New Conditions for Global Stability of Neural Networks with Application to Linear and Quadratic Programming Problems

Mauro Forti and Alberto Tesi, Member, IEEE

Abstract—In this paper, we present new conditions ensuring existence, uniqueness, and Global Asymptotic Stability (GAS) of the equilibrium point for a large class of neural networks. The results are applicable to both symmetric and nonsymmetric interconnection matrices and allow for the consideration of all continuous nondecreasing neuron activation functions. Such functions may be unbounded (but not necessarily surjective), may have infinite intervals with zero slope as in a piece-wise-linear model, or both. The conditions on GAS rely on the concept of Lyapunov Diagonally Stable (or Lyapunov Diagonally Semi-Stable) matrices and are proved by employing a class of Lyapunov functions of the generalized Lur'e-Postnikov type. Several classes of interconnection matrices of applicable interest are shown to satisfy our conditions for GAS. In particular, the results are applied to analyze GAS for the class of neural circuits introduced in [10] for solving linear and quadratic programming problems. In this application, the principal result here obtained is that the networks in [10] are GAS also when the constraint amplifiers are dynamical, as it happens in any practical implementation.

I. INTRODUCTION

ONE OF THE MOST investigated problems in nonlinear circuit theory is that of the existence, uniqueness, and Global Asymptotic Stability (GAS) of the equilibrium point [1]–[7]. In this paper, we address such a problem for nonlinear systems modeling a large class of neural networks of the additive type.

The property of GAS, which means that the domain of attraction of the equilibrium point is the whole space, is of importance from a theoretical as well as an application point of view in several fields [7]–[9]. In particular, in the neural field, GAS networks are known to be well suited for solving some classes of optimization problems in real time [10]–[15], also with connection to adaptive control [16]. In fact, a GAS neural network is guaranteed to compute the global optimal solution independently of the initial condition, which in turn implies that the network is devoid of spurious suboptimal responses. Such GAS neural circuits can also be useful for accomplishing other interesting cognitive or computational tasks [17], [18].

A number of papers in the literature deal with conditions ensuring GAS for neural networks. In [12], [13], GAS is proved for lower triangular interconnection matrices \( T \) in a Hopfield model. In [14], [15], [19], it is shown that for a symmetric \( T \), the negative semidefiniteness of \( T \) is both necessary and sufficient to ensure GAS with respect to the whole class of sigmoidal neuron activation functions. Other papers address conditions for GAS based on row or column dominance conditions for \( T \) [11], [20]. In [17], GAS has been investigated for Generalized Cellular Neural Networks by exploiting classical techniques from control theory such as the circle and small gain criteria. Conditions for GAS in [11], [14], [20] were shown to be special cases of more general theorems on GAS obtained in [21], which are valid for nonsymmetric \( T \) and involve definiteness properties of the symmetric part of \( T \).

