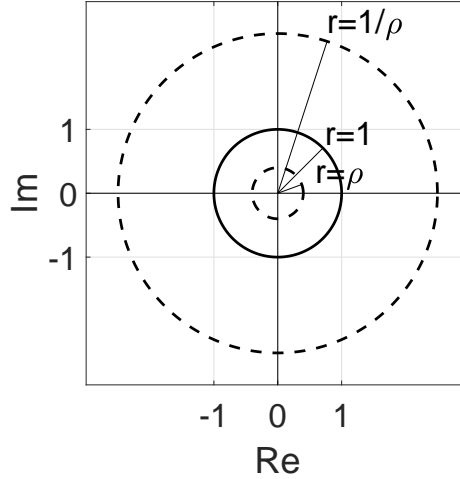
Figure 3. The analysis is on the circle with radius 1 in IQC framework; the analysis is on the circles with radii ρ and $\frac{1}{\rho}$ in ρ-IQC framework.

The second contribution of this note is to provide a unified convex search for noncausal multipliers in Section 4, following some of the techniques used in Carrasco et al. (2019); Wang et al. (2014). As a result, an easy-to-check LMI is provided to estimate a tighter bound of the convergence rate that can be used as a unified structure for FIR Zames-Falb multipliers.

3

Numerical results are demonstrated to lead to less-conservative analysis in Section 5.

This note formalises the search of noncausal multipliers initially drafted in (Zhang, Seiler, & Carrasco, 2019), which has been already used in (Michalowsky, Scherer, & Ebenbauer, 2020) as part of the robustness analysis of optimization algorithms. The main difference with the classical lifting technique in (Carrasco et al., 2019) is the requirement to deal with the causal and anticausal parts in different vectors. Moreover, the factorisation method in this note is applicable for both IQC and ρ-IQC approaches by simply changing a parameter in the LMI.

2. Preliminaries

2.1. Notations and Lur'e systems

Let $\mathbb{Z}$ and $\mathbb{Z}^+$ ($\mathbb{R}$ and $\mathbb{R}^+$) be the set of integer (real) numbers and positive integer (real) numbers including zero, respectively. Let $\mathbb{C}$ be the set of complex numbers. We use the notations $0_{(a \times b)}$ for $a \times b$ zero matrices, and $I_{(n)}$ for $n \times n$ identity matrices, where the dimension is neglected if it is straightforward. Moreover, the empty matrix is denoted by $\emptyset$.

Let ℓ be the space of all real-valued sequences $h : \mathbb{Z} \mapsto \mathbb{R}$. Let ℓ_2 be the space of real-valued square-summable sequences $h : \mathbb{Z}^+ \mapsto \mathbb{R}$. For any $\rho \in (0,1)$, we say that $h \in \ell_2^\rho$ if the sequence $\{\rho^{-k} h_k\}_{k=0}^\infty$ belongs to ℓ_2. Finally, for absolute-summable sequences $h : \mathbb{Z} \mapsto \mathbb{R}$, we define $\|h\|_1 = \sum_{k=-\infty}^\infty |h_k|$.

We say a proper real rational transfer function $G \in \mathbf{RH}_\infty$ if it has all poles inside the open unit disk in the complex plane; $G \in \mathbf{RL}_\infty$ if it has no pole on the unit disk. With the minimal state-space realisation, the transfer function is $G(z) = C(zI - A)^{-1}B + D$, or $G \sim \begin{bmatrix} A\,B; C\,D \end{bmatrix}$ in short. The expression $G^*(z)$ denotes the complex conjugate transpose of $G(z)$ at $|z| = 1$, i.e. $G^*(z) = G^T\left(\frac{1}{z}\right)$, where the superscript T indicates the transpose. Moreover, if a parameter ρ is involved in the variable, the complex conjugate transpose is expressed as $G^*(\rho, z) = G^T\left(\rho, \frac{1}{z}\right)$. In some cases, we write $G(\rho z)$ as $G(\rho, z)$ to stress the form of complex conjugate transpose.

A nonlinear operator $\Delta : \ell(\mathbb{Z}^+) \mapsto \ell(\mathbb{Z}^+)$ is said to be memoryless if there exists a map $N : \mathbb{R} \to \mathbb{R}$ such that $(\Delta v)_k = N(v_k)$, $\forall k \in \mathbb{Z}^+$. Assume that $\Delta(0) = 0$. The memoryless uncertainty Δ is said to be (sector) bounded, denoted by $\Delta \in [\underline{k}, \overline{k}]$ ($0 \leq \underline{k} < \overline{k} < \infty$), if $\underline{k}x \leq \Delta(x) \leq \overline{k}x, \forall x \in \mathbb{R}$. The uncertainty Δ is said to be slope-restricted, denoted by $\Delta \in S[\underline{k}, \overline{k}]$, if $\underline{k}(x_1 - x_2) \leq \Delta(x_1) - \Delta(x_2) \leq \overline{k}(x_1 - x_2), \forall x_1, x_2 \in \mathbb{R}$ and $x_1 \neq x_2$. The slope-restricted uncertainty is also sector bounded, but the reverse is not. Finally, the uncertainty Δ is said to be odd if $\Delta(-x) = -\Delta(x)$, $\forall x \in \mathbb{R}$.

Consider the Lur'e system in Figure 1. It is expressed as

$$v = f + Gw, \quad w = g + \Delta v.$$

The feedback interconnection is well-posed if the inverse map $(v, w) \mapsto (g, f)$ is causal in ℓ.

Definition 2.1. The feedback interconnection in Figure 1 is ℓ_2-stable if it is well-posed, and the signals $(v, w) \in \ell_2$ for any $(g, f) \in \ell_2$.

Definition 2.2 (Vidyasagar (2002)). The feedback interconnection in Figure 1 is globally exponentially stable with convergence rate ρ if there exists some $\rho \in (0,1)$ and $c \geq 1$ such

that when $g = 0$ and $f = 0$,

$$\|x_k\| \leq c\rho^k \|x_0\| \quad \forall k \geq 0, \ \forall x_0 \in \mathbb{R}^n. \tag{4}$$

Remark 1. Condition (4) is equivalent to the fact that the state x_k of G converges to zero exponentially with the rate ρ, i.e. $x_k \rho^{-k} \to 0$ as $k \to \infty$.

Henceforth, the convergence rate of the feedback interconnection in Figure 1 is referred to as $\rho_{\{G,\Delta\}}$.

Comparing the definitions above, ℓ_2-stability is an input-output relation, while exponential stability is an internal property. Therefore, we restate exponential stability in an input-output manner as follows.

Definition 2.3. The feedback interconnection in Figure 1 is ℓ_2^ρ-stable if it is well-posed, and the signals $(v, w) \in \ell_2^\rho$ for any $(g, f) \in \ell_2^\rho$.

