Analysis of Lurie Systems with Magnitude Nonlinearities and Connections to Neural Network Stability Analysis

Carl R. Richardson, Matthew C. Turner and Steve R. Gunn

***Abstract*—This paper considers the interconnection of a continuous time linear time-invariant system and a multivariable magnitude nonlinearity. A number of different quadratic constraints are established for the magnitude nonlinearity and then used to derive stability criteria based on quadratic and Lurie-type Lyapunov functions. The new stability criteria are cast as matrix inequalities and in some cases solved using semi-definite programming. Connections are made between the magnitude nonlinearity and neural network activation functions, such as the rectified linear unit (ReLU) and leaky ReLU, effectively allowing the stability criteria derived here to be used to analyse interconnections of dynamical systems and neural networks. Using the positive homogeneity property, shared by the (leaky) ReLU and magnitude functions, mild conditions are also established to show that the existence of a unique equilibrium point is sufficient for local and global stability to be equivalent. Finally, the new global stability criteria are tested on several numerical examples, including a Hopfield network with 100 states and neurons, and compare favourably with competing criteria from the literature.**

***Index Terms*—Lyapunov methods, neural networks, semi-definite programming, stability of nonlinear systems**

I. INTRODUCTION

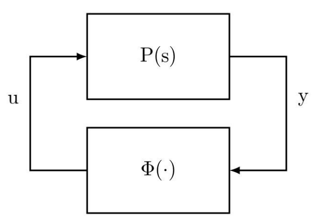


Fig. 1. Lurie system with static nonlinearity

Lurie systems are feedback interconnections involving a linear time-invariant (LTI) system and a, normally static, nonlinear element (Fig. 1). They have been studied extensively in the control systems literature due to their relevance to a number of practical problems and because there are several elegant and computationally efficient ways of analysing their stability: the Circle and Popov Criteria [1]–[3] are routinely taught to students of nonlinear control and the techniques of Zames and Falb [4], Park [5] and others [6], [7] are natural successors to these approaches. Over recent years, a resurgence of interest in Lurie systems has occurred, due to the recognition that dynamical systems involving neural networks (NNs) can often be modelled in the Lurie framework [8]. Indeed recent analyses of such systems have either adopted established approaches [9]–[13] or have derived new but related theoretical results [14]–[19]. Furthermore, in [20], [21], piecewise affine systems have been written in Lurie form, significantly expanding the class of systems covered.

The assumptions normally made in the study of Lurie systems are that the nonlinearity, $\Phi(\cdot)$, is sector bounded and/or slope-restricted. Many NN activation functions satisfy one or more of these assumptions (e.g. tanh, sigmoid), so many approaches to Lurie system analysis can be adopted without modification. An approach for handling activation functions which invalidate the slope-restricted assumption was detailed in [22]. Alternatively, it is possible to focus on certain activation functions and, by deriving additional quadratic constraints (QCs), obtain stability criteria dedicated to the particular activation function. This was the approach taken in [15], [16] where strengthened Circle and Popov Criteria were derived for Lurie systems in which the nonlinearity was of the ReLU type. This line of research was continued in [17], [18] where a more detailed analysis of the QCs was given. In [15], [16], which treated continuous and discrete-time systems respectively, it was observed that a dramatic decrease in conservatism could be achieved without a large increase in computational complexity which is often associated with other more advanced criteria [4], [6], [23], [24].

In this paper, the normal assumptions above are replaced with the assumption that the nonlinearity is the repeated (applied element-wise) magnitude function. This makes the results relevant to systems where the magnitude naturally arises such as electronic circuits containing full-bridge rectifiers, and a class of systems able to produce chaotic behaviour [25]. Furthermore, it is noted that some activation functions, such as the ReLU and leaky ReLU, can also be represented using the magnitude nonlinearity. Therefore, the results introduced here can also be used to analyse the stability of systems containing such activation functions (e.g., neural networks) and can be considered a more general alternative to the results of [15].

Contribution: Novel loop transformations are presented in Section II which enable systems with (leaky) ReLU nonlinearities to be represented as systems with magnitude nonlinearities and vice versa. New QCs for the repeated magnitude, which generalise existing QCs, are presented in Section III. These are leveraged to establish new stability criteria with potentially lower conservatism (Section IV). Similar, but different, relaxations to those used in [15] are presented in Section V to convexify one of the main results. Section VI presents

This work was supported and funded by DSTL and the UKRI centre for doctoral training in Machine Intelligence for Nano-electronic Devices and Systems [EP/S024298/1]. C.R. Richardson is with the Department of Engineering Science, University of Oxford, UK (carl.richardson@eng.ox.ac.uk). M.C. Turner and S.R. Gunn are with the School of Electronics and Computer Science, University of Southampton, UK ({m.c.turner@, srg@ecs.}soton.ac.uk).

some simple conditions for when global and local stability are equivalent; the proof is different but analogous to the discrete-time result presented in [16, Theorem 19]. Finally, Section VII compares numerically the conservatism and complexity of the new criteria against existing methods, including on NN examples which are of practical significance.

II. Preliminaries

A. Notation

The field of real numbers, the space of m dimensional real vectors and the space of real $m \times n$ matrices are indicated by $\mathbb{R}$, $\mathbb{R}^m$ and $\mathbb{R}^{m\times n}$ respectively. The sets of real numbers, m-dimensional vectors and $m \times n$ matrices each with non-negative entries are respectively denoted by $\mathbb{R}_{\geq 0}$, $\mathbb{R}^m_{\geq 0}$ and $\mathbb{R}^{m\times n}_{\geq 0}$. The set of $m \times m$ real symmetric matrices with non-negative elements and positive definite symmetric matrices are respectively denoted by $\mathbb{S}^m_{\geq 0}$ and $\mathbb{S}^m_+$. The diagonal subset of $\mathbb{S}^m_+$ is represented by $\mathbb{D}^m_+$. The set of square matrices which have non-negative off diagonal elements is denoted $\mathbb{M}^m$ (also called *Metzler* matrices). The notation $M \prec\ (\succ)\ 0$ indicates the matrix M is negative (positive) definite and the sum of a matrix M with its transpose is defined as $\mathrm{He}(M) := M + M'$. The space of real rational transfer function matrices, analytic in the right-half complex plane is represented by $\mathcal{RH}_\infty$. Finally, the repeated magnitude function is expressed as the component-wise application of the magnitude function, see definition below for $u \in \mathbb{R}^m$.

$$|u| := \begin{bmatrix} |u_1| & \dots & |u_m| \end{bmatrix}' \qquad |u_i| := \begin{cases} +u_i & u_i \geq 0 \\ -u_i & u_i < 0 \end{cases} \tag{1}$$

B. Problem formulation

The Lurie system in Fig. 1, with $P(s) \in \mathcal{RH}_\infty$, has the following state-space realisation

$$\begin{aligned} \dot{x} &= Ax + B\Phi(y) \\ y &= Cx + D\Phi(y) \end{aligned} \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, $D \in \mathbb{R}^{m\times m}$ and $\Phi(\cdot) : \mathbb{R}^m \mapsto \mathbb{R}^m$. In this paper, a special case of Lurie system is studied, where $\Phi(y) \equiv |y|$. That is

$$\begin{aligned} \dot{x} &= Ax + B|y| \\ y &= Cx + D|y| \end{aligned} \tag{3}$$

It is assumed throughout that (A, B) is stabilisable and (C, A) is detectable. The following assumption is also made.

Assumption 1 (Well-posedness): A unique solution $x(t)$ exists to (3) for all $x(0) \in \mathbb{R}^n$ and all $t \geq 0$.

As the magnitude function is globally Lipschitz, Assumption 1 will hold unconditionally for $D = 0$; when $D \neq 0$, sufficient conditions are discussed in Section II-D. The main problem addressed in the remainder of the paper is stated next.

Problem 1: Find tractable Lyapunov conditions which ensure the origin of (3) is globally asymptotically stable (GAS).