The results on GAS previously quoted [11]–[15], [21] concern the case where the neuron activations are assumed bounded and strictly increasing (sigmoidal activations). Unfortunately, these assumptions make the results unapplicable to some important engineering problems. For example, this is the case of neural networks for solving optimization problems in the presence of constraints (linear, quadratic, or more general programming problems [10]), where unbounded activations modeled by diode-like exponential-type functions are needed to impose constraints satisfaction. The extension of the quoted results to the unbounded case is not straightforward. One problem is that, different from the bounded case where the existence of an equilibrium point (a prerequisite for GAS) is always guaranteed [21], [15], for unbounded activations, it may happen that there are no equilibrium points, so that equilibrium existence becomes a basic preliminary property to prove. Other extensions of the quoted results on GAS are of importance. When considering the widely employed piece-wise-linear (pwl) neural models [22], [23], [10], infinite intervals with zero slope are present in the activations, making it of interest to drop the assumptions of strict increasingness and continuous first derivative for the activation. Moreover, it is desirable that the analysis of GAS be valid in the general case of both symmetric or nonsymmetric interconnection matrices \( T \). Due to tolerances in the electronic implementation, nonsymmetric \( T \) may origin from slight perturbations of an otherwise symmetric matrix. Nonsymmetries of \( T \) may also be deliberately introduced to accomplish special tasks [24] or may be related to the attempt to consider a more realistic model of some classes of neural circuits composed of the interconnection of two different sets of amplifiers (e.g., programming neural networks [10]). Finally, for actual...
implementations, it is needed that the conditions for GAS are robust with respect to variations of the system parameters [25]. Motivated by the above discussion, we analyze existence, uniqueness, and GAS of the neural network equilibrium point in the general case where the interconnection matrix $T$ is allowed to be symmetric or nonsymmetric and the neuron activation may be modeled by any continuous nondecreasing function. Notice that such activations may be unbounded (but not necessarily surjective), may have infinite intervals with zero slope as in a pwl model, or both. As such, the analysis is applicable to a large class of neural networks including the additive neural network model [20], the Hopfield network [26], and the Cellular Neural Network model [22]. Moreover, this class is closely related to neural networks for solving linear and quadratic programming problems [10]. A robustness analysis of conditions for GAS with respect to system parameter variations is also addressed.

The remainder of the paper is organized as follows: Section II introduces notation and gives the problem formulation and preliminary results. The main results are proved in Sections III and IV and concern conditions ensuring that there exists a unique GAS equilibrium point of the neural network for each constant input and for all nondecreasing neuron activations. These conditions rely on the concept of Lyapunov Diagonally Stable or Lyapunov Diagonally Semi-Stable matrices. More specifically, in Section III, existence and uniqueness of the equilibrium point are addressed by employing techniques from nonlinear circuit theory, while in Section IV, conditions on $T$ ensuring GAS are proved by means of Lyapunov functions of the generalized Lur’e-Postnikov type. Two important issues for practical application of the results are discussed in Section V, namely, the computational burden of checking conditions for GAS and their robustness against uncertainties on $T$. In Section VI, the results on GAS are applied to solve an open dynamical problem for the neural circuits for linear and quadratic programming problems in [10]. We prove that such circuits are GAS and hence, show a correct behavior for optimization also when the constraint amplifiers are dynamical, as it happens in any realistic circuit implementation. This general case could not be analyzed by the usual gradient-type approach based on an energy function concept, since a network with a nonsymmetric $T$ is obtained for dynamical constraints. Concluding remarks appear in Section VII.

II. NOTATION AND PRELIMINARIES

Notation

By $R^n$, we denote the real $n$-space. By $x = (x_1, \ldots, x_n)^T \in R^n$, we mean a column vector (the symbol $^T$ denotes transpose). In particular, $e_1, \ldots, e_n$ denote the standard basis vectors of $R^n$. If $A = [a_{ij}]$ is a given $n \times n$ matrix, $A^T$ means the transpose of $A$, $A^{-1}$ means the inverse of $A$, $[A]^S$ means the symmetric part of $A$ defined by $[A]^S = \frac{1}{2}(A^T + A)$, and $|A|$ means the absolute-value matrix given by $|A| = [a_{ij}]$. By $trA$, we denote the trace of $A$. By $diag[a_1, \ldots, a_n]$, we mean a diagonal matrix with diagonal entries $a_i$. If $A$ is a symmetric matrix, $A > 0$ ($A \geq 0$) means that $A$ is positive definite (positive semidefinite). Similarly, $A < 0$ ($A \leq 0$) means that $A$ is negative definite (negative semidefinite). If $A, B$ are symmetric matrices, $A > B$ ($A \geq B$) means that $A - B$ is positive definite (positive semidefinite).

