Aizerman Conjectures for a class of multivariate positive systems

Ross Drummond, Chris Guiver and Matthew C. Turner Member, IEEE

Abstract—The Aizerman Conjecture predicts stability for a class of nonlinear control systems on the basis of linear system stability analysis. The conjecture is known to be false in general. Here, a number of Aizerman conjectures are shown to be true for a class of internally positive multivariate systems, under a natural generalisation of the classical sector condition and, moreover, guarantee positivity in closed loop. These results are stronger and/or more general than existing results. The paper relates the obtained results to other, diverse, results in the literature.

I. INTRODUCTION

Consider the feedback interconnection depicted in Figure 1, where $G(s)$ is a linear-time-invariant (LTI) system and $\Phi(\cdot,\cdot) : \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}^m$ is a static, possibly time-varying, nonlinear element. Such a configuration is commonly referred to as a Lur’ë (or Lurie or Lurye) system and has had much attention devoted to it, since this class of systems arises naturally in many areas of science and engineering [1], [2]. In the single-input single-output case, meaning that $u(t)$ and $y(t)$ in Figure 1 are scalars, the nonlinearity $\Phi = \phi$ is typically assumed to satisfy a sector condition

$$\alpha \leq \frac{\phi(y)}{y} \leq \beta \quad \forall y \in \mathbb{R}, \ y \neq 0 \quad (1)$$

for given $\alpha \leq \beta$. Aizerman conjectured in 1949 in [3] that the system in Figure 1 will be stable if the set of linear systems, formed by replacing the nonlinearity $\phi$ with a linear gain $k$, were themselves stable for all $k \in [\alpha, \beta]$. This conjecture is

Remarkably, a version of the Aizerman Conjecture, known as the complex (or generalised) Aizerman Conjecture, is true, and dates back to the work of Hinrichsen and Pritchard [4]. The well-known Circle Criterion can be derived as a consequence of the complex Aizerman Conjecture; see, for example, [5]. However, the method loses much of its appeal in the complex case, and the conclusions may be conservative.

The more familiar form of the Aizerman Conjecture, when $k$ is restricted to be real, has garnered much interest, and the academic literature on the subject is vast, as highlighted by numerous references in the 2006 survey paper [1]. Indeed, the field of absolute stability theory arguably arose from the initial studies on the Aizerman Conjecture, with early results in the 1950s by Soviet scholars demonstrating the conjecture to be true, up to some assumptions, when $G(s)$ is a second order system (see [6], [7] and the references therein), but false in general, with counterexamples presented in, for example, [8] and [9] (see also [6], [10] for more recent results). However, for the reasons given above, much effort has been devoted to identifying situations in which the Aizerman Conjecture is true. Of these, the most relevant are those where certain positivity assumptions are made on the linear part of the system including [11], [12], [13] — these will be discussed later in the paper.

Building upon these results, here a version of the Aizerman Conjecture is shown to be true (in a sense made precise in Theorem 1) for a class of multivariate positive Lur’e systems. Moreover, the hypotheses are particularly simple to verify and the proofs particularly short. To the best of the authors’ knowledge, the results given here are stronger and/or more general than other comparable results in the literature.

The note is organised as follows. Section II describes the class of systems considered, and contains the main result, Theorem 1. Section III contains further background and seeks to contextualise the work by relating it to relevant known results in the literature. Brief conclusions appear in Section IV.
A. Notation

Notation is mainly standard, but the reader’s attention is drawn to the following. For a real matrix (vector) \( M \), the notation \( M \geq 0 \), \( M > 0 \) and \( M \gg 0 \) means that \( M \) has non-negative elements, non-negative elements and is not equal to the zero matrix (vector), or positive elements, respectively. The symbols \( \leq \), \( < \) and \( \ll \) are defined similarly. It is also convenient to define

\[
\mathbb{R}^{n \times m}_+ = \{ M \in \mathbb{R}^{n \times m} : M \geq 0 \}
\]

which comprises so-called non-negative matrices (vectors).

A square matrix \( M \) is called Hurwitz if every eigenvalue of \( M \) has negative real part. The spectral abscissa (the maximum of the real part of the eigenvalues of \( M \)) is denoted \( s(M) \). For a vector \( v \geq 0 \), \( v_M \) denotes the value of smallest element, while \( v_M \) denotes the largest. Further, \( \| \cdot \| \) denotes the Euclidean norm of a vector, or for a matrix, the norm induced by the Euclidean norm. An \( n \times m \) matrix with unity elements is denoted \( \mathbf{1}_{n \times m} \).

For a transfer function \( G(s) \), the \( \mathcal{H}_\infty \) norm, \( \| G \|_\infty \), is defined as

\[
\| G \|_\infty = \sup_{\omega \in \mathbb{R}} \| G(j\omega) \| = \sup_{\omega \in \mathbb{R}} \bar{\sigma}(G(j\omega))
\]

where \( \bar{\sigma}(M) \) denotes the maximum singular value of \( M \).

The \( \mathcal{L}_\infty \) norm of a locally essentially bounded signal \( z \) is defined as

\[
\| z \|_{\mathcal{L}_\infty(0,T)} := \text{ess sup}_{0 \leq \tau \leq T} \max_{0 \leq t \leq T} | z_t(\tau) |
\]

This note will deal with the well-studied notions of positive systems and positive stability, as in, for example [14, 15], in addition to the familiar notions of asymptotic and exponential stability. Indeed, the system of ordinary differential equations

\[
\dot{x} = f(x(t)) \quad f(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n
\]

is said to be a positive system, or just positive, if \( x(t) \geq 0 \) for all \( t \geq 0 \) whenever \( x(0) \in \mathbb{R}^n_+ \). Similarly, the system \( \dot{x} = f(x(t)) \) is said to be positively globally exponentially stable (pGES) if it is positive and there exist \( \eta, \lambda > 0 \) such that every solution \( x \) of \( \dot{x} = f(x(t)) \) satisfies

\[
\| x(t) \| \leq \eta e^{-\lambda t} \| x(0) \| \quad \forall t \geq 0 \quad \forall x(0) \in \mathbb{R}^n_+
\]

II. THE AZEIMBERG CONJECTURE FOR POSITIVE SYSTEMS

A. Positive Linear Systems

Consider first the positive linear system of differential equations

\[
\dot{x} = Mx
\]

where \( M \in \mathbb{R}^{n \times n} \) is Metzler, that is \( M_{ij} \geq 0 \) for all \( i \neq j \). Metzler matrices are also called essentially non-negative [14 p. 146], [15 p.30] or quasi positive [16 p.60]. They play the same role in nonnegative differential equations as nonnegative matrices in difference equations (discrete-time).