C. Loop transformations between magnitude and ReLU

Prior work, such as [15], has shown that many NN stability problems can be expressed as a Lurie system (2), where $\Phi(\cdot)$ contains the NN activation functions. When the popular leaky ReLU is chosen as the NN activation function, such problems can be framed in the form (3), as shown below. Consider the repeated *leaky ReLU* function, $\Phi_{\mathrm{LR}}(\cdot)$, defined by

$$\Phi_{\mathrm{LR}}(u) := \begin{bmatrix} \phi_{\mathrm{LR}}(u_1) \\ \vdots \\ \phi_{\mathrm{LR}}(u_m) \end{bmatrix} \qquad \phi_{\mathrm{LR}}(u_i) := \begin{cases} u_i & u_i \geq 0 \\ \epsilon u_i & u_i < 0 \end{cases} \tag{4}$$

where $0 \leq \epsilon \ll 1$ is typically assumed. For the special case $\epsilon = 0$, Eq. (4) defines the repeated ReLU function, $\Phi_{\mathrm{ReLU}}(\cdot)$. Since all functions can be written in terms of their odd and even components, it is easy to see that

$$\Phi_{\mathrm{LR}}(u) = \frac{1+\epsilon}{2}u + \frac{1-\epsilon}{2}|u| \tag{5}$$

where the leaky ReLU is a sum of a linear term and magnitude term. Hence, it is clear that NN stability problems involving the repeated leaky ReLU can be written as problems involving the repeated magnitude by leveraging (5). To simplify the exposition, ϵ is set to $\epsilon = 0$ and transformations between (3) and (6) are derived; this could easily be extended to the leaky ReLU where $\epsilon \neq 0$.

$$\begin{aligned} \dot{x} &= \tilde{A}x + \tilde{B}\Phi_{ReLU}(y) \\ y &= \tilde{C}x + \tilde{D}\Phi_{ReLU}(y) \end{aligned} \tag{6}$$

The following two results capture the relationships between the systems (3) and (6).

Proposition 1: Assume $\tilde{E} = I - \frac{1}{2}\tilde{D} \in \mathbb{R}^{m\times m}$ is nonsingular. Then the feedback system (6) can be expressed in the form (3) via the following relationship

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := \left[\begin{array}{c|c} \tilde{A} + \frac{1}{2}\tilde{B}\tilde{E}^{-1}\tilde{C} & \frac{1}{2}\tilde{B}(I + \frac{1}{2}\tilde{E}^{-1}\tilde{D}) \\ \hline \tilde{E}^{-1}\tilde{C} & \frac{1}{2}\tilde{E}^{-1}\tilde{D} \end{array}\right] \tag{7}$$

Proof: Using (5), with $\epsilon = 0$, (6) can be rearranged as

$$(I - \frac{1}{2}\tilde{D})y = \tilde{C}x + \frac{1}{2}\tilde{D}|y|$$

From the definition of $\tilde{E}$ and its assumed nonsingularity

$$y = \tilde{E}^{-1}\tilde{C}x + \frac{1}{2}\tilde{E}^{-1}\tilde{D}|y| \tag{8}$$

The state equation of (6) can be re-written using (5) and (8)

$$\dot{x} = (\tilde{A} + \frac{1}{2}\tilde{B}\tilde{E}^{-1}\tilde{C})x + \frac{1}{2}\tilde{B}(I + \frac{1}{2}\tilde{E}^{-1}\tilde{D})|y| \tag{9}$$

Defining (7) based on (8) and (9) leads to the equivalent representation of (6) in the form (3). □

Conversely, magnitude feedback systems (3) can be expressed in terms of ReLU feedback systems (6).

Proposition 2: Assume $E = I + D \in \mathbb{R}^{m\times m}$ is nonsingular. Then the feedback system (3) can be expressed in the form (6) via the following relationship

$$\left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right] := \left[\begin{array}{c|c} A - BE^{-1}C & 2B(I - E^{-1}D) \\ \hline E^{-1}C & 2E^{-1}D \end{array}\right] \tag{10}$$

Proof: Reversing (5), the output equation of (3) becomes

$$(I + D)y = Cx + 2D\Phi_{\mathrm{ReLU}}(y)$$

Noting the definition and nonsingularity of E then gives

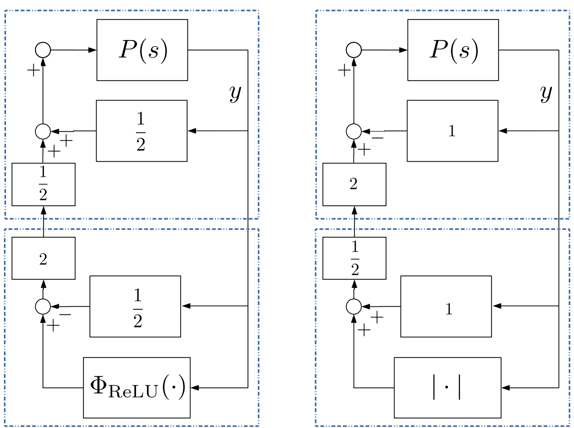
$$y = E^{-1}Cx + 2E^{-1}D\Phi_{\mathrm{ReLU}}(y) \tag{11}$$

The state equation of (3) can also be re-written using the reversal of (5) and (11)

$$\dot{x} = (A - BE^{-1}C)x + 2B(I - E^{-1}D)\Phi_{\mathrm{ReLU}}(y) \tag{12}$$

Thus, defining (10) based on (11) and (12) leads to the equivalent representation of (3) in the form (6). □

Propositions 1 and 2 show that systems with ReLU nonlinearities (6) can be analysed with the results derived here; and systems with magnitude nonlinearities (3) can be analysed with the results of [15]. Both Propositions assume the existence of an inverse matrix. For an L-layer feed-forward NN, the $\tilde{D}$-matrix in (6) has a sparse strictly triangular structure [15, Eq. 3], in which case $\tilde{E}$ exists. There are many cases when E exists, for example, if D in (3) is positive definite or strictly triangular. The loop transformations used in Propositions 1 and 2 are illustrated by Fig. 2.



(a) (b)

Fig. 2. Loop transformations from (a) ReLU to magnitude (Proposition 1) and (b) magnitude to ReLU (Proposition 2).

Remark 1: The loop transformations in Fig. 2 resemble those in [26, Fig. 5] for illustrating the equivalence between the Small Gain Theorem and Passivity Theorem. The key difference is that each block in Fig. 2 only requires one feedback or feed-forward connection; in [26], each block requires both. Loop transformations are used extensively in absolute stability problems and have been used in [11] and [27] to, respectively, modify the representation of a NN controller and compute the upper bound of a NN Lipschitz constant. □

D. Well-posedness

Traditionally, absolute stability problems assume $D = 0$. This is true for recurrent neural networks (RNNs) e.g., [28]; however, systems involving feed-forward NNs can sometimes be expressed as (2) where the NN weights form a sparse, non-zero D matrix [15]. To handle such systems, well-posedness of (2) must be addressed for $D \neq 0$. When $\Phi(\cdot)$ is an incrementally sector bounded nonlinearity in the Sector$[0, I]$, it has been shown in [29] that well-posedness is guaranteed if there exists a matrix $\mathbf{Y} \in \mathbb{D}_+^m$ satisfying the following linear matrix inequality (LMI)

$$2\mathbf{Y} - \mathbf{Y}\tilde{D} - \tilde{D}'\mathbf{Y} \succ 0 \tag{13}$$

Since $\Phi_{\mathrm{LR}}(\cdot)$ is incrementally in the Sector$[0, I]$, it follows that satisfaction of (13) guarantees (6) is well-posed. Thus, using Proposition 2, it is clear that (3) will be well-posed if (13) is satisfied with $\tilde{D} = 2E^{-1}D = 2(I + D)^{-1}D$. That is

$$\mathbf{Y} - \mathbf{Y}E^{-1}D - (E^{-1}D)'\mathbf{Y} \succ 0 \tag{14}$$

Applying the identity $E^{-1}D = I - E^{-1}$, followed by a congruence transform with transformation matrix E, it can be seen that this inequality is equivalent to

$$\mathbf{Y} - D'\mathbf{Y}D \succ 0 \tag{15}$$

Fact 1: Let $E = I + D \in \mathbb{R}^{m\times m}$ be nonsingular. If there exists a $\mathbf{Y} \in \mathcal{D}_+^m$ satisfying (15), then Assumption 1 holds.