Theorem 2.4. *For the Lur'e system in Figure 1, assume G is controllable and observable, and Δ is memoryless and slope-restricted. The unforced system is globally exponentially stable with rate ρ if and only if the forced system is ℓ_2^ρ-stable.*

Proof. The sufficiency can be proved in a similar way with Proposition 5 in Boczar et al. (2017, 2015), and the necessity is proved in outline in Appendix A. $\qquad \square$

2.2. Kalman conjecture for convergence analysis

Definition 2.5 (Nyquist value, K_N). The Nyquist value of a stable transfer function $G(z)$ is

$$K_N = \sup_K \{K > 0 : (1 - \tau K G(z))^{-1} \text{ is stable } \forall \tau \in [0, 1]\}.$$

Conjecture 2.6 (Kalman conjecture (Kalman, 1957)). *Let Δ be memoryless, and $\Delta \in S[0, K]$. The feedback interconnection between G and Δ is asymptotically stable if and only if $K < K_N$.*

Although the Kalman conjecture is false in general, it is valid for some systems (see Carrasco et al. (2016) for a general summary). We say that a system satisfies the Kalman conjecture when it is valid for this particular system.

We can translate the above definition and conjecture into the convergence analysis. Define the absolute convergence rate of the class of systems defined by the Lur'e system as

$$\rho^*_{\{G,K\}} = \sup_{\Delta \in S[0,K]} \{\rho_{\{G,\Delta\}}\}, \tag{5}$$

where $K < K_N$. A lower bound of the $\rho^*_{\{G,K\}}$ is given by the linear case as

$$\underline{\rho^*_{\{G,K\}}} = \max_{\tau \in [0,1]} \left\{ \left| \text{pole}\left(\frac{G}{1 - \tau K G} \right) \right| \right\}. \tag{6}$$

The Kalman conjecture for convergence analysis can be formulated by using definitions in (5) and (6) as follows:

Conjecture 2.7 (Kalman conjecture for convergence analysis). *For any $G \in \mathbf{RH}_\infty$, let $\Delta \in S[0, K]$ with $K < K_N$, then*

$$\rho^*_{\{G,K\}} = \underline{\rho^*_{\{G,K\}}}. \tag{7}$$

A counter-example for this conjecture will be provided in Section 3. It shows that the correct statement is $\rho^*_{\{G,K\}} \geq \underline{\rho^*_{\{G,K\}}}$, and it motivates the analysis below.

2.3. IQC framework for convergence analysis

In the last years, IQC and ρ-IQC analysis have been used to estimate an upper bound of $\rho^*_{\{G,K\}}$, denoted by $\overline{\rho^*_{\{G,K\}}}$. In this section, we will focus on the stability technique in the IQC framework, and a class of ρ-IQC multipliers defined in literature will be mentioned for comparison in the next section.

The relation below makes it possible to analyse exponential stability in the IQC framework.

Definition 2.8 (Boczar et al. (2017, 2015)). In Figure 2, with $\rho \in (0, 1)$, define $\Delta_\rho \equiv \rho_- \circ (\Delta \circ \rho_+)$, and $G_\rho(z) \equiv \rho_- \circ (G(z) \circ \rho_+)$, where the operators ρ_+ and ρ_- are defined as follows

$$\rho_+(w_\rho) \equiv \rho^k w_{\rho,k}, \quad \rho_-(y) \equiv \rho^{-k} y_k.$$

Due to the linearity of G, $G_\rho(z) = G(\rho z)$.

Lemma 2.9. Δ_ρ *is odd if and only if Δ is odd.*

Proof. For each given ρ, the operators ρ_+ and ρ_- are constant at each time instant k. As a result, Δ_ρ is a time-varying but memoryless operator. In addition, Δ_ρ is odd at each instant k if and only if Δ is odd. $\square$

Theorem 2.10 (Boczar et al. (2017, 2015); Hu and Seiler (2016)). *The system in Figure 1 is well-posed if and only if the scaled system in Figure 2 is well-posed. Furthermore, the system in Figure 1 is ℓ_2^ρ-stable if and only if the scaled system in Figure 2 is ℓ_2-stable.*

Henceforth, ℓ_2^ρ stability of the system in Figure 1 is studied by ℓ_2 stability of the scaled system in Figure 2, where an IQC should be constructed for the scaled uncertainty Δ_ρ at first.

Definition 2.11 (IQC Megretski and Rantzer (1997)). Let $\Pi(z)$ be a Hermitian (self-adjoint) bounded measurable operator. Then, for a bounded and causal operator $\Delta_\rho : \ell \mapsto \ell$, it is said to satisfy the IQC defined by Π, if

$$\int_{|z|=1} \begin{bmatrix} \widehat{v_\rho}(z) \\ \widehat{\Delta_\rho(v_\rho)}(z) \end{bmatrix}^* \Pi(z) \begin{bmatrix} \widehat{v_\rho}(z) \\ \widehat{\Delta_\rho(v_\rho)}(z) \end{bmatrix} dz \geq 0, \ \forall v_\rho \in \ell_2, \tag{8}$$

where $\widehat{v_\rho}$ and $\widehat{\Delta_\rho(v_\rho)}$ denote the z-transform of v_ρ and $\Delta_\rho(v_\rho)$ respectively.

Theorem 2.12 (Megretski and Rantzer (1997)). *For the feedback interconnection between G_ρ and Δ_ρ in Figure 2, let $G_\rho(z) \in \mathbf{RH}_\infty$, and Δ_ρ be a causal bounded operator. Assume that $\forall \tau \in [0, 1]$,*

(1) the feedback interconnection between G_ρ and $\tau \Delta_\rho$ is well-posed;
(2) the operator $\tau \Delta_\rho$ satisfies the IQC defined by Π;

6

(3) there exists $\varepsilon > 0$, such that

$$\begin{bmatrix} G_\rho(z) \\ I \end{bmatrix}^* \Pi(z) \begin{bmatrix} G_\rho(z) \\ I \end{bmatrix} \leq -\varepsilon I, \quad \forall |z| = 1. \tag{9}$$

Then, the system in Figure 2 is ℓ_2-stable.

In order to make the FDI (9) computable, the dynamic IQC $\Pi(z)$ should be factorised, and the KYP lemma Rantzer (1996) will be applied to convert the FDI to an LMI. The factorisation of an IQC $\Pi(z)$ is defined below.

Definition 2.13 (Scherer and Weiland (2011))**.** The pair $(\Psi(z), K_p)$ is said to be a factorisation of $\Pi(z) \in \mathbf{RL}_\infty$ if

$$\Pi(z) = \Psi^*(z) K_p \Psi(z), \tag{10}$$

where $K_p = K_p^T$ is constant, and $\Psi(z)$ is a stable LTI system.