The following facts are well known

Fact 1. [4] is a positive system if, and only if, \( M \) is a Metzler matrix. Further, if \( M \) is additionally Hurwitz, then there exists \( v \in \mathbb{R}^n_+ \), \( v \gg 0 \), such that \( v^TM \ll 0 \).

Proofs of these claims may be found in, for instance, [14 Theorem 3.1, p.146] and [15 Lemma 2.2, p.31], respectively. We highlight that stable positive linear systems admit linear Lyapunov functions constructed in terms of vectors \( v \in \mathbb{R}^n_+ \) as above; see, for instance, [17]. These ideas will be employed frequently throughout the paper.

B. Positive Lust’e Systems

Consider the class of systems depicted in Figure 1 with \( r_1 = 0 \) and \( r_2 = 0 \). The interconnection of the linear element \( \dot{x} = Ax + Bu \)

\[
\begin{aligned}
\dot{x} &= Ax + Bu \\
y &= Cx
\end{aligned}
\]

(5)

where \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \) and \( C \in \mathbb{R}^{p \times n} \). As usual, \( u(t), x(t) \) and \( y(t) \) in (5) denote the input, state and output variables, which take values in \( \mathbb{R}^n, \mathbb{R}^n \) and \( \mathbb{R}^p \), respectively. We recall that (5) is called

- internally positive if \( x \geq 0 \) for all \( x(0) \geq 0 \) and all \( u \geq 0 \)
- externally positive if \( y \geq 0 \) whenever \( x(0) = 0 \) and for all \( u \geq 0 \)

Here \( x \geq 0 \) means \( x(t) \geq 0 \) for all \( t \geq 0 \), and similarly for \( u \) and \( y \). These systems are well-studied in the literature; see, for instance [18, 19]. Internal positivity is equivalent to \( A \) being Metzler, and \( B \) and \( C \) being nonnegative [18 Theorem 2], and external positivity is equivalent to the impulse response \( h(t) = Ce^AMB \) taking nonnegative values [18 Theorem 1]. Exponential positivity does not stipulate positivity of the state, and is harder to characterise further [20, 21], but has been considered in the work of [22] to some extent.
stability. Consequently, the pGES estimate (3) and sector condition (7) are only required to hold for nonnegative initial states and arguments, respectively; see also Remark 1 below. We note that (6) need not be a monotone control system, in the sense of [24] (at least in the usual nonnegative orthant \( \mathbb{R}^n_+ \)), even when \( \Phi \) satisfies (7).

Before stating results for Lur’e systems, a useful stepping-stone is to consider positivity and pGES for linear feedback systems. It transpires that, for positive systems, verifying the linear positivity and stability conditions amount to checking the properties of matrices at the “end-points” of the sector, as described in the next lemma.

For the remainder of the work, we let \( \Sigma_1, \Sigma_2 \in \mathbb{R}^{m \times p} \) with \( \Sigma_1 \leq \Sigma_2 \) be given.

**Lemma 1.** Consider (5) with \( B, C \geq 0 \), and the hypothesis (H) \( A + B\Sigma C \) is Hurwitz and Metzler for all \( \Sigma \) such that \( \Sigma_1 \leq \Sigma \leq \Sigma_2 \).

The following statements hold.

1. **Hypothesis** (H) is equivalent to \( A + B\Sigma_1 C \) being Metzler and \( A + B\Sigma_2 C \) being Hurwitz.
2. Hypothesis (H) is necessary and sufficient for the linear feedback system of (5) and \( y = \Sigma u \), that is, with \( \Phi(z,t) = \Sigma z \), to be pGES for all \( \Sigma_1 \leq \Sigma \leq \Sigma_2 \).

**Proof.** (1) That (H) is sufficient for the claimed properties is clear, as \( \Sigma = \Sigma_1 \) and \( \Sigma = \Sigma_2 \) are included in (H). Conversely, for \( \Sigma \) such that \( \Sigma_1 \leq \Sigma \leq \Sigma_2 \), the hypothesis \( B, C \geq 0 \) yields that

\[ A + B\Sigma_1 C \leq A + B\Sigma C \leq A + B\Sigma_2 C \]

Hence, \( A + B\Sigma C \) is Metzler, as \( A + B\Sigma_1 C \) is. Moreover, a consequence of [16, Corollary 4.3.2] is that \( s(A + B\Sigma C) \leq s(A + B\Sigma_2 C) \leq 0 \) — the last inequality following by hypothesis. Therefore, \( A + B\Sigma C \) is Hurwitz, and since \( \Sigma \) was arbitrary, the proof is complete.

Statement (2) follows immediately from Fact 1. \( \square \)

**Theorem 1 (Positive Aizerman).** Consider the Lur’e system (6) with \( B, C \geq 0 \). If \( A + B\Sigma_1 C \) is Hurwitz and \( A + B\Sigma_2 C \) is Hurwitz, then (6) is pGES for every \( \Phi \in \text{Sector}[\Sigma_1, \Sigma_2] \).

The above theorem shows that the positive Aizerman Conjecture is true. Namely, for positive systems, the hypothesis (H) — a necessary and sufficient condition for positivity and global exponential stability (pGES) of the linear feedback system (3) for all feedback gains \( \Sigma \) such that \( \Sigma_1 \leq \Sigma \leq \Sigma_2 \) — implies that the Lur’e system (6) is itself pGES for all \( \Phi \) in the same sector, that is, in the sense of (7). Note that \( \Phi \) itself is not required to be Metzler.