By a factor of a given matrix $A > 0$, we mean any nonsingular matrix $A_f$ such that $A$ has the factorization $A = A_f^T A_f$. The minimum and maximum real eigenvalues of a symmetric matrix $A$ are denoted by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$, respectively. By $E_n$, we denote the $n \times n$ identity matrix. By $||x||$, we denote a vector norm, while $||A||$ denotes a matrix norm. In particular, $||x||_2 = [x_1^2 + \cdots + x_n^2]^{1/2}$ is the vector 2-norm, $||x||_\infty = \max_{1 \leq i \leq n} |x_i|$ is the vector $\infty$-norm, $||A||_2 = \lambda_{\max}^{1/2}(A^T A)$ is the matrix 2-norm and $||A||_F = [a_{11}^2 + a_{22}^2 + \cdots + a_{nn}^2]^{1/2}$ is the Frobenius norm. By $C^1$, we denote the set of functions with continuous first derivative, while $C^0$ denotes the set of continuous functions. If $G: R^n \rightarrow R^n$ is $C^1$, by $J_G(x)$, we mean the Jacobian of $G$ at $x$. If $F: R^n \rightarrow R$ is $C^1$, $\nabla F(x)$ denotes the gradient of $F$ at $x$. 

We will consider neural networks described by the system of nonlinear differential equations

$$\dot{x} = -Dx + T g(x) + I \quad (\dot{\cdot} = d/dt) \tag{1}$$

where $x \in R^n$, $D = \text{diag}(d_1, \ldots, d_n)$ is a constant $n \times n$ diagonal matrix with $d_i > 0$, $i = 1, \ldots, n$, $T$ is a constant $n \times n$ matrix, $g(x) = (g_1(x_1), \ldots, g_n(x_n))^T: R^n \rightarrow R^n$ is a locally Lipschitz continuous nonlinear mapping with $g(0) = 0$, and $I \in R^n$ is a constant vector.

The matrix $T$ is referred to as the interconnection matrix. The functions $g_i$ represent the neuron input-output activations, while $I$ describes constant inputs to the neural network. The diagonal entries of $D$ model neuron self-inhibitions.

We assume that $g$ belongs to the class $[G_{m}]$ defined by the property that $g \in [G_{m}]$ if for $i = 1, \ldots, n$, the functions $g_i: R \rightarrow R$ are monotonic nondecreasing. If there exist constants $\overline{G}_i, 0 < \overline{G}_i < \infty$, $i = 1, \ldots, n$, such that the incremental ratio for $g_i$ satisfies

$$0 \leq \frac{g_i(\xi_1) - g_i(\xi_2)}{\xi_1 - \xi_2} \leq \overline{G}_i$$

for each $\xi_1, \xi_2 \in R, \xi_1 \neq \xi_2$ and for $i = 1, \ldots, n$, we say that $g \in [G_{m}]$ in the finite sector case. In this case, we denote by $\overline{G} = \text{diag}(\overline{G}_1, \ldots, \overline{G}_n)$ the matrix of maximum allowable $g_i$ slopes. If otherwise, the incremental ratio for $g_i$ is nonnegative but is allowed to take arbitrarily large positive values, we say that $g \in [G_{m}]$ in the infinite sector case.

Functions $g \in [G_{m}]$, which are typically used in neural network models are depicted in Fig. 1. Sigmoidal (i.e., bounded and strictly increasing) functions as in Fig. 1(a) are considered in the Hopfield model [26] and in the additive neural network model [20]. Unbounded exponential-type functions (Fig. 1(c)) that model diode-like nonlinear devices, play an essential role to impose constraints satisfaction in neural networks for solving programming problems [10]. The pwl approximations of a sigmoidal or an exponential-type function (see Fig. 1(b), (d)) are of special interest, since they are widely employed to model neuron activations. Consider, e.g., the Cellular Neural...
Fig. 1. Typical nonlinear functions in the class \( \{G_m\} \). (a) \( C^1 \) bounded and strictly increasing (sigmoidal) function. (b) \( C^0 \) pwl counterpart of the sigmoidal function in (a). (c) \( C^1 \) strictly increasing unbounded exponential-type function which belongs to \( \{G_m\} \) in the infinite sector case. (d) Pwl counterpart of (c), which belongs to \( \{G_m\} \) in the finite sector case.