The state-space representation of $\Psi(z)$ is in the form

$$\Psi(z) \sim \begin{bmatrix} A_\Psi & B_\Psi \\ C_\Psi & D_\Psi \end{bmatrix} = \begin{bmatrix} A_\Psi & B_{\Psi_1} & B_{\Psi_2} \\ C_\Psi & D_{\Psi_1} & D_{\Psi_2} \end{bmatrix}. \tag{11}$$

On the other hand, in the ρ-IQC approach, a class of multipliers is defined in Freeman (2018) to preserve the positivity of memoryless and slope-restricted uncertainties $\Delta : \ell_2^\rho \mapsto \ell_2^\rho$. With a fixed $\rho \in (0,1)$, the multiplier is in the form

$$M(\rho z) = \sum_{k \in \mathbb{Z}} \tilde{m}_k \rho^{-k} z^{-k} \tag{12}$$

where $\tilde{m}_0 > 0$. In addition, $\tilde{m}_k \leq 0 \; \forall k \neq 0$, or Δ is odd. Moreover, the ℓ_1 norm condition should be satisfied

$$\sum_{k \in \mathbb{Z}} |\tilde{m}_k| \min\{1, \rho^{-2k}\} < 2\tilde{m}_0, \tag{13}$$

where we can set $\tilde{m}_0 = 1$ without loss of generality. The notation here is slightly different from Freeman (2018), and refer to it for other details.

In Section 4, the details to estimate $\rho^*_{\{G,K\}}$ in the IQC framework will be discussed: the Zames-Falb IQC for the class of slope-restricted nonlinearities will be defined, and a unified factorisation will be provided to search noncausal FIR multipliers. This subset of Zames-Falb multipliers will be proved equivalent to (12) by variable change.

3. Motivating example

The Kalman conjecture, originally stated as a trivial theorem, has attracted much attention since the Sixties. In discrete time, it is known that it is true for first order systems, but fails for second order systems.

As the convergence analysis is mainly used in optimisation techniques, we show the first counterexample of the Kalman conjecture for convergence analysis. **In the existing counterexamples (Lessard et al., 2016; Scoy et al., 2018), the systems reach limit cycles where**

$\rho = 1$, **similar to the counterexamples given in (Carrasco et al., 2015; Heath et al., 2015), which means the system is not asymptotically stable, and there is no Zames-Falb multipliers in this case. In contrast, the system here converges with a rate $\rho < 1$ asymptotically, but slower than the linearised case. In short, we will show that $\rho^*_{\{G,K\}} < 1$ but $\rho^*_{\{G,K\}} > \underline{\rho^*_{\{G,K\}}}$ is possible in Conjecture 2.7. The possible difference between $\rho^*_{\{G,K\}}$ and $\underline{\rho^*_{\{G,K\}}}$ motivates to bound the worst-case convergence rate of this class of systems by utilising multipliers.**

We consider a Heavy-Ball algorithm, i.e.

$$x_{k+1} = x_k - \alpha \nabla f(x_k) + \beta(x_k - x_{k-1}), \tag{14}$$

where we assume that ∇f is a slope-restricted nonlinearity in $S[m,L]$ [2] with $0 < m < L$, and the parameters α and β are optimally selected as

$$\alpha = \frac{4}{(\sqrt{L} - \sqrt{m})^2}, \quad \beta = \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^2,$$

where $\kappa = L/m$. Then the heavy-ball algorithm can be described by the Lur'e system in Figure 1, where

$$G \sim \left[\begin{array}{cc:cc} 1+\beta & -\beta & -\alpha \\ 1 & 0 & 0 \\ \hdashline 1 & 0 & 0 \end{array}\right],$$

and $\Delta(x) = \nabla f(x)$. So far, the use of Zames-Falb multipliers is motivated by examples where the algorithm reaches a limit cycle when it is designed for strongly-convex quadratic functions (Lessard et al., 2016).

To the best of authors' knowledge, this is the first example in the literature where it is demonstrated that the asymptotic rate of convergence above the expected rate of convergence for the family of quadratic functions.

Let us take $m = 0.1$, $L = 36.1$, hence $\kappa = 361$, $\beta = 0.81$, and $\alpha = 0.1$. Moreover, let

$$\phi(x) = \nabla f(x) = \begin{cases} 12.6x & \text{if } x \geq 0; \\ 6.9x & \text{if } x < 0. \end{cases} \tag{15}$$

The slower asymptotic converge rate of the Lur'e system with Δ replaced by any linear gain in the interval $[0.1, 36.1]$ is given by 0.9.

Let us consider the initial condition

$$x_1 = -1, x_2 = \frac{19899351}{5(26860988 + \sqrt{446144163740329})},$$

then it generate a cyclic convergence defined by $x_3 > 0$, $x_4 > 0$, $x_5 < 0$, $x_6 < 0$, and $x_7 < 0$, and $x_8 = \lambda x_2$; since

$$((A + 6.9BC)^3(A + 12.6BC)^3)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

---

[2]Note that it is standard in optimisation to use this sector, but it can be trivially translated into a $[0,K]$ sector typically used in absolute stability literature via standard loop transformations.

8

where $\lambda = \frac{266558677 + \sqrt{446144163740329}}{500000000} \simeq 0.5753616$. **These values were found symbolically using Maple to satisfy the cycle convergence equations.** As a result, the a cyclic convergence is preserved with the rate $\rho = \lambda^{1/6} \simeq 0.91199 > 0.9$. The behaviour of the system is shown in Figure 4 as a numeric verification of the symbolic solution.
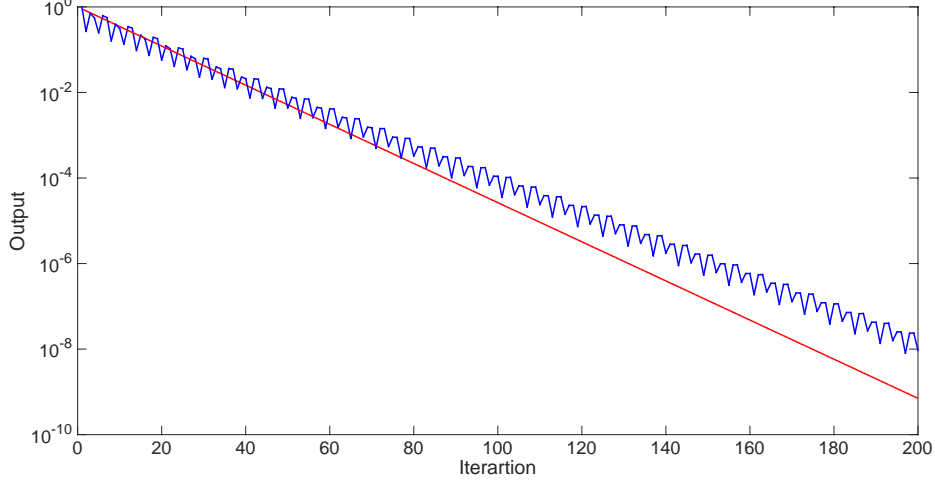

Figure 4. Decaying limit cycle in blue; the slower expected rate predicted by the analysis of convex quadratic functions in red.