Before proving Theorem 1 further commentary is given.

**Remark 1.**

(a) In the SISO case, meaning \( n = p = 1 \), writing \( \phi := \Phi, \sigma_1 := \Sigma_1 \leq \Sigma_2 := \sigma_2 \) the sector condition (7) may be rewritten in the more familiar form

\[ \sigma_1 \leq \frac{\phi(z,t)}{z} \leq \sigma_2 \quad \forall z > 0, \quad \forall t \in \mathbb{R}^+ \]

The results proved here thus hold for the SISO case, but actually hold for the (possibly non-square) MIMO case.

(b) **Sign conventions.** A positive feedback convention has been adopted in this paper. However, no assumptions on the sign(s) of the sector data \( \Sigma_1 \) and \( \Sigma_2 \) are made. Often \( \Phi \) will satisfy a so-called one-sided sector condition; that is \( \Phi \in \text{Sector}[-\Sigma_2, \Sigma_2] \) for \( \Sigma_2 \geq 0 \), but negative feedback will be used. This is actually equivalent to the current configuration by taking \( \Phi \in \text{Sector}[\Sigma_2, 0] \); the positive feedback convention does not limit generality. Furthermore, the nonlinear feedback \( \Phi \) is defined as a function of all real arguments, but since (6) is required to be a positive system, the sector condition (7) is only required for nonnegative arguments.

(c) If, in addition to the hypotheses of Theorem 1, the matrix \( A + B\Sigma_2 C \) is assumed irreducible, then the exponential rate of decay in the pGES estimate for (6) may be chosen equal to \( s(A + B\Sigma_2 C) \), and this is the smallest decay rate which “works” for all \( \Phi \) satisfying the sector condition (7). Recall that a square matrix \( M \in \mathbb{R}^{n \times n} \) is called irreducible if it is not reducible. A matrix is reducible if it is similar, via a permutation matrix, to an upper block triangular matrix with non-zero (and non-trivial) block diagonal terms. \( \square \)

**Theorem 2.** Consider the Lur’e system (6) with \( B, C \geq 0 \). Then (6) is a positive system for all \( \Phi \in \text{Sector}[\Sigma_1, \Sigma_2] \) if, and only if, \( A + B\Sigma_1 C \) is Metzler.

**Proof.** That \( A + B\Sigma_1 C \) is Metzler is necessary for (6) to be positive is clear from Fact 1 as the linear function \( \Phi(z,t) = \Sigma z \) belongs to \( \text{Sector}[\Sigma_1, \Sigma_2] \).

For sufficiency, suppose that \( A + B\Sigma_1 C \) is Metzler, let \( \Phi \in \text{Sector}[\Sigma_1, \Sigma_2] \) be fixed, and consider the \( i \)-th state equation associated with the Lur’e system (6)

\[ \dot{x}_i = A_{ii} x_i + \sum_{i \neq j = 1}^n A_{ij} x_j + (B \Phi(Cx,t))_i \]

Since \( x(0) \in \mathbb{R}^n_+ \), were some component \( x_i(t) \) of \( x \) the first to become negative, then, by continuity of solutions, there exists \( t_1 \geq 0 \) such that \( x_i(t_1) = 0, x_i(t_1) \geq 0, \) and \( Cx(t) \geq 0 \) on \([0, t_1]) \). From the sector condition (7) and the assumed nonnegativity, it follows that

\[ \dot{x}_i(t_1) = (A + B\Sigma_1 C)_{ii} x_i(t_1) + \sum_{i \neq j = 1}^n (A + B\Sigma_1 C)_{ij} x_j(t_1) \]

\[ = \sum_{i \neq j = 1}^n (A + B\Sigma_1 C)_{ij} x_j(t_1) \]

However, by the Metzler property of \( A + B\Sigma_1 C \) and because \( x_j(t_1) \geq 0 \), the above inequality implies that \( \dot{x}_i(t_1) \geq 0 \) and thus \( x_i \) can, in fact, never become negative. \( \square \)

**Remark 2.** An alternative proof to the above lemma is obtained by using a differential inequality. For brevity, we only sketch the details, by noting that the solution \( v \) of \( \dot{v} = f(v) := Av + B\Sigma_1 Cv \) from \( x_0 \in \mathbb{R}^n_+ \) is nonnegative by Fact 1 and as \( A + B\Sigma_1 C \) is assumed Metzler. The function \( f \) is quasi-monotone increasing, see [25, p.94]. It can be shown that the solution of \( x(t) \) satisfies \( \dot{x} \geq f(x) \) and the desired claim that \( x \geq v \geq 0 \) follows from [25, Theorem VIa, p.96]. \( \square \)
Fact. There exist \( M \) such that \( \Phi \) is useful. In this section it will be shown that, roughly, the hypotheses hold. In short, one has a direct estimate of the convergence rate.

Let \( C \) be given and let \( r := B \Sigma_2 r_1 + r_2 \). Let \( v \in \mathbb{R}^n_+ \), \( v \gg 0 \) and \( \varepsilon > 0 \) such that

\[
\begin{align*}
v^T M &\leq -\varepsilon v^T \\
&\leq \varepsilon v^T M \quad (9)
\end{align*}
\]

Since \( v \) is strictly positive, we have

\[
\begin{align*}
v_m \| z \| &\leq v^T z \leq \varepsilon M \| z \| \quad \forall z \in \mathbb{R}^n_+ \\
&\leq \varepsilon v^T M \quad (10)
\end{align*}
\]

Routine calculations invoking \( 8 \) and \( 9 \) give that

\[
\frac{d}{dt} e^{\varepsilon t} v^T x(t) = e^{\varepsilon t} v^T x(t) + e^{\varepsilon t} v^T \dot{x}(t) \leq e^{\varepsilon t} (v^T + v^T M) \\
&\leq 0 \quad \text{almost all } t \geq 0
\]

Since \( t \mapsto e^{\varepsilon t} v^T x(t) \) is nonnegative valued, \( 10 \) gives

\[
\begin{align*}
e^{\varepsilon t} v_m \| x(t) \| &\leq e^{\varepsilon t} v^T x(t) \leq v^T x(0) \leq \varepsilon M \| x(0) \| \quad \forall t \geq 0
\end{align*}
\]

from which pGES follows. \( \square \)

If \( A + B \Sigma_2 C \) is assumed irreducible then by, for example, Theorem 3.4 there exists \( v \in \mathbb{R}^n_+ \), \( v \gg 0 \) such that

\[
v^T (A + B \Sigma_2 C) = s(A + B \Sigma_2 C) v^T
\]

In short, one has a direct estimate of the convergence rate.