It was noted in [15] that many feedback loops involving NNs naturally ensure that (13) is satisfied and thus are well-posed.

III. Quadratic constraints

Quadratic constraints (QCs), which characterise the nonlinearity of the Lurie system, are the key ingredient in the derivation of many absolute stability results. They are adjoined to the derivative of the Lyapunov candidate to obtain matrix inequalities. This is exploited in the standard Circle/Popov Criteria, but also other QCs have been established for different activation functions in [12], and specific QCs for the ReLU function were derived in [15]. In this section, new QCs are derived for the magnitude nonlinearity, which generalise existing QCs such as the "small gain" bound.

There are three obvious properties the scalar magnitude function satisfies, for all $u_i \in \mathbb{R}$

$$|u_i| \geq 0 \tag{16}$$

$$|u_i| + u_i \geq 0 \tag{17}$$

$$|u_i| - u_i \geq 0 \tag{18}$$

Since these hold component-wise, it is easy to assemble these as the following QC, characterising the repeated magnitude.

Fact 2 (Positivity QC): If $\mathbf{Q} \in \mathbb{R}_{\geq 0}^{3m\times 3m}$ then the following inequality holds

$$\begin{bmatrix} |u| \\ |u| + u \\ |u| - u \end{bmatrix}' \underbrace{\begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \mathbf{Q}_{13} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \mathbf{Q}_{23} \\ \mathbf{Q}_{31} & \mathbf{Q}_{32} & \mathbf{Q}_{33} \end{bmatrix}}_{\mathbf{Q}} \begin{bmatrix} |u| \\ |u| + u \\ |u| - u \end{bmatrix} \geq 0 \; \forall u \in \mathbb{R}^m \tag{19}$$

For any $\alpha \in \mathbb{R}$ a further property satisfied by the scalar magnitude function is

$$\alpha(|u_i| + u_i)(|u_i| - u_i) = \alpha(|u_i|^2 - u_i^2) = 0 \quad \forall u_i \in \mathbb{R} \tag{20}$$

This, combined with (17) and (18) yields the following fact.

Fact 3 (Metzler QC): If $\mathbf{V} \in \mathbb{M}^m$, then the following inequality holds

$$(|u| + u)'\mathbf{V}(|u| - u) \geq 0 \quad \forall u \in \mathbb{R}^m \tag{21}$$

Note that Fact 3 affords substantially more freedom in the choice of $\mathbf{V}$ than a "small gain" bound on $|u|$. In the small gain approach, see for instance [30, Section 4.2.4], one observes that $|u_i|^2 \leq u_i^2$ and then obtains a QC of the form (21) with $\mathbf{V}$ *diagonal*. In Fact 3, $\mathbf{V}$ needs only to be *Metzler*, providing more freedom when solving the resulting matrix inequalities.

$$F_Q(\cdot) = \begin{bmatrix} \mathrm{He}\left(\mathbf{P}A - C'(\mathbf{V}-\mathbf{Q}_{22})C\right) & \mathbf{P}B - C'\mathbf{V}'(I+D) + C'\mathbf{V}(I-D) + C'\mathbf{Q}_{22}D + C'(2\mathbf{Q}_{22}-\mathbf{Q}_{13}) \\ \star & \mathrm{He}\left((I+D)'\mathbf{V}(I-D) + \mathbf{Q}_{22} + \mathbf{Q}_{13} + D'\mathbf{Q}_{22}D + D'(2\mathbf{Q}_{22}-\mathbf{Q}_{13})\right) \end{bmatrix} \quad (26)$$

Remark 2: Other, seemingly less useful, QCs also exist for the repeated magnitude; for example if $\mathbf{W} \in \mathbb{D}_+^m$, it is straightforward to observe that

$$u'\mathbf{W}(u+|u|) \geq 0 \quad \forall u \in \mathbb{R}^m$$
$$u'\mathbf{W}(u-|u|) \geq 0 \quad \forall u \in \mathbb{R}^m$$

These inequalities do not contain negative definite terms in $|\cdot|$ which hampers the derivation of useful LMIs. □

A. Simplifying the Positivity QC

Fact 2 provides nine different QCs and Fact 3 presents an additional one. While it is possible to incorporate all ten constraints into the stability criteria, each constraint is associated with an extra m^2 decision variables, so it is desirable to consolidate these constraints.

Lemma 1 (Simplified Positivity QC): If $\mathbf{Q}$ satisfies Fact 2 and $\mathbf{Q}_{22} \in \mathbb{S}_{\geq 0}^m$, then the following inequality holds

$$\begin{bmatrix} u \\ |u| \end{bmatrix}' \begin{bmatrix} 2\mathbf{Q}_{22} & 2\mathbf{Q}_{22}-\mathbf{Q}_{13} \\ \star & \mathrm{He}(\mathbf{Q}_{13}+\mathbf{Q}_{22}) \end{bmatrix} \begin{bmatrix} u \\ |u| \end{bmatrix} \geq 0 \quad \forall u \in \mathbb{R}^m \quad (22)$$

Proof: The first step is to note the following observations

- The QCs involving $\mathbf{Q}_{21}$, $\mathbf{Q}_{31}$ and $\mathbf{Q}_{32}$ can be omitted since they duplicate those involving $\mathbf{Q}_{12}$, $\mathbf{Q}_{13}$ and $\mathbf{Q}_{23}$.
- Secondly, the QC involving $\mathbf{Q}_{23}$

  $$(|u|+u)'\mathbf{Q}_{23}(|u|-u) \geq 0$$

  can also be removed since it is a more restrictive version of the QC presented in Fact 3. That is, $\mathbf{Q}_{23} \in \mathbb{R}_{\geq 0}^{m\times m}$ whereas $\mathbf{V} \in \mathbb{M}^m$.
- Note that the matrices $\mathbf{Q}_{11}$, $\mathbf{Q}_{22}$ and $\mathbf{Q}_{33}$ can, without loss of generality, be taken as symmetric since

  $$w_i'\mathbf{Q}_{ii}w_i = \frac{1}{2}w_i'(\mathbf{Q}_{ii}+\mathbf{Q}_{ii}')w_i$$

  where $i \in \{1,2,3\}$ and w_i represents the associated term $|u|$ or $|u| \pm u$.
- Finally, it is also possible to remove the QC involving $\mathbf{Q}_{11}$ ($|u|'\mathbf{Q}_{11}|u| \geq 0$). When adjoined to the derivative of the Lyapunov candidate, this QC does nothing to assist in making the derivative negative definite.

Applying each of these observations, condenses (19) to

$$\begin{bmatrix} u \\ |u| \end{bmatrix}' \begin{bmatrix} \mathbf{Q}_{22}+\mathbf{Q}_{33} & \frac{1}{2}(\mathbf{Q}_{12}-\mathbf{Q}_{13})+\mathbf{Q}_{22}-\mathbf{Q}_{33} \\ \star & \frac{1}{2}\mathrm{He}(\mathbf{Q}_{12}+\mathbf{Q}_{13})+\mathbf{Q}_{22}+\mathbf{Q}_{33} \end{bmatrix} \begin{bmatrix} u \\ |u| \end{bmatrix} \geq 0 \quad \forall u \in \mathbb{R}^m$$

This can be further simplified, with the loss of some generality, by noting that $\mathbf{Q}_{33}$ plays the same role as $\mathbf{Q}_{22}$ in the (1,1) and (2,2) entries. In the off-diagonal blocks, $\mathbf{Q}_{33}$ plays the same role as $\mathbf{Q}_{13}$ and $\mathbf{Q}_{12}$ plays the same role as $\mathbf{Q}_{22}$. Finally, $\mathbf{Q}_{12}$ plays the same role as $\mathbf{Q}_{13}$ in the (2,2) entry. Therefore, the above inequality can be replaced by (22). □□

B. Sector-like Lemma for magnitude nonlinearities

The following Lemma will be a useful addition to the previous QCs.