This example shows that the heavy-ball algorithm can converge asymptotically at a slower rate than predicted by the analysis of the class of convex quadratic functions.

It is important to highlight that the original statement of the Kalman conjecture requires the nonlinearity to be differentiable everywhere, although this detail is normally ignored nowadays as the stability analysis (or the convergence analysis) deals with possible non-smooth nonlinearities. The nonlinearity can be "smoothed" around the non-differentiable point at the origin. This will cause the algorithm to reproduce the slower convergence rate when outside the "smoothed" region. However, the faster, asymptotic linear convergence is achieved once the iterates enter the smoothed region of the function.

As the slope condition is invariant to the scale, if a nonlinearity exhibits some local rate, then it is possible to find a nonlinearity where this behaviour is reproduced arbitrarily close to the origin. Loosely speaking, if we restrict our attention to smooth nonlinearities, slower convergence rates will be found arbitrarily closed to the origin.

As an example, Fig. 5 shows the behavior of the family of smooth nonlinearities with slope in the interval $[0.1, 36.1]$, where the nonlinearities are given by

$$\phi_\varepsilon(x) = \nabla f_\varepsilon(x) = \begin{cases} 12.6x - 2.85\varepsilon, & \textbf{if } x \geq \varepsilon; \\ \frac{5.7}{2\varepsilon}x^2 + 6.9x, & \textbf{if } 0 \leq x \leq \varepsilon; \\ 6.9x & \textbf{if } x < 0. \end{cases} \tag{16}$$

The iterates converge at the slower rate as long as $x \geq \varepsilon$ or $x < 0$. However, the faster linear convergence rate is achieved once $x \in [0, \varepsilon]$, i.e. the slope of the simulation is parallel to the worst linear case.
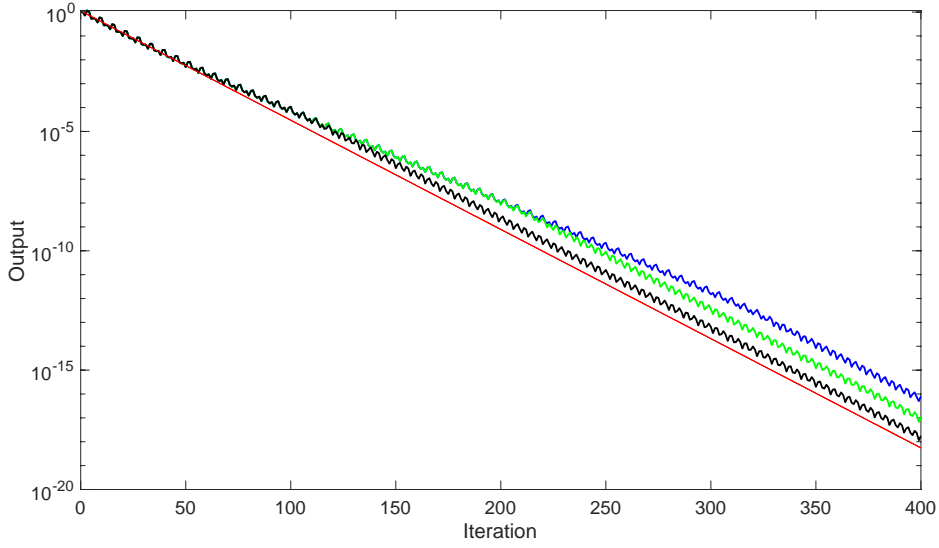
9

Figure 5. Simulations of the Lur'e system for ϕ_ε with $\varepsilon = 10^{-4}$ in black, $\varepsilon = 10^{-8}$ in green, and $\varepsilon = 10^{-12}$ in blue. In red, convergence rate for the linarised system with $k = 6.9$.

4. Zames-Falb multipliers for convergence analysis in IQC framework

As mentioned in Section 2.3, the convergence analysis can be conducted in the IQC framework on the scaled system in Figure 2. In this section, we consider the Zames-Flab IQC for the scaled uncertainty Δ_ρ, where the the multiplier belongs to a subclass of Zames-Falb multipliers whose ℓ_1 norm condition is penalised with ρ.

Definition 4.1. With a given $\rho \in (0,1)$, an LTI convolution operator $M : \ell_2(\mathbb{Z}) \mapsto \ell_2(\mathbb{Z})$ is said to be in the class $\mathscr{M}_\rho$ if its impulse response $m : \mathbb{Z} \mapsto \mathbb{R}$ satisfies the ℓ_1 norm condition $\sum_{k=-\infty}^{\infty} |m_k| \rho^{-|k|} < 2m_0$. Without loss of generality, we can set $m_0 = 1$.

Referring to the definition of the class of discrete-time Zames-Falb multipliers $\mathscr{M}$ (Willems & Brockett, 1968; Willems, 1971), it is clear $\mathscr{M}_\rho \subset \mathscr{M}$.

Lemma 4.2 (Zames-Falb IQC for Δ_ρ). *Assume the uncertainty Δ is memoryless and $\Delta \in S[0,K]$. The corresponding scaled uncertainty Δ_ρ satisfies the Zames-Falb IQC defined by Π as*

$$\Pi(z) = \begin{bmatrix} 0 & KM^*(z) \\ KM(z) & -(M(z) + M^*(z)) \end{bmatrix}, \tag{17}$$

where $M \in \mathscr{M}_\rho$ with either $m_k \leq 0, \forall k \neq 0$ or Δ is also odd.

Proof. The proof is linked with the argument in Freeman (2018). By setting $m_k = \tilde{m}_k \rho^{-k}$ $\forall k$, the multiplier $M(\rho z)$ in (12) will be independent with ρ, and be in the form of $M \in \mathscr{M}_\rho$. Then, the ℓ_1 norm condition of $M \in \mathscr{M}_\rho$ is proved by (13) in turn. $\square$

Here, we focus on the search of a wide subset of $\mathscr{M}_\rho$.

Definition 4.3. An LTI convolution operator $M : \ell_2(\mathbb{Z}) \mapsto \ell_2(\mathbb{Z})$ is said to be an FIR multiplier

10

$(M \in \mathcal{M}_{\rho,FIR})$ if it is in the form

$$M(z) = m_{-n_f}z^{n_f} + \cdots + m_{-1}z + m_0 + m_1 z^{-1} + \cdots + m_{n_b}z^{-n_b}, \tag{18}$$

where $m_{-i_f} \leq 0$ ($i_f = 1, \cdots, n_f$) and $m_{i_b} \leq 0$ ($i_b = 1, \cdots, n_b$), or Δ_ρ is odd. In addition, the impulse response of M should satisfy the ℓ_1 norm condition

$$\sum_{i_f=1}^{n_f} |m_{-i_f}|\rho^{-i_f} + \sum_{i_b=1}^{n_b} |m_{i_b}|\rho^{-i_b} < m_0, \tag{19}$$

where we can set $m_0 = 1$ without loss of generality.