The following loop-shifting corollary is an immediate consequence of Theorem 1 and follows by replacing \( A \) and \( \Phi \) by \( A + BK \) and \( (z, t) \mapsto (\Phi(z, t) - K z) \), respectively.

**Corollary 1.** Consider the Lur’e system with \( B, C \geq 0 \). Let \( K \in \mathbb{R}^{m \times p} \) be such that \( A + BK \) is Metzler. If \( A + B(K + \Sigma) \) and \( A + B(K + \Sigma) \) is Hurwitz, then for every \( \Phi \) such that \( (z, t) \mapsto (\Phi(z, t) - K z) \in \text{Sector}[\Sigma_1, \Sigma_2] \) the Lur’e system \( 6 \) is pGES.

**C. An exponential ISS result**

Consider the system in Figure 1 where \( r_1 : \mathbb{R}_+ \to \mathbb{R}^p \) and \( r_2 : \mathbb{R}_+ \to \mathbb{R}^n \) are piecewise continuous signals modelling exogenous inputs. This leads to the Lur’e system

\[
\dot{x} = Ax + B \Phi (Cx + r_1, t) + r_2
\]

In this section it will be shown that, roughly, the hypotheses of Theorem 1 are sufficient for the stronger stability notion of exponential input-to-state stability (ISS) of \( 11 \), provided that the state \( x \) is nonnegative. For this purpose, the following lemma is useful.

**Lemma 3.** Consider the Lur’e system \( 11 \) with \( B, C \geq 0 \). If \( A + B \Sigma_1 C \) is Metzler, \( \Phi \in \text{Sector}[\Sigma_1, \Sigma_2] \), \( x(0) \geq 0 \), \( r_1 \geq 0 \) and \( B \Sigma_1 r_1 + r_2 \geq 0 \), then \( x(t) \geq 0 \) for all \( t \geq 0 \).

**Proof.** The proof is similar to that of Lemma 2 or alternatively by arguing as in Remark 2. \( \square \)

**Proposition 1.** Imposing the notation and assumptions of Theorem 7 there exist \( \Gamma, \gamma > 0 \) such that, for all \( x(0) \geq 0 \), \( r_1, r_2 \) with \( r_1, B \Sigma_1 r_1 + r_2 \geq 0 \) for all \( t \geq 0 \), the solution \( x \) of \( 11 \) satisfies \( x(t) \geq 0 \) and

\[
\| x(t) \| \leq \Gamma (e^{-\gamma t} \| x(0) \| + \| (r_1, r_2) \|_{\mathcal{L}^\infty(0,t)}) \quad \forall t \geq 0
\]

If \( A + B \Sigma_2 C \) is irreducible, then \( \gamma \) above may be chosen equal to \(-s(A + B \Sigma_2 C) > 0\).

**Proof.** Let \( x(0) \in \mathbb{R}^n_+ \), \( r_1, r_2 \) and \( \Phi \in \text{Sector}[\Sigma_1, \Sigma_2] \) be as in the statement of the result. By Lemma 3 it follows that \( x(t) > 0 \) for all \( t \geq 0 \). Equation \( 11 \) can be re-written as

\[
\dot{x} = (A + B \Sigma_2 C) x + B \Sigma_2 r_1 + r_2 \\
+ B \Phi (Cx + r_1, t) - \Sigma_2 (Cx + r_1)
\]

Note that the final term on the right hand side of \( 13 \) is nonpositive, and hence the variation of parameters formula entails that \( x \) admits the estimate, for all \( t \geq 0 \):

\[
0 \leq x(t) \leq e^{M t} x(0) + \int_0^t e^{M (t-\tau)} r(\tau) \, d\tau
\]

where \( M := A + B \Sigma_2 C \) and \( r := B \Sigma_2 r_1 + r_2 \). Let \( v \in \mathbb{R}^n_+ \), \( v \gg 0 \) and \( \varepsilon > 0 \) be as in \( 9 \). Applying \( v^T \) to both sides of the above, and invoking \( 10 \), yields that

\[
v_m \| x(t) \| \leq v^T x(t) \leq e^{-\varepsilon t} v^T x(0) + \int_0^t e^{-\varepsilon (t-\tau)} v^T r(\tau) \, d\tau
\]

\[
\leq e^{-\varepsilon t} v_m \| x(0) \| + \frac{1}{\varepsilon} \| v^T r \|_{\mathcal{L}^\infty(0,t)}
\]

for all \( t \geq 0 \), from which the estimate \( 12 \) follows. \( \square \)

**D. Stability in the large and global asymptotic stability**

Theorem 1 provides sufficient conditions for positivity and global exponential stability of the Lur’e system \( 6 \). Since \( 6 \) includes linear systems as a special case, global exponential stability is qualitatively the best expected in general.