Lemma 2: Suppose that $\mathbf{H} \in \mathbb{R}_{\geq 0}^{m\times m}$ is invertible. Define $\Psi(\mathbf{H},u) := \mathbf{H}^{-1}|\mathbf{H}u|$, then the following inequality holds

$$\Psi(\mathbf{H},u)'\mathbf{W}(|u|-\Psi(\mathbf{H},u)) \geq 0 \quad \forall u \in \mathbb{R}^m \quad (23)$$

for all $\mathbf{W} = \mathbf{H}'\mathbf{M}\mathbf{H}$ where $\mathbf{M} \in \mathbb{R}_{\geq 0}^{m\times m}$.

Proof: Denoting each row of $\mathbf{H} \in \mathbb{R}_{\geq 0}^{m\times m}$ by $\mathbf{H}_i'$

$$|\mathbf{H}_i'u| = |\sum_{j=1}^m \mathbf{H}_{ij}u_j| \leq \sum_{j=1}^m |\mathbf{H}_{ij}u_j| = \sum_{j=1}^m \mathbf{H}_{ij}|u_j| = \mathbf{H}_i'|u|$$

As this applies to each row of $\mathbf{H}$, one can conclude

$$|\mathbf{H}u| \leq \mathbf{H}|u| \quad (24)$$

Next we introduce $\mathbf{H}$ to define a generalised form of the QC involving $\mathbf{Q}_{13}$, from Fact 2, in terms of Ψ.

$$\begin{aligned} S(\mathbf{W},\mathbf{H},u) &:= \Psi(\mathbf{H},u)'\mathbf{W}(|u|-\Psi(\mathbf{H},u)) \\ &= (\mathbf{H}^{-1}|\mathbf{H}u|)'\mathbf{W}(|u|-\mathbf{H}^{-1}|\mathbf{H}u|) \end{aligned}$$

Expanding the terms leads to

$$\begin{aligned} S(\mathbf{W},\mathbf{H},u) = &|\mathbf{H}u|'(\mathbf{H}')^{-1}\mathbf{W}\mathbf{H}^{-1}\mathbf{H}|u| \\ &- |\mathbf{H}u|'(\mathbf{H}')^{-1}\mathbf{W}\mathbf{H}^{-1}|\mathbf{H}u| \end{aligned}$$

Substituting for $\mathbf{M}$ and leveraging (24)

$$\begin{aligned} S(\mathbf{W},\mathbf{H},u) &= |\mathbf{H}u|'\mathbf{M}\mathbf{H}|u| - |\mathbf{H}u|'\mathbf{M}|\mathbf{H}u| \\ &\geq |\mathbf{H}u|'\mathbf{M}|\mathbf{H}u| - |\mathbf{H}u|'\mathbf{M}|\mathbf{H}u| = 0 \end{aligned}$$

It should be noted that this QC has the same form as the sector QC, so it is referred to as the "sector-like" inequality. □□

IV. Global Stability Criteria

This section reports global stability conditions based on quadratic and Lurie-type Lyapunov candidates.

A. Quadratic Lyapunov function

Theorem 1 (Quadratic Criterion): Consider the feedback system (3) and let Assumption 1 be satisfied. If there exists $\mathbf{P} \in \mathbb{S}_+^n$, $\mathbf{V} \in \mathbb{M}^m$, $\mathbf{Q}_{13} \in \mathbb{R}_{\geq 0}^{m\times m}$ and $\mathbf{Q}_{22} \in \mathbb{S}_{\geq 0}^m$ such that

$$F_Q(\mathbf{P},\mathbf{V},\mathbf{Q}_{13},\mathbf{Q}_{22}) \prec 0 \quad (25)$$

where $F_Q(\cdot)$ is defined in (26), then the origin of (3) is GAS.

Proof: Consider the quadratic Lyapunov function $V_Q(x) = x'\mathbf{P}x$ and consider its derivative along the trajectories of (3). Appending the QCs from Fact 3, denoted by $M(y,\mathbf{V})$, and Lemma 1, denoted by $SP(y,\mathbf{Q}_{13},\mathbf{Q}_{22})$, is equivalent to the final quadratic inequality below.

$$\begin{aligned} \dot{V}_Q(x) &= 2x'\mathbf{P}(Ax+B|y|) \\ &\leq 2x'\mathbf{P}(Ax+B|y|) + 2M(y,\mathbf{V}) + SP(y,\mathbf{Q}_{13},\mathbf{Q}_{22}) \\ &= \begin{bmatrix} x \\ |y| \end{bmatrix}' F_Q(\mathbf{P},\mathbf{V},\mathbf{Q}_{13},\mathbf{Q}_{22}) \begin{bmatrix} x \\ |y| \end{bmatrix} \end{aligned}$$

□□

$$F_L(\cdot)=\begin{bmatrix} \mathrm{He}\Big(\mathbf{P}A+C'\mathbf{H}'(\mathbf{\Lambda_1}+\mathbf{\Lambda_2})\mathbf{H}CA-C'(\mathbf{V}-\mathbf{Q}_{22})C\Big) & \star & \star \\ B'\mathbf{P}+B'C'\mathbf{H}'(\mathbf{\Lambda}_1+\mathbf{\Lambda}_2)\mathbf{H}C-\mathbf{V}C+\mathbf{V}'C+(2\mathbf{Q}_{22}-\mathbf{Q}_{13})'C & \mathrm{He}(\mathbf{V}+\mathbf{Q}_{22}+\mathbf{Q}_{13}) & \star \\ \mathbf{H}'(\mathbf{\Lambda}_1-\mathbf{\Lambda}_2)\mathbf{H}CA & \mathbf{H}'(\mathbf{\Lambda}_1-\mathbf{\Lambda}_2)\mathbf{H}CB+\mathbf{H}'\mathbf{M}\mathbf{H} & -\mathrm{He}(\mathbf{H}'\mathbf{M}\mathbf{H}) \end{bmatrix} \quad (31)$$

Remark 3: Theorem 1 is in the form of an LMI and bears some resemblance to the LMI associated with the *Small Gain Theorem* (SGT) [1]. In the small gain approach, which for the magnitude nonlinearity is equivalent to considering diagonal norm-bounded constraints as described in [30, Section 5.1], one has the matrix $\mathbf{V}$ simply being diagonal and constraints associated with the matrices $\mathbf{Q}_{ij}$ do not exist; that is $\mathbf{Q}_{ij}\equiv 0$. In this case, Theorem 1 simply reduces to

$$\begin{bmatrix} \mathbf{P}A+A'\mathbf{P}-C'\mathbf{V}C & \mathbf{P}B-C'\mathbf{V}'D \\ \star & \mathbf{V}-D'\mathbf{V}D \end{bmatrix}\prec 0 \quad (27)$$

This is in the same form as the matrix inequality in [30, Eq. 5.15] and is clearly a special case of (25). Therefore, Theorem 1 is less conservative than the Small Gain Theorem for systems of the form (3). Satisfying (27) also automatically implies well-posedness of (3) since the well-posedness condition (15) naturally arises in the $(2,2)$ block. □

B. Lurie-type Lyapunov function

The Popov Criterion (see, for instance [3], [31]) uses the following Lurie-type Lyapunov function to derive a stability criterion

$$V_a(x)=x'\mathbf{P}x+2\int_0^y \mathbf{\Lambda}\Phi(\sigma)\cdot d\sigma$$

with $\mathbf{\Lambda}\in\mathbb{D}_+^m$. However, when $\Phi(\sigma)\equiv|\sigma|$, it is clear that such a choice is not appropriate. As shown by the scalar case, the integral term

$$2\int_0^y|\sigma|d\sigma=|y|\ y$$

is not guaranteed to be non-negative. Instead, the following alternative choice of Lyapunov function could be made

$$V_b(x)=x'\mathbf{P}x+2\int_0^{\mathbf{H}y}\mathbf{\Lambda}(|\sigma|+\sigma)\cdot d\sigma \quad (28)$$

again with $\mathbf{\Lambda}\in\mathbb{D}_+^m$. The second term can be directly integrated to get

$$\begin{aligned} V_b(x)&=x'\mathbf{P}x+2\sum_{i=1}^m\int_0^{\mathbf{H}_i'y}\mathbf{\Lambda}_{ii}(|\sigma_i|+\sigma_i)d\sigma_i \\ &=x'\mathbf{P}x+\sum_{i=1}^m\mathbf{\Lambda}_{ii}\Big[|\sigma_i|(\sigma_i+|\sigma_i|)\Big]_0^{\mathbf{H}_i'y} \end{aligned}$$

where each row of $\mathbf{H}$ is denoted by $\mathbf{H}_i'$. The final term is positive by virtue of inequality (17). In fact, positivity of the Lyapunov function (28) is not surprising as (5) implies

$$V_b(x)=x'\mathbf{P}x+4\int_0^{\mathbf{H}y}\mathbf{\Lambda}\ \Phi_{\mathrm{ReLU}}(\sigma)\cdot d\sigma$$

This gives $V_b(x)$ the same form as the Lurie-type Lyapunov function used in [15].