By the condition (19), it is clear that $\mathcal{M}_{\rho,FIR} \subset \mathcal{M}_\rho$. Moreover, the causal part is with the backward-shift operators z^{-i_b}, and the anticausal part is with the forward-shift operators z^{i_f}. The FIR multiplier here is a restricted form of the original FIR multiplier in Carrasco et al. (2019); Wang et al. (2014).

The following result is bound the convergence rate of a Lur'e system. The result is developed to cover both techniques in the Literature by using an additional parameter μ.

Definition 4.4. Let $\mu \in \mathbb{R}$. Given $n_f \in \mathbb{Z}^+$, $n_b \in \mathbb{Z}^+$, let $i_f = 1, \cdots, n_f$, $i_b = 1, \cdots, n_b$. Set $\hat{n} = \max\{n_f, n_b\}$. Define the set of matrices

$$A_S = \begin{bmatrix} 0 & 0 \\ I_{(\hat{n}-1)} & 0 \end{bmatrix}_{(\hat{n} \times \hat{n})}, \quad B_S = \begin{bmatrix} 1 \\ 0_{(\hat{n}-1) \times 1} \end{bmatrix}_{(\hat{n} \times 1)},$$

$$C_{S1} = \begin{bmatrix} 0 & 0^\square \\ \frac{1}{\mu}I_{(n_b)} & \end{bmatrix}_{((n_b+1) \times \hat{n})}, \quad C_{S2} = \begin{bmatrix} 0 & 0^\triangle \\ \mathrm{diag}\left(\mu^{2i_f-1}\right) & \end{bmatrix}_{((n_f+1) \times \hat{n})}, \tag{20}$$

$$D_{S1} = \begin{bmatrix} 1 \\ 0_{(n_b \times 1)} \end{bmatrix}_{((n_b+1) \times 1)}, \quad D_{S2} = \begin{bmatrix} 1 \\ 0_{(n_f \times 1)} \end{bmatrix}_{((n_f+1) \times 1)},$$

where the matrices $0^\square$ and $0^\triangle$ are

$$0^\square = \begin{cases} \varnothing & \text{if } n_b \geq n_f \\ 0_{((n_b+1) \times (\hat{n}-n_b))} & \text{if } n_b < n_f \end{cases},$$

$$0^\triangle = \begin{cases} \varnothing & \text{if } n_f \geq n_b \\ 0_{((n_f+1) \times (\hat{n}-n_f))} & \text{if } n_f < n_b \end{cases}.$$

Theorem 4.5. *Let $\rho \in (0,1)$. Given $n_f \in \mathbb{Z}^+$, $n_b \in \mathbb{Z}^+$, let $i_f = 1, \cdots, n_f$, $i_b = 1, \cdots, n_b$. Set $\hat{n} = \max\{n_f, n_b\}$, and $\mu = 1$. For the feedback interconnection between G and Δ in Figure 1, let $G(z) \in \mathbf{RH}_\infty$, and $G(z) \sim \begin{bmatrix} A & B; & C & D \end{bmatrix}$; let Δ be memoryless and $\Delta \in S[0,K]$. Assume there exist a matrix $P = P^T$, such that*

$$\begin{bmatrix} \hat{A}^T P \hat{A} - P & \hat{A}^T P \hat{B} \\ \hat{B}^T P \hat{A} & \hat{B}^T P \hat{B} \end{bmatrix} + \begin{bmatrix} \hat{C}^T \\ \hat{D}^T \end{bmatrix} K_p \begin{bmatrix} \hat{C} & \hat{D} \end{bmatrix} < 0, \tag{21}$$

where K_p is given in (23) and the variables $\boldsymbol{m}_{-i_f} = [m_{-1} \; m_{-2} \; \cdots \; m_{-n_f}]^T$, $\boldsymbol{m}_{i_b} = [m_1 \; m_2 \; \cdots \; m_{n_b}]^T$ satisfy $\sum_{i_f=1}^{n_f} |m_{-i_f}|(\rho/\mu)^{-i_f} + \sum_{i_b=1}^{n_b} |m_{i_b}|(\rho\mu)^{-i_b} < m_0$, and set $m_0 = 1$.

Meanwhile, $m_{i_b} \leq 0$ and $m_{-i_f} \leq 0$, or Δ is odd. In addition, the matrices $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$ are

$$\hat{A} = \begin{bmatrix} \rho^{-1}A & 0 \\ B_{\Psi_1}C & A_\Psi \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \rho^{-1}B \\ B_{\Psi_2} + B_{\Psi_1}D \end{bmatrix}, \tag{22}$$

$$\hat{C} = \begin{bmatrix} D_{\Psi_1}C & C_\Psi \end{bmatrix}, \quad \hat{D} = D_{\Psi_2} + D_{\Psi_1}D,$$

where A_Ψ, $B_\Psi = \begin{bmatrix} B_{\Psi_1} & B_{\Psi_2} \end{bmatrix}$, C_Ψ, $D_\Psi = \begin{bmatrix} D_{\Psi_1} & D_{\Psi_2} \end{bmatrix}$ are in (24) with matrices A_S, B_S, C_{S1}, C_{S2}, D_{S1}, D_{S2} in (20). Then, a convergence rate of the unforced system in Figure 1 is ρ.

$$K_p = \begin{bmatrix} 0 & 0 & Km_0/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Km_{i_b} & 0 & 0 & 0 & 0 & 0 \\ Km_0/2 & Km_{i_b}^T & -m_0 & -m_{i_b}^T & 0 & 0 & 0 & 0 \\ 0 & 0 & -m_{i_b} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & Km_0/2 & Km_{-i_f}^T \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Km_0/2 & 0 & -m_0 & -m_{-i_f}^T \\ 0 & 0 & 0 & 0 & Km_{-i_f} & 0 & -m_{-i_f} & 0 \end{bmatrix}, \tag{23}$$

$$A_\Psi = \frac{1}{\mu}\begin{bmatrix} A_S & 0 \\ 0 & A_S \end{bmatrix}, \quad B_\Psi = \begin{bmatrix} B_S & 0 \\ 0 & B_S \end{bmatrix},$$

$$C_\Psi = \begin{bmatrix} C_{S1} & 0 \\ 0 & C_{S1} \\ C_{S2} & 0 \\ 0 & C_{S2} \end{bmatrix}, \quad D_\Psi = \begin{bmatrix} D_{S1} & 0 \\ 0 & D_{S1} \\ D_{S2} & 0 \\ 0 & D_{S2} \end{bmatrix}. \tag{24}$$

Remark 2. The additional parameter $\mu = 1$ is used in Theorem 4.5. In contrast, by setting $\mu = \rho$, Theorem 4.5 becomes a ρ-IQC analysis as provided Freeman (2018), but the details are not discussed here.