Networks [22] or the pwl model of the Hopfield network [23]. Also consider the approach in [10] to solve linear and nonlinear programming problems. It should be noticed that pwl functions of practical interest in neural network applications sigmoidal function in (a). (c) that (1) has a unique equilibrium point which is GAS. An equilibrium point of (1) tend towards e.g., \([27]\). By global attractivity, we mean that all trajectories locally stable in the sense of Lyapunov and globally attractive. The equilibrium point for (1). Before analyzing GAS, some remarks are in order to highlight nonphysical situations that may occur in case of unbounded \( g \).

Remark 1: If \( g \) is bounded, it is straightforward to show (e.g., by using Brouwer fixed point theorem [21]) that (1) has at least one equilibrium point. On the contrary, for unbounded \( g \), it may happen that there are no equilibrium points. For example, consider the first order equation \( \dot{x} = -x + g(x) \), where

\[
g(x) = \begin{cases} x^2 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}
\]

is such that \( g \in \{G_m\} \) in the infinite sector case. Each solution with initial condition \( x_0 > 1 \) at \( t = 0 \) blows up at \( t_0 = \log(x_0/(x_0 - 1)) > 0 \). It is worth noticing that the conditions for GAS we will find in Section IV also ensure that such nonphysical situations are ruled out and hence guarantee in particular well-posedness (see [29, p. 1016]) of the neural model (1).

We conclude this section by collecting definitions and results needed in the following developments. To prove existence and uniqueness of the equilibrium point, we make use of these concepts from topology.

**Definition 2:** A map \( H: R^n \rightarrow R^n \) is a homeomorphism of \( R^n \) onto itself if \( H \) is \( C^0 \), \( H \) is one-to-one, \( H \) is onto and the inverse map \( H^{-1} \) is \( C^0 \).

**Definition 3:** A map \( H: R^n \rightarrow R^n \) is a diffeomorphism of \( R^n \) onto itself if \( H \) is \( C^1 \), \( H \) is a homeomorphism of \( R^n \) onto itself and \( H^{-1} \) is \( C^1 \).

A characterization of homeomorphisms in \( \mathbb{R}^n \) has been given by Hadamard [30], where the concept is related to the properness of the map \( H \).

**Theorem 1 ([30]):** A locally invertible \( C^0 \) map \( H: \mathbb{R}^n \rightarrow \mathbb{R}^n \) is a homeomorphism of \( \mathbb{R}^n \) onto itself if and only if it is proper.

Recall that a \( C^0 \) map \( H \) is proper if the inverse image under \( H, H^{-1}(K) \), is compact for any compact set \( K \) in \( \mathbb{R}^n \). In practice, properness of \( H \) can be checked when some information is available on the divergence of the sequence \( \|H(x_p)\| \rightarrow \|x_p\| \rightarrow \infty \). In fact, a \( C^0 \) map \( H \) is proper if and only if \( \|H(x_p)\| \rightarrow \infty \) for \( ||x_p|| \rightarrow \infty \).

A similar result holds for diffeomorphisms, as stated in the next theorem due to Palais, which has been widely employed in nonlinear circuit analysis.

**Theorem 2 ([31]):** A \( C^1 \) map \( H: R^n \rightarrow R^n \) is a diffeomorphism of \( R^n \) onto itself if and only if \( \det J_H(x) \neq 0 \) for each \( x \in R^n \) and \( H \) is proper.

To prove GAS of the equilibrium point (suppose \( x^e = 0 \)), we will exploit the Lyapunov direct method. According to a classical result [27], GAS is guaranteed if there exists a \( C^1 \) Lyapunov function \( V: R^n \rightarrow R \) which is positive definite \( V(x) > 0 \) for \( x \neq 0, V(0) = 0 \) and radially unbounded \( V(x) \rightarrow \infty \) as \( ||x|| \rightarrow \infty \), and whose time-derivative \( V \) along solutions of (1), which is given by \( \dot{V}(x) = \nabla V(x)^T \dot{x} \), is negative definite \( \dot{V}(x) < 0 \) for \( x \neq 0, V(0) = 0 \).