Unfortunately, the Lyapunov function (28) still requires the matrix $\mathbf{\Lambda}$ to be positive definite, despite the extra freedom afforded by the matrix $\mathbf{H}$. For this reason, the following Lyapunov function is proposed

$$\begin{aligned} V_L(x)=x'\mathbf{P}x&+2\int_0^{\mathbf{H}_1y}\mathbf{\Lambda}_1(|\sigma|+\sigma)\cdot d\sigma \\ &+2\int_0^{\mathbf{H}_2y}\mathbf{\Lambda}_2(\sigma-|\sigma|)\cdot d\sigma \quad (29) \end{aligned}$$

with $\mathbf{\Lambda}_1,\mathbf{\Lambda}_2\in\mathbb{D}_+^m$. These two integral terms are useful because while $\mathbf{\Lambda}_1,\mathbf{\Lambda}_2$ are both positive definite they lead to terms $\mathbf{\Lambda}_1-\mathbf{\Lambda}_2$ in the resulting matrix inequality, potentially inducing a reduction in conservatism. The same approach was used in [3], [31] to, effectively, enable $\mathbf{\Lambda}$ to be indefinite. The integrals in $V_L(x)$ can be evaluated to obtain

$$\begin{aligned} V_L(x)=x'\mathbf{P}x&+\sum_{i=1}^m\mathbf{\Lambda}_{1,ii}\Big[|\sigma_i|(|\sigma_i|+\sigma_i)\Big]_0^{\mathbf{H}_{1,i}'y} \\ &+\sum_{i=1}^m\mathbf{\Lambda}_{2,ii}\Big[|\sigma_i|(|\sigma_i|-\sigma_i)\Big]_0^{\mathbf{H}_{2,i}'y} \end{aligned}$$

The final term is positive by virtue of inequality (18). Furthermore, this final expression verifies that the integrals in (29) are path independent [32]. Using the Lyapunov function $V_L(x)$, the following result can be derived.

Theorem 2 (Lurie-based Criterion): Consider the feedback system (3) and let $D=0$. If there exists $\mathbf{P}\in\mathbb{S}_+^n$; $\mathbf{\Lambda}_1,\mathbf{\Lambda}_2\in\mathbb{D}_+^m$; $\mathbf{V}\in\mathbb{M}^m$; $\mathbf{Q}_{22}\in\mathbb{S}_{\geq 0}^m$; $\mathbf{Q}_{13},\mathbf{M},\mathbf{H}\in\mathbb{R}_{\geq 0}^{m\times m}$ with $\mathbf{H}$ nonsingular such that

$$F_L(\mathbf{P},\mathbf{H},\mathbf{\Lambda}_1,\mathbf{\Lambda}_2,\mathbf{M},\mathbf{V},\mathbf{Q}_{22},\mathbf{Q}_{13})\prec 0 \quad (30)$$

where $F_L(\cdot)$ is given in (31), then the origin of (3) is GAS.

Proof: At the expense of some conservatism, it is assumed that $\mathbf{H}:=\mathbf{H}_1=\mathbf{H}_2$, leading to the following derivative of the Lyapunov function (29)

$$\dot V_L(x)=2x'\mathbf{P}\dot x+2\Big(y'\mathbf{H}'(\mathbf{\Lambda}_1+\mathbf{\Lambda}_2)+|\mathbf{H}y|'(\mathbf{\Lambda}_1-\mathbf{\Lambda}_2)\Big)\mathbf{H}C\dot x$$

Appending the QCs from Fact 3, Lemma 1 and Lemma 2, respectively denoted by $M(y,\mathbf{V})$, $SP(y,\mathbf{Q}_{13},\mathbf{Q}_{22})$ and $S(y,\mathbf{W},\mathbf{H})$ results in

$$\begin{aligned} \dot V_L(x)\leq \dot V_L(x)+2M(y,\mathbf{V})&+SP(y,\mathbf{Q}_{13},\mathbf{Q}_{22}) \\ &+2S(y,\mathbf{W},\mathbf{H}) \end{aligned}$$

After some algebra, the above inequality can be expressed in the quadratic form

$$\dot V_L(x)\leq\begin{bmatrix}x\\|y|\\\Psi\end{bmatrix}'F_L(\cdot)\begin{bmatrix}x\\|y|\\\Psi\end{bmatrix} \quad (32)$$

where $\Psi=\Psi(\mathbf{H},y)$ and $F_L(\cdot)$ is defined in (31). □□

$$G(\cdot)=\begin{bmatrix}\mathrm{He}\Big(\mathbf{P}A+2C'\mathbf{X}CA-C'(\mathbf{V}-\mathbf{Q}_{22})C\Big) & \star & \star\\ B'\mathbf{P}+2B'C'\mathbf{X}C-\mathbf{V}C+\mathbf{V}'C+(2\mathbf{Q}_{22}-\mathbf{Q}_{13})'C & \mathrm{He}\Big(\mathbf{V}+\mathbf{Q}_{22}+\mathbf{Q}_{13}\Big) & \star\\ 0 & \mu\mathbf{X} & -\mathrm{He}(\mu\mathbf{X})\end{bmatrix} \tag{39}$$

$$J(\cdot)=\begin{bmatrix}\mathrm{He}\Big(\mathbf{P}A+C'(\mathbf{\Lambda_1}+\mathbf{\Lambda}_2)CA-C'(\mathbf{V}-\mathbf{Q}_{22})C\Big) & \star\\ B'\mathbf{P}+B'C'(\mathbf{\Lambda}_1+\mathbf{\Lambda}_2)C-\mathbf{V}C+\mathbf{V}'C+(2\mathbf{Q}_{22}-\mathbf{Q}_{13})'C+(\mathbf{\Lambda}_1-\mathbf{\Lambda}_2)CA & \mathrm{He}\Big(\mathbf{V}+\mathbf{Q}_{22}+\mathbf{Q}_{13}+(\mathbf{\Lambda}_1-\mathbf{\Lambda}_2)CB\Big)\end{bmatrix} \tag{41}$$

Remark 4: It is possible to enhance Theorem 2 further by noting that, since $\mathbf{H}$ is assumed positive and invertible, the quadratic inequalities (16)-(18) can be used to obtain

$$\mathbf{H}\Psi(\mathbf{H},u)\geq 0 \tag{33}$$
$$\mathbf{H}\Psi(\mathbf{H},u)+\mathbf{H}u\geq 0 \tag{34}$$
$$\mathbf{H}\Psi(\mathbf{H},u)-\mathbf{H}u\geq 0 \tag{35}$$

where $\Psi(\mathbf{H},u)$ is defined in Lemma 2. By application of Fact 2, this implies that for all $\mathbf{R}\in\mathbb{R}^{3m\times 3m}_{\geq 0}$ and for all $u\in\mathbb{R}^m$

$$\begin{bmatrix}\mathbf{H}\Psi\\ \mathbf{H}\Psi+\mathbf{H}u\\ \mathbf{H}\Psi-\mathbf{H}u\end{bmatrix}'\underbrace{\begin{bmatrix}\mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13}\\ \mathbf{R}_{21} & \mathbf{R}_{22} & \mathbf{R}_{23}\\ \mathbf{R}_{31} & \mathbf{R}_{32} & \mathbf{R}_{33}\end{bmatrix}}_{\mathbf{R}}\begin{bmatrix}\mathbf{H}\Psi\\ \mathbf{H}\Psi+\mathbf{H}u\\ \mathbf{H}\Psi-\mathbf{H}u\end{bmatrix}\geq 0 \tag{36}$$

where the shorthand $\Psi=\Psi(\mathbf{H},u)$ has been used. Simplifying in the same way as Lemma 1, the reduced set of QCs become

$$\begin{bmatrix}u\\ \Psi\end{bmatrix}'\begin{bmatrix}2\mathbf{H}'\mathbf{R}_{22}\mathbf{H} & \mathbf{H}'(2\mathbf{R}_{22}-\mathbf{R}_{13})\mathbf{H}\\ \star & \mathbf{H}'\,\mathrm{He}(\mathbf{R}_{13}+\mathbf{R}_{22})\mathbf{H}\end{bmatrix}\begin{bmatrix}u\\ \Psi\end{bmatrix}\geq 0 \tag{37}$$