Proof. First, the matrices in (24) correspond to the state-space representation of the matrix function, called "lifting matrix",

$$\Psi(\mu, z) = \begin{bmatrix} 1 & 0 \\ \mu^{-i_b}\mathbf{Z}^{-i_b} & 0 \\ 0 & 1 \\ 0 & \mu^{-i_b}\mathbf{Z}^{-i_b} \\ 1 & 0 \\ \mu^{i_f}\mathbf{Z}^{-i_f} & 0 \\ 0 & 1 \\ 0 & \mu^{i_f}\mathbf{Z}^{-i_f} \end{bmatrix}. \tag{25}$$

With $\mu = 1$, it is equivalent to write $\Psi(\mu, z)$ as $\Psi(z)$. Then, it is straightforward to show that the pair $(\Psi(\mu, z), K_p)$ in (25) and (23) satisfies the definition of factorisation in (10) for the IQC in the form (17) with M in the form (18). Moreover, because M satisfies the conditions in Definition 4.3, M here belongs to the class $\mathcal{M}_\rho$. Then, according to Definition 2.8, Lemma 2.9 and 4.2, the scaled uncertainty Δ_ρ satisfies this IQC. This argument is equivalent to Condition 2 of Theorem 2.12.

Next, with $G(z) \sim [A \, B; \, C \, D]$, the scaled plant $G_\rho(z) = G(\rho z) \sim [\rho^{-1}A \, \rho^{-1}B; \, C \, D]$. Hence, $\Psi(\mu, z) \begin{bmatrix} G_\rho(z) \\ I \end{bmatrix} \sim \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix}$ holds. By the KYP lemma, the LMI (21) is equivalent to the FDI (9) with the corresponding IQC in Condition 3 of Theorem 2.12.

Then, as the assumptions on G and Δ ensure the well-posedness, the system between G_ρ and Δ_ρ is also well-posed, which corresponds to Condition 1 of Theorem 2.12.

Finally, based on the analysis above, the scaled system in Figure 2 is ℓ_2-stable. By Theorem 2.10, the original system in Figure 1 is ℓ_2^ρ-stable. Therefore, according to Theorem 2.4, the unforced system has the convergence ρ. $\qquad\square$

We call the pair $(\Psi(\mu, z), K_p)$ defined in (25) and (23) "lifting factorisation" or "factorisation by lifting".

Remark 3. The lifting factorisation $(\Psi(\mu, z), K_p)$ can be used as a unified structure of factorisation for the Zames-Falb IQC in (17) with causal, anticausal, noncausal multipliers $M \in \mathscr{M}_\rho$, where causal or anticausal multipliers can be obtained by setting $n_f = 0$ or $n_b = 0$ with the corresponding matrices being empty respectively.

In short, by the IQC analysis, everything keeps the same as in the conventional IQC case except that the ℓ_1 norm condition of the subset FIR Zames-Falb multipliers is penalised symmetrically on causal and anticausal parts with ρ. Then, the factorisation $(\Psi(\mu, z), K_p)$ with $\mu = 1$ defined in (25) and (23) can be used to construct the stability LMI (21). The minimum ρ obtained by Theorem 4.5 corresponds to $\overline{\rho^*_{\{G,K\}}}$.

5. Numerical results

In this section, we compare the results by causal, anticausal and noncausal multipliers via standard IQC analysis, i.e. $\mu = 1$. The examples are listed in Table 1 with other relating information. The Kalman conjecture is false for Ex.$1 - 4$, and true for Ex. 5. Here, we make the causal and anticausal steps equal in noncausal multipliers $(n_b = n_f = n_z)$. According to preliminary studies, small deterioration in results may occur with a large n_z (see (Carrasco et al., 2019, Fig.3) for instance), so n_z is tuned for each example.

Ex	$G(z)$	K	n_z
1	$-\dfrac{0.1z}{z^2 - 1.8z + 0.81}$	12	7
2	$-\dfrac{-0.08658z + 0.007162}{z^2 + 1.415z + 0.5523}$	10	3
3	$-\dfrac{z^3 - 1.95z^2 + 0.9z + 0.05}{z^4 - 2.8z^3 + 3.5z^2 - 2.412z + 0.7209}$	0.6	5
4	$-\dfrac{z^4 - 1.5z^3 + 0.5z^2 - 0.5z + 0.5}{4.4z^5 - 8.957z^4 + 9.893z^3 - 5.671z^2 + 2.207z - 0.5}$	3.5	4
5	$-\dfrac{10z^2 + 19z + 9}{100z^3 - 80z^2 + 17z - 1}$	3	1

Table 1. The plant G, the stable slope K of nonlinearity and the order n_z of multipliers in each example

The best bound of $\overline{\rho^*_{\{G,K\}}}$ are obtained via Theorem 4.5 by a bisection search starting from the initial range $(\rho^*_{\{G,K\}}, 1)$, where $\rho^*_{\{G,K\}}$ is given in (6). The software package CVX (Grant & Boyd, 2014) with the solver sdpt3 (Tütüncü, Toh, & Todd, 2003) is used to solve the LMI (21). The bound are listed in Table 2, where the best results of each example are in bold.

In the table, Ex. $1 - 5$ are with non-odd uncertainties, while Ex. $1_{odd} - 5_{odd}$ are with odd uncertainties. As the Kalman conjecture is false for Ex. $1 - 4$, ρ^* is not tight. In contrast, the Kalman conjecture is true for Ex. 5, so $\underline{\rho^*}$ is tight, and the multiplier technique is not needed.

Ex	$\underline{\rho^*_{\{G,K\}}}$	$\overline{\rho^*_{\{G,K\}}}$		
		Causal	Anticausal	Noncausal
1	0.900000	0.992474	invalid	**0.992465**
1_{odd}	0.900000	0.991441	invalid	**0.990514**
2	0.789886	0.934003	0.934985	**0.929675**
2_{odd}	0.789886	**0.863946**	0.906619	**0.863946**
3	0.942971	0.986340	0.991168	**0.982478**
3_{odd}	0.942971	0.971305	0.970435	**0.969681**
4	0.977129	invalid	0.985964	**0.985785**
4_{odd}	0.977129	0.998126	0.985792	**0.985724**
5	0.975367	**0.975368**	invalid	**0.975368**
5_{odd}	0.975367	**0.975368**	0.977501	**0.975368**

Table 2. Best bounds of $\overline{\rho^*_{\{G,K\}}}$ by causal, anticausal and noncausal multipliers compared with $\underline{\rho^*_{\{G,K\}}}$

In addition, the best estimated convergence rates with odd uncertainties are no greater than those with non-odd uncertainties, which is reflected by the more strict condition on multiplier parameters in the non-odd case. As for different multipliers, the introduction of anticausal steps are efficient to reduce the conservatism (also see (Freeman, 2018, Fig.2)). As a result, noncausal multipliers are less conservative in general, especially for Ex. $1, 4, 5$, where the search on causal and anticausal multipliers fails in some cases.