A necessary condition to avoid linear instability is evidently that \( s(A + B \Sigma_2 C) \leq 0 \), and the situation wherein \( s(A + B \Sigma_2 C) < 0 \) has been considered in Theorem 1. Here it is demonstrated that other (weaker) stability notions are guaranteed, under certain assumptions, when \( s(A + B \Sigma_2 C) = 0 \). Even in the linear setting, as the situation with nontrivial Jordan blocks indicates, for instance

\[
\begin{pmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{pmatrix} =
\begin{pmatrix}
0 & 1 \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
\]

so that

\[
\begin{pmatrix}
x_1(t) \\
x_2(t)
\end{pmatrix} =
\begin{pmatrix}
0 & 1 \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
x_1(0) \\
x_2(0)
\end{pmatrix}
\quad \forall t \geq 0
\]

global stability notions cannot be expected without suitable additional assumptions. The additional assumption presently imposed is that \( A + B \Sigma_2 C \) is irreducible, see Remark 2.

**Proposition 2.** Consider the Lur’e system \( 6 \) with \( B, C \geq 0 \). Assume further that \( A + B \Sigma_1 C \) is Metzler, that \( A + B \Sigma_2 C \) is irreducible, and that \( s(A + B \Sigma_2 C) = 0 \). The following statements hold.

1. The square matrix appearing in the first equation above is not irreducible.
(1) For every $\Phi \in \text{Sector}[\Sigma_1, \Sigma_2]$ the Lur’e system \((6)\) is positively stable in the large, meaning that it is a positive system and there exists $\Gamma > 0$ such that
\[\|x(t)\| \leq \Gamma \|x(0)\| \quad \forall \ t \geq 0 \quad \forall \ x(0) \in \mathbb{R}^n.\]

(2) Assume that $A + B\Sigma_2 C$ is Hurwitz. For every $\Phi \in \text{Sector}[\Sigma_1, \Sigma_2]$ such that
- $\sup_{t \in \mathbb{R}} \Phi(z, t) < \Sigma_2 z$ for all $z \in \mathbb{R}^p$, $z \neq 0$, and;
- $z \mapsto \sup_{t \in \mathbb{R}} v^T B\Psi(z, t)$ is bounded from above by a continuous function which is zero at zero and negative for nonzero arguments. Here $v^T$ is such that $v^T (A + B\Sigma_2 C) = 0$ and $\Psi(z, t) := \Phi(z, t) - \Sigma_2 z$; the solution $x$ of \((6)\) satisfies $x(t) \to 0$ as $t \to \infty$ for every $x(0) \in \mathbb{R}^n$.

Although statement \((1)\) guarantees that the zero equilibrium of \((6)\) enjoys certain stability properties, the hypotheses on $A + B\Sigma_2 C$ ensure that there is a strictly positive $w \in \mathbb{R}^n$ such that $(A + B\Sigma_2 C)w = 0$. In particular, $w$ is another equilibrium of \((6)\) when $\Phi(z) = \Sigma_2 z$. The simplest situation wherein the bullet-pointed hypotheses on $\Phi$ are satisfied is when $B$ has no zero columns, $\Phi = \Phi(z)$ is independent of $t$, is continuous in $z$, and satisfies $\Phi(z) < \Sigma_2 z$ for all nonzero $z$. For time-varying $\Phi$, note that the supremum in the first bullet-point above is understood componentwise.

**Proof.** Let $\Phi \in \text{Sector}[\Sigma_1, \Sigma_2]$. It follows from Lemma \((2)\) that $x(t) \geq 0$ for all $t \geq 0$. A consequence of $A + B\Sigma_2 C$ being Metzler, irreducible and having zero spectral absicass is that there exists $v^T \gg 0$ with $v^T (A + B\Sigma_2 C) = 0$ and so
\[v^T e^{(A + B\Sigma_2 C)t} = v^T \quad \forall \ t \geq 0.\]

Statement \((1)\). It remains to prove the bound for $\|x(t)\|$. The proof is the same as that of Theorem \((1)\) but taking $\varepsilon = 0$ in \((9)\).

Statement \((2)\). The following proof draws inspiration from \((5)\) proof of Theorem \((5)\). Since statement \((1)\) holds, for every $x(0) \in \mathbb{R}^n$, the solution $x$ of \((6)\) is bounded by some $\rho > 0$. To show the claimed convergence, we prove that
\[\lim_{t \to \infty} \sup_{t \in \mathbb{R}} \|Cx(t)\| = 0\]
which implies that $x(t) \to 0$ as $t \to \infty$, in light of
\[\dot{x}(t) = (A + B\Sigma_1 C)x(t) + B(\Phi(Cx(t), t) - \Sigma_1 Cx(t))\]
as $A + B\Sigma_1 C$ is assumed Hurwitz and the hypotheses on $\Phi$. Seeking a contradiction, suppose that \((15)\) fails. Thus, there exists a strictly increasing, unbounded sequence $(t_k)_{k \in \mathbb{N}}$ and $\varepsilon \in (0, \|C\|/\rho)$ such that $\|Cx(t_k)\| \geq 2\varepsilon > 0$ for all $k \in \mathbb{N}$. Since $x$ is bounded, it follows from \((5)\) and \((7)\) that $\dot{x}$ is bounded, and hence $x$ is uniformly continuous. Therefore, there exists $\delta > 0$ such that
\[\|Cx(t)\| \geq \varepsilon \quad \forall \ t \in [t_k, t_k + \delta] \quad \forall \ k \in \mathbb{N}.\]

Passing to a subsequence if needed, we may assume that $t_{k+1} \geq t_k + \delta$ for all $k \in \mathbb{N}$. Set $\mathcal{M} := \{z \in \mathbb{R}^p : \varepsilon \leq \|z\| \leq \|C\|/\rho\}$, which is non-empty and compact. Recall the notation $\Psi(z, t) := \Phi(z, t) - \Sigma_2 z$.