Using the S-procedure, this constraint could additionally be adjoined to (32), although at the expense of adding several more decision variables to the arising matrix inequality. □

Remark 5: For $D=0$, Theorem 1 is a special case of Theorem 2 with $\mathbf{\Lambda}_1=\mathbf{\Lambda}_2=\mathbf{W}=0$. Hence, Theorem 2 will always be less conservative under this condition. □

Remark 6: The key advantage of Theorem 1 over Theorem 2 is that it accounts for $D\neq 0$ and can be applied to systems involving feed-forward NNs. When $D=0$, such as RNNs and full-bridge rectifiers, Theorem 2 will be less conservative. □

V. Convex relaxations for Theorem 2

Theorem 2 is potentially much less conservative than Theorem 1, but suffers from a crucial weakness: (30) is a nonlinear matrix inequality since it contains products of two or three matrix variables. This section presents two relaxations which recover LMIs from (30) and make Theorem 2 tractable.

A. Specific choice of $\mathbf{\Lambda}$ and $\mathbf{M}$

Corollary 1: Consider the feedback system (3) and let $D=0$. If there exist $\mathbf{P}\in\mathbb{S}^n_+$; $\mathbf{X},\mathbf{Q}_{22}\in\mathbb{S}^m_{\geq 0}$; $\mathbf{V}\in\mathbb{M}^m$; $\mathbf{Q}_{13}\in\mathbb{R}^{m\times m}_{\geq 0}$ and a scalar $\mu>0$ such that

$$G(\mathbf{P},\mathbf{X},\mathbf{V},\mathbf{Q}_{22},\mathbf{Q}_{13},\mu)\prec 0 \tag{38}$$

where $G(\cdot)$ is defined in (39), then the origin of (3) is GAS.

Proof: The matrix inequality (38) is a special case of (30) where $\mathbf{\Lambda}_1=\mathbf{\Lambda}_2=I$, $\mathbf{M}=\mu I$, and $\mathbf{X}:=\mathbf{H}'\mathbf{H}\in\mathbb{S}^m_{\geq 0}$. □□

With this choice, the only remaining bilinear term is $\mu\mathbf{X}$. Hence, (38) can be solved as an LMI plus a line search.

B. Specific choice of $\mathbf{H}$

Corollary 2: Consider the feedback system (3) and let $D=0$. If there exists $\mathbf{P}\in\mathbb{S}^n_+$; $\mathbf{\Lambda}_1,\mathbf{\Lambda}_2\in\mathbb{D}^m_+$; $\mathbf{V}\in\mathbb{M}^m$; $\mathbf{Q}_{22}\in\mathbb{S}^m_{\geq 0}$ and $\mathbf{Q}_{13}\in\mathbb{R}^{m\times m}_{\geq 0}$ such that

$$J(\mathbf{P},\mathbf{\Lambda}_1,\mathbf{\Lambda}_2,\mathbf{V},\mathbf{Q}_{22},\mathbf{Q}_{13})\prec 0 \tag{40}$$

where $J(\cdot)$ is defined in (41), then the origin of (3) is GAS.

Proof: The LMI (40) is a special case of (30) when $\mathbf{H}=I$. Note that this relaxation implies $\Psi(I,y)=|y|$ which makes Lemma 2 redundant. □□

To motivate this choice, first restrict $\mathbf{H}\in\mathbb{D}^m_+$. This implies that the products $\mathbf{H}'\mathbf{\Lambda}_1\mathbf{H},\mathbf{H}'\mathbf{\Lambda}_2\mathbf{H}\in\mathbb{D}^m_+$ and the product $\mathbf{H}'\mathbf{M}\mathbf{H}\in\mathbb{R}^{m\times m}_{\geq 0}$. Since these matrices do not appear elsewhere in inequality (31), the choice $\mathbf{H}=I$ can be made without loss of generality. The appeal of the matrix inequality (40) is that it is entirely *linear*, making Corollary 2 preferable to Corollary 1.

Remark 7: Theorem 1 is also a special case of Corollary 2, when $D=0$, under the same relaxations as Remark 5. Hence, Corollary 2 will always be less conservative. □

VI. Equivalence of global and local stability

It is obvious that the magnitude nonlinearity is *positive homogenous*, as is the (leaky) ReLU nonlinearity: this implies $\alpha|u|=|\alpha u|$ for all scalars $\alpha>0$. This fact can be harnessed to prove that, under mild conditions, if $x=0$ is a local equilibrium point of the system (3), then it will in fact be a global equilibrium point. This is unusual for typical Lurie systems where, if a global analysis fails, one might still try to establish local stability for a given region of attraction. The first step in establishing that global and local stability are equivalent is the following lemma.

Lemma 3: Assume $(I-DU)$ is invertible and define

$$\mathcal{H}=\big\{A+BU(I-DU)^{-1}C\ :\ U\in\mathcal{U}\big\} \tag{42}$$

and

$$\mathcal{U}=\{\mathrm{diag}(u_1,\ldots,u_m)\mid u_i\in -1,1\text{ and } i\in 1,\ldots,m\} \tag{43}$$

If all matrices in $\mathcal{H}$ are full rank, then $x=0$ is a unique equilibrium point of (3).

Proof: For $U\in\mathcal{U}$ and by definition of the magnitude function, the repeated magnitude may be expressed as

$$|y|=Uy \tag{44}$$

Now, at equilibrium, the state equation of (3) becomes

$$0=Ax+B|y| \tag{45}$$

Assume $x\neq 0$ is an equilibrium point, then (44) is leveraged to equivalently express (45) as

$$0=\big(A+BU(I-DU)^{-1}C\big)x,\quad U\in\mathcal{U} \tag{46}$$

				Maximum series gain (left) and number of decision variables (right)																
Ex	n	m	Source	SGT [1]		Theorem 1 [15]		Corollary 1 [15]		Corollary 2 [15]		Park [5]		Zames-Falb [33]		Theorem 1		Corollary 2		Nyquist Gain
1	9	3	[5] Ex. 3	20.8766	48	39.5246	63	52,430.1224	81	9,990.0000	66	448.7543	87	435.8305	252	40.1000	69	**90,283.3800**	75	100,000+
2	3	3	[29] Ex. 3	**89.9000**	9	**89.9000**	24	**89.9000**	42	**89.9000**	27	**89.9000**	30	**89.9000**	45	**89.9000**	30	**89.9000**	36	89.9000
3	3	4	[7] Ex. 4.9	0.5236	10	**0.6818**	38	**0.6818**	70	**0.6818**	42	0.5236	40	0.5526	48	**0.6818**	48	**0.6818**	56	0.6983
4	8	4	[34] Ex. 22	0.0010	40	0.0012	68	0.0012	100	**0.0015**	72	0.0010	90	0.0010	208	0.0012	78	**0.0015**	86	0.0020
5	6	4	[34] Ex. 17	0.0813	25	0.0814	53	0.0814	85	0.0830	57	**0.0845**	67	**0.0845**	129	0.0814	63	**0.0845**	71	0.0869
6	6	4	[34] Ex. 18	0.1946	25	0.3901	53	0.4158	85	0.5048	57	0.2266	67	0.2785	129	0.4082	63	**0.6129**	71	0.8202
7	8	4	[34] Ex. 22	0.0966	40	0.1232	68	0.1232	100	0.1462	72	0.1035	90	0.1040	208	0.1232	78	**0.1502**	86	0.2002
8	5	5	[29] Ex. 2	**2.0221**	20	**2.0221**	65	**2.0221**	115	**2.0221**	70	**2.0221**	70	**2.0221**	100	**2.0221**	80	**2.0221**	90	2.0221
9	40	40		1.4695	860	**2.0516**	4,020	**2.0516**	7,220	**2.0516**	4,060	1.4695	3,360	1.4695	4,300	**2.0516**	4,840	**2.0516**	4,920	2.0600
10	60	60		1.3820	1,890	2.0224	9,030	2.0343	16,230	**2.2240**	9,090	1.3820	7,440	1.3820	9,450	2.0699	10,860	2.1088	10,980	2.7000
11	80	80		1.4531	3,320	2.1189	16,040	2.0449	28,840	**2.2092**	16,120	1.4605	13,120	1.4605	16,600	2.1189	19,280	2.2067	19,440	2.2200
12	100	100		1.4613	5,150	2.2144	25,050	2.1582	45,050	**2.3380**	25,150	1.4613	20,400	1.4613	25,750	2.2256	30,100	2.3099	30,300	2.7500