In the results in Sections III-V, we will require for (1) that the interconnection matrix \( T \) (or a suitable matrix depending on \( T \)) belongs to one of the classes of matrices defined below. Algebraic properties of such matrices play a key role to prove existence, uniqueness, and GAS of the equilibrium point.
**III. EXISTENCE AND UNIQUEQNESS OF THE EQUILIBRIUM POINT**

The main result in this section (Theorem 3) gives a condition ensuring that the vector field associated with (1)

$$H(x) = -Dx + Tg(x) + I$$  \hspace{1cm} (2)

is a homeomorphism or a diffeomorphism of $R^n$ onto itself. Under this condition, existence and uniqueness of the equilibrium point $x^*$ of (1) are guaranteed for each $I$ as stated in Corollary 1.

**Proposition 1:** Suppose that either (a): $g \in \{G_m\}$ in the finite sector case and $-T + DG^{-1} \in P$ or (b): $g \in \{G_m\}$ in the infinite sector case and $-T \in P_0$. Then, $H$ is injective, i.e., $H(x_1) \neq H(x_2)$ for $x_1 \neq x_2$.

**Proof:** See Appendix A. \hspace{1cm} $\square$

Proposition 1 ensures injectivity of $H$. However, it does not guarantee surjectivity of $H$ (i.e., $H(R^n) \equiv R^n$) and hence, the existence of the equilibrium point for each $I$. The problem of equilibrium existence is of importance since, unless suitable conditions are satisfied by $T$, there may be no equilibrium points for some $I$ (cf. Remark 1). The following theorem and corollary state that if $-T + DG^{-1} \in LDS$ or $-T \in LDSS$, the existence of the equilibrium point is ensured for each $I$.

**Theorem 3:** Suppose that either (a): $g \in \{G_m\}$ in the finite sector case and $-T + DG^{-1} \in LDS$ or (b): $g \in \{G_m\}$ in the infinite sector case and $-T \in LDSS$. Then, (i) $H$ is a homeomorphism of $R^n$ onto itself; (ii) $H$ is a diffeomorphism of $R^n$ onto itself if $g \in C^1$.

We first discuss the principal consequence of Theorem 3, while postponing the theorem proof at the end of the section.

**Corollary 1:** Under the same hypotheses as in Theorem 3, (1) has a unique equilibrium point for each $I \in R^n$.

**Proof:** Theorem 3 ensures that $H$ is one-to-one and onto (see Definitions 2 and 3). Hence, the equation $H(x) = 0$ has a unique solution. \hspace{1cm} $\square$

Notice that the unique equilibrium point has the additional property of being a $C^0$ function of the input vector $I$ if $H$ is a homeomorphism as in point (i) of Theorem 3 or even a $C^1$ function if $H$ is a diffeomorphism as in point (ii).

**Remark 3:** There is an extensive literature on the analysis of algebraic equations modeling nonlinear dc networks (see, e.g., [1]–[3] and references therein) where problems similar to those in the present section are considered. To the authors knowledge, however, the results in Proposition 1, Theorem 3, and Corollary 1 cannot be derived from existing ones. In particular, results concerning nonlinear functions satisfying a finite sector condition do not seem to have received much attention.

**Proof of Theorem 3:** Claim (i) Let us show that $H$ is locally invertible and proper, so that the result in (i) follows from Theorem 1.

Assume (a). It is known that $-T + DG^{-1} \in LDS$ implies $-T + DG^{-1} \in P$ [33]. Then, from Proposition 1, $H$ is injective and in particular is locally invertible.