By hypothesis the function $z \mapsto \sup_{t \in \mathbb{R}} v^T B\Psi(z, t)$ is bounded from above by a continuous, real-valued function which is negative for non-zero arguments, denoted $g$, say. Hence, there exists $\eta > 0$ such that
\[\sup_{z \in \mathcal{M}, t \in \mathbb{R}} v^T B\Psi(z, t) \leq \sup_{z \in \mathcal{M}} g(z) = -\eta < 0\]

Now applying $v^T$ to both sides of the variation of parameters formula for $x$ between $t_k$ and $t_{k+1}$ for $k \in \mathbb{N}$, and invoking \((13)\), that $Cx(\tau) \in \mathcal{M}$ for $\tau \in [t_k, t_k + \delta]$ and \((16)\), gives
\[v^T x(t_{k+1}) = v^T x(t_k) + \int_{t_k}^{t_{k+1}} v^T B\Psi(Cx(\tau), \tau) d\tau \leq v^T x(t_k) + \int_{t_k}^{t_{k+1}} v^T B\Psi(Cx(\tau), \tau) d\tau \leq v^T x(t_k) - \delta \eta \quad \forall \ k \in \mathbb{N}.
\]
The above inequality, for sufficiently large $k$, contradicts the nonnegativity of $v^T x(t)$. The proof is complete. $\square$

### E. Maximal elements for sectors

Here a key assumption of the present work is considered in more depth, namely the linear positivity and stabilisability assumption \((H)\). Observe that the sectors $\text{Sector}[\Sigma_1, \Sigma_2]$ are nested, in the sense that if $\Sigma_0 \leq \Sigma_1$ and $\Sigma_0 \leq \Sigma_3$, then $\text{Sector}[\Sigma_1, \Sigma_2] \subseteq \text{Sector}[\Sigma_0, \Sigma_3]$. Hence, a natural question is what is the “biggest” sector possible under which \((H)\) holds? The stability aspect of this question can be addressed by appealing to the well-known concept of the stability radius, dating back to \((25)\), and \((27)\) for positive systems.

For simplicity, assume that $A$ is Metzler and Hurwitz and that the transfer function $G$ associated with \((5)\) is nonzero. In particular, if $B, C \geq 0$, then \((27)\) Theorem \((5)\) yields that $A + B\Delta C$ is Hurwitz for all $\Delta \in \mathbb{R}^{m \times p}$ with $\|\Delta\| < 1/\|G(0)\|$. This estimate is sharp as the next lemma demonstrates.

**Lemma 4.** Consider \((5)\) with $A$ Metzler and Hurwitz, and $B, C \geq 0$. If $G(0) \neq 0$, then there exists rank-one $\Delta \in \mathbb{R}^{m \times p}$ such that $\|\Delta\| = 1/\|G(0)\|$ and zero is an eigenvalue of $A + B\Delta C$.

**Proof.** Let $v \in \mathbb{R}^m$ be such that $\|v\| = 1$ and $\|G(0)v\| = \|G(0)\|$. It follows that $v \in \mathbb{R}^n$ as $G(0) = C(-A)^{-1}B \geq 0$. Note that $(-A)^{-1}Bv \neq 0$. Define
\[\Delta := \frac{1}{\|G(0)\|^2} v(G(0)v)^T\]

which is evidently real, nonnegative and rank one. It is routine to verify that $\|\Delta\| = 1/\|G(0)\|$, that $w := (-A)^{-1}Bv \neq 0$, and, finally, that zero is an eigenvalue of $A + B\Delta C$ as
\[(A + B\Delta C)w = -Bv + B\Delta G(0)v = 0\]

Although Lemma \((4)\) provides an explicit definition of $\Delta$, it requires finding $v \in \mathbb{R}^n$ such that $\|G(0)v\| = \|G(0)\|$. In the SISO case, $\Delta$ is simply given by $\Delta = 1/\|G(0)\|$.

To use Lemma \((4)\) as a design tool requires a relationship between spectral absicass and componentwise orderings. These objects interact nicely with one another, in the sense that for Metzler $M_1, M_2$, it follows that if $M_1 \leq M_2$, then $s(M_1) \leq s(M_2)$. However, some care needs to be taken when seeking to infer the strict inequality that $M_1 < M_2$. 


implies $s(M_1) < s(M_2)$, which is false in general as upper triangular matrices show, but is true if $M_1$ is irreducible; see, for example [16 Corollary 4.3.2].

Thus, in light of Lemma 4, $A + BΣC$ is Hurwitz for all $Σ < Δ$ such that $A + BΣC$ is Metzler and irreducible. The hypothesis [II] cannot hold with $Σ_2 ≥ Δ$ and, further, if $A + BΔC$ is irreducible, then $s(A + BΣC) > 0$ for all $Σ > Δ$. In the MIMO case, these considerations cannot be applied to determine the stability of $A + BΣC$ for $Σ ∈ ℝ^{m×p}$ which satisfy $Σ ≤ Δ$ and $Σ ≥ Δ$.

III. Connections to Other Work

The Aizerman Conjecture is known to be false in general but various papers have identified particular classes of systems where it holds true. This section compares the new conditions derived in this paper to those already available in the literature, with summarising observations given in Table I.

A. A real Aizerman conjecture

There is overlap between Theorem 1 and [11 Theorem 1], which also presents a real Aizerman Conjecture for positive Lur’e systems. The work [11] only considers the case of diagonal nonlinearities, essentially meaning $Φ_i(z, t) = φ_i(z_i, t)$ for all $1 ≤ i ≤ m$ and $z ∈ ℝ^m$ with components $z_i$, and only concludes positive global asymptotic stability, but the ideas are otherwise similar to those used presently. The work [11] is brief, and does not consider the other facets considered here — exponential ISS in in Proposition 1 or the other stability considerations in Proposition 2.

B. Externally positive systems

Significant work on the Aizerman Conjecture for externally positive systems (recall, meaning the impulse response is non-negative valued) has been undertaken by Gil’, dating back to the 1980s, see [12 Chapter 6] and the references therein. The result [12 Theorem 6.3.1] shows that, for given $Q ∈ ℝ^{m×m}$, the Lur’e system (6) is GES for all $Φ$ such that

\[-Q|z| ≤ Φ(z, t) ≤ Q|z| \quad ∀ z ∈ ℝ^m, \forall t ≥ 0 \quad (17)\]

if, and only if, $\det(P(s) - L(s))Q$ is a Hurwitz polynomial, where $G(s) = P^{-1}(s)L(s)$. The result [12 Theorem 6.3.1] is different to the situation considered here, as [12 Theorem 6.3.1] does not require the Lur’e system (6) to be positive (which makes it more general), but does not address when (4) is positive — a natural requirement in many applied settings.