TABLE I

COMPARISON OF THE MAXIMUM SERIES GAIN AND THE NUMBER OF DECISION VARIABLES FOR VARIOUS CRITERIA. EACH SYSTEM HAS STATE DIMENSION n AND OUTPUT DIMENSION m.

However, as all matrices in $\mathcal{H}$ are full rank; there is a contradiction and hence $x \neq 0$ cannot be an equilibrium point. Thus, $x = 0$ is a unique equilibrium point of (3). □□

Remark 8: A sufficient condition for $(I - DU)^{-1}$ to exist is for Fact 1 to be satisfied i.e., (3) is well-posed. This can simply be seen by leveraging (44) to show that under Fact 1, a solution to $y = Cx + DUy$ must exist for all $x \in \mathbb{R}^n$ and $U \in \mathcal{U}$; hence, $(I - DU)$ must be nonsingular $\forall\, U \in \mathcal{U}$. □

As an example, Hopfield RNNs [28] can be expressed in the form (3) with $A = -I + \frac{1}{2}\tilde{B}$, $B = \frac{1}{2}\tilde{B}$, $C = I$, $D = 0$, where $\tilde{B} \in \mathbb{R}^{n\times n}$. Therefore, Lemma 3 requires $(\tilde{B}U - I)$ to be full rank for all $U \in \mathcal{U}$; this will typically be the case unless $\tilde{B}$ has a very special structure. Lemma 3 can then be used to establish the following interesting result.

Theorem 3: Consider the Lurie system (3) and let Assumption 1 be satisfied. If $x = 0$ is a unique equilibrium point and it is locally stable, then it is also globally stable.

Proof: Assume $x = 0$ is a locally stable equilibrium point of (3), this implies the existence of a ball

$$\mathcal{B}(x, c) := \{x \in \mathbb{R}^n : \|x\| < c\}$$

such that $\forall\, x(0) \in \mathcal{B}(x, c)$ it follows that $\lim_{t\to\infty} x(t) = 0$; that is, $\mathcal{B}(x, c)$ is a region of attraction of $x = 0$. Assumption 1 implies that a unique solution, y, exists to the implicit equation

$$\theta(y) := y - D|y| = Cx$$

for all x; equivalently, $\theta(\cdot) : \mathbb{R}^m \mapsto \mathbb{R}^m$ is invertible. It is easy to see that $\theta(\cdot)$ is positive homogenous and as $\theta^{-1}(\cdot)$ is assumed to exist, it must also be positive homogenous (see Appendix A). The Lurie system (3) can thus be written as

$$\dot{x} = Ax + B|\theta^{-1}(Cx)| \tag{47}$$

Defining the change of coordinates $z := \alpha x$ for $\alpha > 0$, the dynamics in the scaled coordinates are

$$\dot{z} = \alpha\Big(Ax + B|\theta^{-1}(Cx)|\Big) = Az + B|\theta^{-1}(Cz)| \tag{48}$$

As this is of the same form as (47), it must have an equilibrium point $z = 0$ and associated region of attraction $\mathcal{B}(z, c)$. Scrutinising the region of attraction

$$\begin{aligned}\mathcal{B}(z, c) &= \{z \in \mathbb{R}^n : \|z\| < c\} \\ &= \Big\{x \in \mathbb{R}^n : \|x\| < \frac{c}{\alpha}\Big\} = \mathcal{B}\big(x, \frac{c}{\alpha}\big)\end{aligned} \tag{49}$$

Since α can be any (arbitrary small) positive scalar, it is clear that $\mathcal{B}(x, c/\alpha) \to \mathbb{R}^n$ as $\alpha \to 0$. □□

If it is possible to establish $x = 0$ as a unique locally stable equilibrium point of (3), then this section has shown that it is in fact globally stable. Equivalently, if the conditions of Lemma 3 are satisfied, but it is not possible to prove global stability of (3), then it is futile to attempt a local stability analysis. Similar results were proved for discrete-time Lurie systems containing the ReLU nonlinearity in [16].

VII. Numerical Examples

Twelve example Lurie systems were used to compare Theorem 1 and Corollary 2 against existing LMI based stability criteria, see Table I. Examples 1-8 appeared in [15] and Examples 9-12 represent four GAS Hopfield networks, which were randomly generated according to the SVD Combo parametrisation [35]. Each example was assumed to have a repeated ReLU nonlinearity so, the results in this paper were applied using the loop-transformations of Section II-C. The code[1] for generating the results is found below.

The maximum series gain (MSG) and number of decision variables were used to, respectively, compare the conservatism and complexity of the criteria. The comparison was performed by inserting a series gain (α) into the feedback loop, replacing B with αB and D with αD, then computing the MSG for which stability was guaranteed. This was implemented by a bisection method in conjunction with the YALMIP parser [36] and Mosek LMI solver [37]. Various methods were compared to Theorem 1 and Corollary 2: the classical SGT [30, Section 5.1], the ReLU results of [15], which strengthened the Circle and Popov Criteria, and the more advanced approaches of Park [5] and Zames-Falb, using the multiplier search of [33].

Theorem 1 was the least conservative criteria, on all examples, out of those which can be applied when $D \neq 0$ (SGT, Theorem 1 [15]). In Examples 1-8, it is noteworthy that Corollary 2 has conservatism lower than or equal to all other criteria. However, Corollary 2 [15] marginally outperforms Corollary 2 in the high-dimensional Hopfield network examples. The LMIs in Therorem 1 and Corollary 2 contain more decision variables than their counterparts in Theorem 1 and Corollary 2 of [15], so this is an unwelcome deficiency with the approach.

VIII. Conclusion

This paper analysed the stability of Lurie systems with repeated magnitude nonlinearities. The relationship with Lurie systems having (leaky) ReLU nonlinearities was highlighted through loop transformations between the two systems; as a result, Lurie systems with (leaky) ReLU nonlinearities can be analysed using the results in this paper. By positive homogeneity of the magnitude function, it was shown that if the Lurie system had a unique equilibrium point, then it must

[1] https://github.com/CR-Richardson/Max-Series-Gain-Magnitude

either be globally stable or unstable, rendering local stability analysis futile. Finally, global stability criteria, in the form of linear matrix inequalities, were established based on novel quadratic constraints which characterised the repeated magnitude. The new criteria were compared against other state-of-the-art criteria and Theorem 1 was the least conservative, out of those which can be applied when $D \neq 0$ (e.g., systems with feed-forward NNs). Overall, Corollary 2 was the least conservative on all the low dimensional numerical examples. For high-dimensional systems with $D = 0$, the main advantage of Corollary 2 is that it can be applied to a wider class of activation functions (ReLU and Leaky ReLU), allowing more flexible NN design, whilst only being marginally more conservative on ReLU problems. Future work will consider if sparsity can be leveraged to further improve scalability for systems involving larger neural networks.