The similar conclusions are also inflected by the convergence rates with uncertainties that have different maximum slopes in Figure 6 and 7, where K starts from 0.01. In the figures, solid lines are for non-odd uncertainties, while dotted lines are for odd uncertainties. Meanwhile, the results by causal, anticausal and noncausal multipliers are in blue, green and red, respectively. In Figure 6, anticausal multipliers are seriously conservative, while causal and noncausal multipliers lead to similar less-conservative results. In Figure 7, all of the multipliers have satisfactory results with small K, but causal and anticausal multipliers become conservative for large K. In summary, noncausal multipliers are less conservative in both odd and non-odd cases. On the other hand, causal multipliers are more reliable than anticausal multipliers in general.
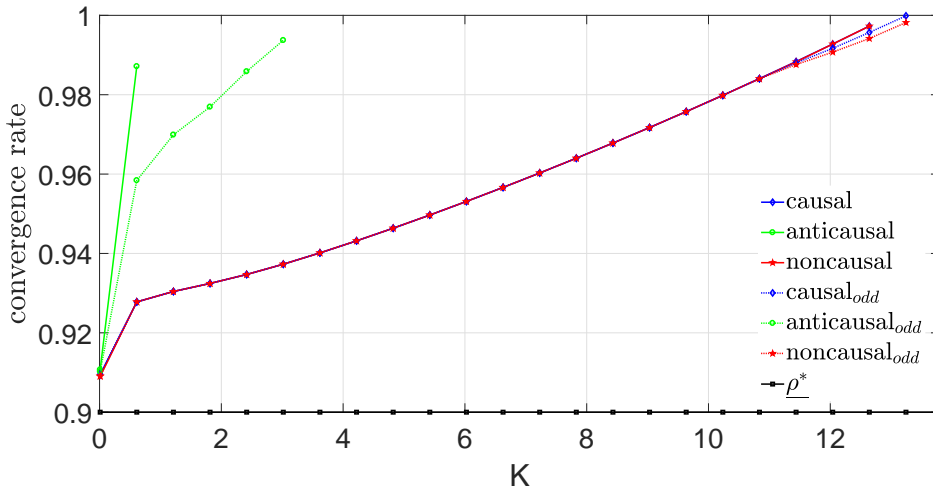

Figure 6. Best bounds of $\overline{\rho^*_{\{G,K\}}}$ with different K and multipliers compared with $\underline{\rho^*_{\{G,K\}}}$ of Ex.1
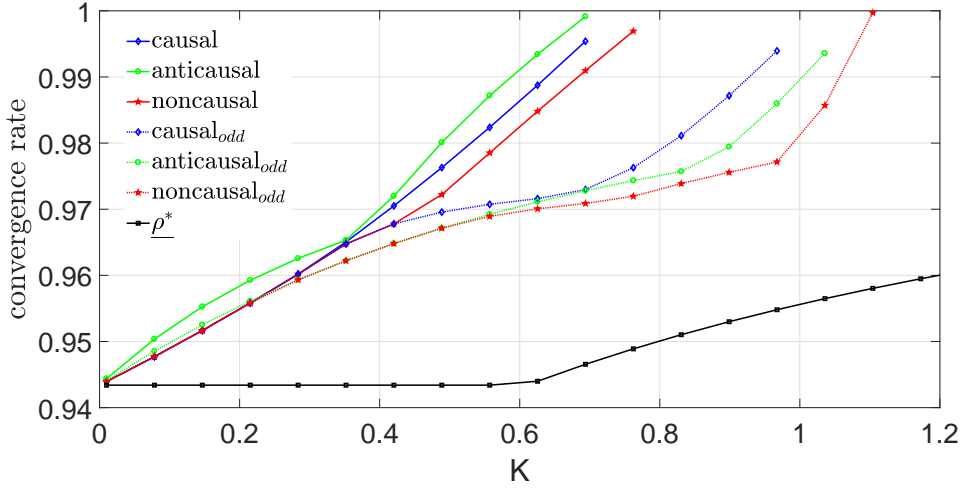
Figure 7. Best bounds of $\overline{\rho^*_{\{G,K\}}}$ with different K and multipliers compared with $\underline{\rho^*_{\{G,K\}}}$ of Ex.3

In short, for an arbitrary plant whose properties are unknown, it is less conservative to use noncausal multipliers to obtain its bounds of the convergence rates. If the given plant verifies the Kalman conjecture, the lower bound in Conjecture 2.7 is tight, and the multiplier techniques are not needed. However, it is an open question to determine whether a plant satisfies the Kalman conjecture or not.

We must highlight that we have found no counterexample of the Kalman conjecture for convergence rate with odd nonlinearities. However, numerical results do not achieve the value of the Kalman conjecture, hence it remains open whether the converge analysis for odd non-linearities is tight or not. Further research is required to clarify this question.

6. Conclusion

In this technical note, we provide novel insights in the convergence analysis of Lur'e systems. On one hand, we provide the first counterexample of the extension of the Kalman conjecture to the analysis of convergence rates, showing the significance of the existing techniques to bound the convergence rate. On the other hand, we have developed a convex search over the set of noncausal multipliers. Finally, the numerical examples indicate that noncausal multipliers are efficient to achieve less-conservative estimation of the upper bound convergence rates.

References

Altshuller, D. (2013). *Frequency domain criteria for absolute stability-a delay -integral-quadratic constraints approach*. London: Springer.

Badithela, A., & Seiler, P. (2019, July). Analysis of the heavy-ball algorithm using integral quadratic constraints. In *2019 American Control Conference (ACC)* (p. 4081-4085).

Boczar, R., Lessard, L., Packard, A., & Recht, B. (2017). Exponential stability analysis via integral quadratic constraints. *eprint arXiv: 1706.0133*. Retrieved from `http://iviv.org/abs/1706.01337`

Boczar, R., Lessard, L., & Recht, B. (2015, Dec). Exponential convergence bounds using integral quadratic constraints. In *2015 54th IEEE Conference on Decision and Control (CDC)* (p. 7516-7521).

Carrasco, J., Heath, W. P., & de la Sen, M. (2015). Second-order counterexample to the discrete-time kalman conjecture. In *2015 european control conference (ecc)* (p. 981-985).