C. Nonnegative Lur’e systems

There is some overlap with the results proved here and those in [13]. The paper [13] considers stability, in various senses, of the forced positive (there called nonnegative) Lur’e systems, in the SISO case. Although stability of the zero equilibrium is considered [13], so that there is overlap between Theorem 1 and Proposition 2, and [13 Theorem 4.4]; the emphasis of that work is on the existence and stability of a nonzero equilibrium, which arises naturally in many ecological and biochemical contexts. Indeed, in that sense the work [13] is more in the spirit of positive dynamical systems, and considers so-called trichotomies of stability as in [28]. Another difference is that [13] considers positive feedback connections (only), meaning the nonlinear term maps $ℝ_+ → ℝ_+$.

D. Stability radii and the real supremum value property

One approach to the Aizerman Conjecture is to first consider additive, structured perturbations

\[ \dot{x} = Ax + BΓ[Cx] \quad (18)\]

of the unperturbed or nominal differential equation $\dot{x} = Ax$. Here $Γ[\cdot]$ in (18) is a placeholder for a number of different classes of perturbation, from matrix multiplication to a nonlinear function $Γ[ CZ ] = Φ(C z)$. In this light, it is clear that (18) encompasses the Lur’e system (6). So-called stability radii are a tool for determining local robustness, that is, determining the maximal bound for which all perturbations “within that bound” will preserve some property, in this case, stability. (For brevity in this discussion we are not precise with what is meant by stability.) Hinrichsen and Pritchard introduced stability radii for a number of perturbation classes (see [30 Section 6]), and a key finding is, unsurprisingly, that different perturbation classes have different stability radii in general.

In this perspective, the real and complex Aizerman Conjectures, roughly, ask when does stability for all perturbations of a certain linear type ensure stability for all perturbations of a corresponding nonlinear type? Thus, in the language of stability radii, the real Aizerman Conjecture is that the so-called “real static nonlinear stability radius”, denoted $r_{R, Φ}$, equals the “real linear stability radius”, denoted $r_R$. Note that $r_{R, Φ} ≤ r_R$ always holds, since a linear perturbation can be viewed as a nonlinear perturbation, but not conversely. The analysis in [4 Example 4.1] of a counterexample to the real Aizerman Conjecture proposed in [9] shows that the ratio $r_R/r_{R, Φ}$ can be arbitrarily large, so that the real Aizerman Conjecture can fail “dramatically”. In other words, whilst [15] may be stable for some fixed $A, B$ and $C$ and “large” linear, real perturbations $Γ_1[\cdot]$, there are arbitrarily small, real nonlinear perturbations $Γ_2[\cdot]$ which destabilise (18).

As stated in the Introduction, the complex Aizerman conjecture is true and, in the current perspective, is true because the complex linear stability radius $r_C$ satisfies the (nontrivial) inequality $r_C ≤ r_{R, Φ}$. Put differently, if every complex feedback gain in a (complex) ball of feedback gains is stabilising, then all nonlinear feedbacks in the same “ball” are stabilising. However, the strict inequality $r_C < r_{R, Φ}$ is possible and, in this case, the complex Aizerman conjecture is conservative.

Therefore, in light of the known bounds $r_C ≤ r_{R, Φ} ≤ r_R$, one approach to the real Aizerman Conjecture is to establish situations wherein $r_C = r_R$. In this case, the hypotheses of the real conjecture imply that the hypotheses of the complex conjecture hold which is true. A sufficient condition for $r_R = r_C$ is the so-called real supremum value property, namely, that

\[ \|G\|_∞ = \|G(jω)\| \quad \text{and} \quad G(jω) ∈ ℝ^{p×m} \quad (19)\]

for some $ω ∈ ℝ$. The real supremum value property is satisfied for certain classes of systems, such as those listed in [29 Example 3.7] — including internally positive and
Table I: Comparable real Aizerman Conjecture-type results for multivariate (MIMO) Lur’e systems.

<table>
<thead>
<tr>
<th>Class of system</th>
<th>Class of Nonlinearity</th>
<th>Stability properties guaranteed</th>
<th>Positivity guaranteed</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internally positive</td>
<td>Norm bounded</td>
<td>Exponential, input-to-state, asymptotic</td>
<td>× (no assumptions on sign of feedback)</td>
<td>Section II-D</td>
</tr>
<tr>
<td>Externally positive</td>
<td>Sector bounded</td>
<td>Exponential [12, Theorem 6.3.2].</td>
<td>× (no assumptions on internal properties)</td>
<td>Section II-B</td>
</tr>
<tr>
<td>Internally positive</td>
<td>Sector bounded</td>
<td>Exponential, input-to-state, asymptotic</td>
<td>✓</td>
<td>Theorem I, Proposition 1</td>
</tr>
</tbody>
</table>

E. The complex Aizerman conjecture

Here we demonstrate how the stability assumptions of Theorem 1 are equivalent to those of the complex Aizerman Conjecture in the SISO case, but not in the MIMO case, where our conditions are more general. Note that the complex Aizerman Conjecture only ensures stability, and does not address positivity, of $\Phi$.

For simplicity, assume that $\Phi$ is independent of $t$. To make the connection, we need to centre the sector conditions. For which purpose, set $M := (\Sigma_2+\Sigma_1)/2$, $D := (\Sigma_2-\Sigma_1)/2 \geq 0$. The stability component of hypothesis (11) is equivalent to

$$A + BS\Sigma C$$ is Hurwitz for all $-D \leq M \leq D$ \hspace{1cm} (20)

Similarly, it follows that $\Phi \in \text{Sector}[\Sigma_1, \Sigma_2]$ if, and only if,

$$-Dy \leq \Phi(y) - My \leq Dy \ \forall \ y \in \mathbb{R}^p$$ \hspace{1cm} (21)

The inequalities in (21) may be extended to all $y \in \mathbb{R}^p$ by (re)defining $\Phi(y) = My$ for $y \in \mathbb{R}^p$, $y \geq 0$. This is unproblematic when the Lur’e system (6) is positive, as nonnegative solutions are independent of how $\Phi$ is defined for nonpositive arguments.