References

[1] H. K. Khalil, "Nonlinear systems," *Patience Hall*, vol. 115, 2002.

[2] G. Leonov, D. Ponomarenko, and V. Smirnova, *Frequency-Domain Methods for Nonlinear Analysis*. Singapore: World Scientific, 1996.

[3] W. P. Heath and G. Li, "Lyapunov functions for the multivariable Popov Criterion with indefinite multipliers," *Automatica*, vol. 45, no. 12, pp. 2977–2981, 2009.

[4] G. Zames and P. Falb, "Stability conditions for systems with monotone and slope restricted nonlinearities," *SIAM Journal of Control*, vol. 6, no. 1, pp. 89–108, 1968.

[5] P. Park, "Stability criteria of sector-and slope-restricted Lur'e systems," *IEEE Transactions on Automatic Control*, vol. 47, pp. 308–313, 2002.

[6] J. Carrasco, M. C. Turner, and W. P. Heath, "Zames–Falb multipliers for absolute stability: From O'Shea's contribution to convex searches," *European Journal of Control*, vol. 28, pp. 1–19, 2016.

[7] M. Fetzer and C. Scherer, "Full-block multipliers for repeated, slope restricted scalar nonlinearities," *International Journal of Robust Nonlinear Control*, vol. 27, no. 17, pp. 3376–3411, 2017.

[8] C. R. Richardson, M. C. Turner, and S. R. Gunn, "Lurie networks with robust convergent dynamics," *TMLR*, 2025.

[9] P. Pauli, D. Gramlich, J. Berberich, and F. Allgöwer, "Linear systems with neural network nonlinearities: Improved stability analysis via acausal Zames-Falb multipliers," in *60th IEEE Conference on Decision and Control*. IEEE, 2021, pp. 3611–3618.

[10] H. Yin, P. Seiler, and M. Arcak, "Stability analysis using quadratic constraints for systems with neural network controllers," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1980–1987, 2021.

[11] H. Yin, P. Seiler, M. Jin, and M. Arcak, "Imitation learning with stability and safety guarantees," *IEEE Control Systems Letters*, vol. 6, pp. 409–414, 2021.

[12] M. Fazlyab, M. Morari, and G. J. Pappas, "Safety verification and robustness analysis of neural networks via quadratic constraints and semidefinite programming," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 1–15, 2020.

[13] H. M. Hedesh and M. Siami, "Ensuring both positivity and stability using sector-bounded nonlinearity for systems with neural network controllers," *arXiv preprint arXiv:2406.12744*, 2024.

[14] M. Newton and A. Papachristodoulou, "Sparse polynomial optimisation for neural network verification," *Automatica*, vol. 157, p. 111233, 2023.

[15] C. R. Richardson, M. C. Turner, and S. R. Gunn, "Strengthened Circle and Popov Criteria for the stability analysis of feedback systems with ReLU neural networks," *IEEE Control Systems Letters*, 2023.

[16] C. R. Richardson, M. C. Turner, S. R. Gunn, and R. Drummond, "Strengthened stability analysis of discrete-time Lurie systems involving ReLU neural networks," in *Learning for Decision and Control*, 2024.

[17] S. V. Noori, B. Hu, G. Dullerud, and P. Seiler, "Stability and performance analysis of discrete-time ReLU recurrent neural networks," *arXiv preprint arXiv:2405.05236*, 2024.

[18] ——, "A complete set of quadratic constraints for repeated ReLU," *arXiv preprint arXiv:2407.06888*, 2024.

[19] T. Yuno, K. Fukuchi, and Y. Ebihara, "A Lyapunov-based method of reducing activation functions of recurrent neural networks for stability analysis," *IEEE Control Systems Letters*, 2024.

[20] L. B. Groff, J. M. Gomes da Silva, and G. Valmorbida, "Regional stability of discrete-time linear systems subject to asymmetric input saturation," in *IEEE Conference on Decision and Control*, 2019, pp. 169–174.

[21] L. Cabral, G. Valmorbida, and J. M. Gomes da Silva, "Exponential stability of continuous-time piecewise affine systems," *IEEE Control Systems Letters*, vol. 8, pp. 1649–1654, 2024.

[22] P. Pauli, A. Havens, A. Araujo, S. Garg, F. Khorrami, F. Allgöwer, and B. Hu, "Novel quadratic constraints for extending lipsdp beyond slope-restricted activations," *arXiv preprint arXiv:2401.14033*, 2024.

[23] R. O'Shea, "A combined frequency-time domain stability criterion for autonomous continuous systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 3, pp. 477–484, 1966.

[24] A. L. J. Bertolin, R. C. L. F. Oliveira, G. Valmorbida, and P. L. D. Peres, "Dynamic output-feedback control of continuous-time Lur'e systems using Zames-Falb multipliers by means of an LMI-based algorithm," *IFAC-PapersOnLine*, vol. 55, no. 25, pp. 109–114, 2022.

[25] S. J. Linz and J. Sprott, "Elementary chaotic flow," *Physics Letters A*, vol. 259, no. 3-4, pp. 240–245, 1999.

[26] B. D. Anderson, "The small-gain theorem, the passivity theorem and their equivalence," *Journal of the Franklin Institute*, vol. 293, no. 2, pp. 105–115, 1972.

[27] M. Fazlyab, T. Entesari, A. Roy, and R. Chellappa, "Certified robustness via dynamic margin maximization and improved lipschitz regularization," *Advances in Neural Information Processing Systems*, vol. 36, pp. 34 451–34 464, 2023.

[28] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical journal*, vol. 12, no. 1, pp. 1–24, 1972.

[29] R. Drummond, C. Guiver, and M. C. Turner, "Aizerman conjectures for a class of multivariate positive systems," *IEEE Transactions on Automatic Control*, 2022.

[30] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[31] P. Park, "A revisited Popov criterion for nonlinear Lur'e systems with sector-restrictions," *International Journal of Control*, vol. 68, no. 3, pp. 461–470, 1997.

[32] R. Wrede and M. Spiegel, *Schaum's Outline of Advanced Calculus*. McGraw-Hill Education, 2010.

[33] M. C. Turner and R. Drummond, "Analysis of MIMO Lurie systems with slope restricted nonlinearities using concepts of external positivity," in *IEEE 58th Conference on Decision and Control*, 2019, pp. 163–168.

[34] M. C. Turner, M. L. Kerr, and J. Sofrony, "Tractable stability analysis for systems containing repeated scalar slope-restricted nonlinearities," *International Journal of Robust Nonlinear Control*, vol. 25, no. 7, pp. 971–986, 2015.

[35] L. Kozachkov, M. Ennis, and J.-J. Slotine, "RNNs of RNNs: Recursive construction of stable assemblies of RNNs," *Advances in Neural Information Processing Systems*, vol. 35, pp. 30 512–30 527, 2022.

[36] J. Lofberg, "YALMIP: A toolbox for modelling and optimisation in MATLAB," in *2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508)*. IEEE, 2004, pp. 284–289.

[37] E. D. Andersen and K. D. Andersen, "The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm," *High perfermance optimization*, pp. 197–232, 2000.

Appendix A
Positive homogeneity and inverse functions

Assume that $\theta(\cdot) : \mathbb{R}^m \mapsto \mathbb{R}^m$ is bijective and positive homogenous. Let $y, z \in \mathbb{R}^p$ and $\alpha > 0$ such that

$$\alpha\theta(y) = z \tag{50}$$

By positive homogeneity, it is also clear that

$$\theta(\alpha y) = z \tag{51}$$

Since $\theta(\cdot)$ is bijective, (50) and (51) respectively imply

$$y = \theta^{-1}(\frac{1}{\alpha}z) = \frac{1}{\alpha}\theta^{-1}(z) \tag{52}$$

Defining $\beta := 1/\alpha > 0$, this clearly implies

$$\beta\theta^{-1}(z) = \theta^{-1}(\beta z)$$

That is, $\theta^{-1}(\cdot)$ is positive homogeneous. This fact was also proved, slightly differently, in [16].