Carrasco, J., Heath, W. P., Zhang, J., Ahmad, N. S., & Wang, S. (2019). Convex searches for discrete-time zames-falb multipliers. *IEEE Transactions on Automatic Control*, 1-1.

Carrasco, J., Turner, M. C., & Heath, W. P. (2016). Zames-Falb multipliers for absolute stability: from O'Shea's contribution to convex searches. *European Journal of Control*, *28*, 1-19.

Cyrus, S., Hu, B., Scoy, B. V., & Lessard, L. (2018, June). A robust accelerated optimization algorithm for strongly convex functions. In *2018 Annual American Control Conference (ACC)* (p. 1376-1381).

Fazlyab, M., Ribeiro, A., Morari, M., & Preciado, V. (2018). Analysis of optimization algorithms via integral quadratic constraints: Nonstrongly convex problems. *SIAM Journal on Optimization*, *28*(3), 2654-2689.

Freeman, R. A. (2018, June). Noncausal Zames-Falb multipliers for tighter estimates of exponential convergence rates. In *2018 Annual American Control Conference (ACC)* (p. 2984-2989).

Grant, M., & Boyd, S. (2014, March). *CVX: Matlab software for disciplined convex programming, version 2.1.* http://cvxr.com/cvx.

G.Zames, & P.L.Falb. (1968). Stability conditions for systems with monotone and slope-restricted nonlinearities. *SIAM J. Control*, *6*, 89-108.

Heath, W. P., Carrasco, J., & de la Sen, M. (2015). Second-order counterexamples to the discrete-time kalman conjecture. *Automatica*, *60*, 140 - 144. Retrieved from http://www.sciencedirect.com/science/article/pii/S0005109815002812

Hu, B., & Seiler, P. (2016, Nov). Exponential decay rate conditions for uncertain linear systems using integral quadratic constraints. *IEEE Transactions on Automatic Control*, *61*(11), 3631-3637.

Kalman, R. E. (1957). Physical and mathematical mechanisms of instability in nonlinear automatic control systems. *Transactions of ASME*, *79*, 553-566.

Lessard, L., Recht, B., & Packard, A. (2016). Analysis and design of optimization algorithms via integral quadratic constraints. *SIAM Journal on Optimization*, *26*(1), 57-95.

Lurie, A. I., & Postnikov, V. N. (1944). On the stability theory of control systems. *Russian Prikl. Matem. i Mekh.*, *8*.

Megretski, A., & Rantzer, A. (1997). System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, *42*, 819-830.

Michalowsky, S., Scherer, C., & Ebenbauer, C. (2020). Robust and structure exploiting optimisation algorithms: an integral quadratic constraint approach. *International Journal of Control*, 1-24.

Nelson, Z. E., & Mallada, E. (2018, June). An integral quadratic constraint framework for real-time steady-state optimization of linear time-invariant systems. In *2018 Annual American Control Conference (ACC)* (p. 597-603).

Rantzer, A. (1996). On the Kalman-Yakubovich-Popov lemma. *Systems & Control Letters*, *28*(1), 7 - 10.

Scherer, C., & Weiland, S. (2011). Linear matrix inequalities in control. In *The control systems handbook, second edition: Control system advanced methods* (pp. 24/1–24/30). London: CRC Press.

Scoy, B. V., Freeman, R. A., & Lynch, K. M. (2018). The fastest known globally convergent first-order method for minimizing strongly convex functions. *IEEE Control Systems Letters*, *2*(1), 49-54.

Tütüncü, R. H., Toh, K. C., & Todd, M. J. (2003). Solving semidefinite-quadratic-linear programs using sdpt3. *MATHEMATICAL PROGRAMMING*, *95*, 189–217.

Vidyasagar, M. (2002). *Nonlinear systems analysis*. NJ: Society for Industrial and Applied Mathematics.

Wang, S., Heath, W. P., & Carrasco, J. (2014, Dec). A complete and convex search for discrete-time noncausal FIR Zames-Falb multipliers. In *53rd IEEE Conference on Decision and Control* (p. 3918-3923).

Willems, J., & Brockett, R. (1968, October). Some new rearrangement inequalities having application in stability analysis. *IEEE Transactions on Automatic Control*, *13*(5), 539-549.

Willems, J. C. (1971). Stability criterion obtained using multipliers. In *The analysis of feedback*

systems (p. 136-155). Cambridge: M.I.T. press.

Zhang, J., Seiler, P., & Carrasco, J. (2019). *Noncausal fir zames-falb multiplier search for exponential convergence rate.*

Appendix A. Proof of Theorem 2.4 (Outline)

The proof follows (Vidyasagar, 2002, Theorem 6.3.46) and the theorems therein, which link internal stability to input-output ℓ-stability in continuous time. Here we relate the relations with the convergence rate in discrete time. The argument in continuous time can be converted to discrete time trivially, which is not stressed here.

In order to keep consistent with the notations in Vidyasagar (2002), we consider the system $-G \sim \begin{bmatrix} A\ B;\ -C\ 0 \end{bmatrix}$. The minus signs indicate negative feedback structure in (Vidyasagar, 2002, Fig.6.4), and the D matrix can be set to 0 without loss of generality. Moreover, because G is linear and stable, the disturbance signal f in Figure 1 is assumed to be zero without loss of generality. Then, the feedback interconnection in Figure 1 can be expressed as

$$x_{k+1} = Ax_k + Bg_k + B\Delta(v_k), \quad y_k = -Cx_k. \tag{A1}$$

Necessity: The expression (A1) satisfies the general form (6.3.7) in Vidyasagar (2002). In addition, the conditions (6.3.16, 17) are satisfied, because the state-space matrices of G are finite and Δ is slope-restricted.

Following the proof and notations in (Vidyasagar, 2002, Theorem 6.3.15), (6.3.19) implies the exponentially stability condition (4) of the unforced system where $c = \frac{\beta}{\alpha}$, $\rho = \frac{1}{2\beta^2}$ ($0 < \alpha < \beta$ are constant). Then, in (6.3.32), $\|x_k\| \leq \frac{W_k}{\alpha}$, where $W_k \leq h_k$, and h_k is the output of a first order system with input $\|g_k\|$ and pole $\frac{-1}{2\beta^2}$. In other words, the exponential rate limited by this transfer function is $\rho = \frac{1}{2\beta^2}$. Henceforth, let the input to this transfer function is $g \in \ell_2^\rho$, then the solution $h \in \ell_2^\rho$, thus $W \in \ell_2^\rho$ and $x \in \ell_2^\rho$. Moreover, according to (6.3.33), $v \equiv y \in \ell_2^\rho$. Next, because Δ is memoryless and slope-restricted, and $g \in \ell_2^\rho$, the signal $w \in \ell_2^\rho$. This proves that the forced system is ℓ_2^ρ-stable.