To invoke complex Aizerman Conjectures requires norm conditions. In particular, by [29, Theorem 5.1], the hypothesis

$$A + BS\Sigma C$$ is Hurwitz for all $\|\Sigma - M\| \leq \|D\|$ \hspace{1cm} (22)

(which is a natural generalisation of (20)) guarantees that the Lur’e system (6) is GES for all $\Phi$ such that

$$\|\Phi(y) - My\| \leq \|D\|\|y\| \ \forall \ y \in \mathbb{R}^p$$ \hspace{1cm} (23)

(Note that it is, in fact, required that (22) holds for all complex $\Sigma$, but this can be relaxed to all real $\Sigma$ under the usual assumption that $A + BS\Sigma C$ is Metzler and $B, C \geq 0$ by the real supremum value property.)

In the SISO $(m = p = 1)$ case the quantities $M, D, \Sigma,$ and $\Sigma$ are all scalar, and the conditions (20) and (22) are both equivalent to $A + BS\Sigma C$ being Hurwitz for every real $\Sigma$ in the interval $[M - D, M + D]$. Moreover, here both (21) and (23) are equivalent to $|\Phi(y) - My| \leq D|y|$ for all $y \in \mathbb{R}$.

However, in the MIMO $(m, p > 1)$ case, it is routine to verify that (22) implies (20), and that (23) holds $\|\Sigma - M\| \leq \|D\|$ for both cases which is false. In particular, the complex Aizerman Conjecture results listed above are not applicable when only (20) and (21) are assumed. Intuitively, the condition (20) requires stability for all $\Sigma - M$ only in the “directions” determined by $-D \leq \Sigma - M \leq D$. The condition (22) requires stability for all $\Sigma - M$ in all directions, as determined by a norm — a stronger requirement.

F. Comparison to Zames-Falb multipliers

Zames-Falb multipliers may be used to predict stability of the Lur’e system in Figure 1 when the the nonlinearity $\Phi$ satisfies the stronger requirement that it is time-invariant and slope restricted. In the SISO case, this equivalent to

$$\alpha \leq \frac{\phi(z_1) - \phi(z_2)}{z_1 - z_2} \leq \beta \ \forall \ z_1, z_2 \in \mathbb{R}, \ z_1 \neq z_2$$ \hspace{1cm} (24)

for some $\alpha < \beta$. It appears ([32], [33]) that Zames-Falb multipliers are the least conservative method for guaranteeing stability of a Lur’e system under the assumption (24). Furthermore, in the MIMO case, when $m = p$ and assuming a “repeated scalar” structure for $\Phi$, that is

$$\Phi(z) = [\phi(z_1), \phi(z_2), \ldots, \phi(z_m)]' \ \forall \ z \in \mathbb{R}^m$$ \hspace{1cm} (25)

where $\phi$ satisfies (24), similar results can be obtained as for the SISO case ([34], [35]). However, the success of the Zames-Falb approach hinges on a search over Zames-Falb multipliers which can be complex and time-consuming ([35], [36]). Furthermore, Zames-Falb results are rather difficult to use for controller synthesis.

The work in this paper provides an alternative to Zames-Falb multipliers when the linear systems $G(s)$ are internally positive, and enables both pGES and positivity to be established. Table I lists some example MIMO systems which are all internally positive and, by Lemma 2, will result in positivity when connected in the manner depicted in Figure 1. Hence Theorem 1 applies. It is assumed that $\Phi$ is such that it belongs to Sector $[0, \sigma_2 I_m]$ (where $p = m$ for simplicity), or satisfies the slope conditions (24) and (25) with $\alpha = 0$ and $\beta = \sigma_2$.

Table I gives the maximum $\sigma_2$ for which stability can be ascertained using various approaches: the standard Circle Criterion, the Lyapunov-based approach of Park [37], and also that of Zames-Falb. Here the Zames-Falb multipliers are computed using the MIMO method of [38] (approaches such as [35] could equally be used). Some observations are in order:

- The Circle Criterion, as expected, provides the most conservative results, but the class of nonlinearities for which it
Table II: Linear data for example MIMO Lur’e systems.

<table>
<thead>
<tr>
<th>Example</th>
<th>Circle Criterion</th>
<th>Park</th>
<th>Zames-Falb</th>
<th>Theorem 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9158</td>
<td>0.9236</td>
<td>0.9236</td>
<td>0.9236</td>
</tr>
<tr>
<td>2</td>
<td>0.7997</td>
<td>2.0220</td>
<td>2.0221</td>
<td>2.0221</td>
</tr>
<tr>
<td>3</td>
<td>19.2764</td>
<td>89.8987</td>
<td>89.8987</td>
<td>89.8999</td>
</tr>
</tbody>
</table>

Table III: Maximum value of $\sigma_2$ for which stability can be numerically verified, according to various different approaches.

<table>
<thead>
<tr>
<th>Example</th>
<th>$\sigma_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9158</td>
</tr>
<tr>
<td>2</td>
<td>0.7997</td>
</tr>
<tr>
<td>3</td>
<td>19.2764</td>
</tr>
</tbody>
</table>

IV. Conclusion

This paper has shown that a suite of Aizerman Conjectures hold for a class of multivariate, positive nonlinear control systems; essentially ensuring positivity and various nonlinear stability notions depending on positivity and stability assumptions on the plant and a linear sector of matrices. In its simplest form, global exponential stability is guaranteed if two matrices at the extremes of the sector are both Hurwitz and Metzler. The contribution of the current work to other related literature was also discussed.

References