To show that $H$ is proper ($\lim_{x \to \pm \infty} H(x) = \pm \infty$), we suppose that $H(x) = -Dx + Tg(x)$ is proper. Suppose, for purposes of contradiction that $H$ is not a proper map. Then, there exists a sequence $\{x_p\} \to \infty$ such that $\|H(x_p)\| \to 0$, i.e.,

$$\|H(x_p)\| \leq M. \hspace{1cm} (3)$$

Now notice that the sequence $g(x_p)$ is necessarily unbounded. Indeed, if it were bounded, we would have $\lim_{p \to \infty} \|H(x_p)\| = \lim_{p \to \infty} \|Dx_p + Tg(x_p)\| = \infty$, since $\|Dx_p\| \to \infty$ and $\|Tg(x_p)\|$ is bounded. But this would contradict (3). Being $g(x_p)$ unbounded, there exists a subsequence $\{x_{p(k)}\} \to \infty$ such that

$$\|g(x_{p(k)})\| \to \infty. \hspace{1cm} (4)$$

We want to prove that (3) and (4) are in contradiction with $-T + DG^{-1} \in LDS$. If $-T + DG^{-1} \in LDS$, there exists $\alpha = \text{diag}(\alpha_1, \ldots, \alpha_n) > 0$ such that

$$[\alpha(I - DG^{-1})]^S = [\alpha T]^S - \alpha DG^{-1} \leq -\epsilon E_n < 0 \hspace{1cm} (5)$$

for sufficiently small $\epsilon > 0$ ($E_n$ is the identity matrix). Let us assume to employ the vector and matrix 2-norm and consider that

$$[\alpha g(x)]^T \hat{H}(x) = [\alpha g(x)]^T (-Dx + Tg(x)) = -g(x) \alpha Dx + g(x) [\alpha T]^S g(x) \leq -g(x) \alpha Dx + g(x) [\epsilon E_n + \alpha DG^{-1}] g(x) = -g(x) \alpha Dx + g(x) \alpha DG^{-1} g(x) - \epsilon \|g(x)\|^2 \hspace{1cm} (6)$$

where the inequality is a consequence of (5). By hypothesis $[\alpha g(x)]^T \hat{H}(x) = -g(x) \alpha Dx + g(x) \alpha DG^{-1} g(x)$ and then $g(x) \alpha Dx \leq -g(x) \alpha DG^{-1} g(x)$. Therefore, from (6)

$$\|g(x)\|^2 \leq \|g(x)\|^2 \hat{H}(x) \leq \|g(x)\|^2 \alpha DG^{-1} g(x) - \epsilon \|g(x)\|^2 \hspace{1cm} (7)$$

and then by taking absolute values and using Schwartz inequality,

$$\|g(x)\|^2 \leq \|g(x)\|^2 \hat{H}(x) \leq \|g(x)\|^2 \|\hat{H}(x)\|_2. \hspace{1cm} (7)$$

Now, by evaluating (7) along points $x_{p(k)}$ of the sequence previously constructed, we easily reach contradiction. To see this, notice that from (3) and (7)

$$\|g(x_{p(k)})\|_2 \neq 0 \text{ for large } k \text{ so that} \hspace{1cm} \|g(x_{p(k)})\|_2 \leq \frac{M}{\epsilon} \|\alpha\|_2 \hspace{1cm} (8)$$

However, (3) implies $\|g(x_{p(k)})\|_2 \neq 0$ for large $k$ so that

$$\|g(x_{p(k)})\|_2 \leq \frac{M}{\epsilon} \|\alpha\|_2 \hspace{1cm} (8)$$

Therefore, the result in (i) follows from Theorem 1.
which is clearly in contradiction with (4). This contradiction proves that $H$ and $H$ are proper maps and completes the proof of Claim (i) when (a) holds.

Assume (b). Since $-T \in \text{LDSS}$ implies $-T \in \mathcal{P}_0$ [33], local invertibility of $H$ is again a consequence of Proposition 1. To show that $H$ is proper, let us proceed by contradiction as in case (a) by constructing $x_{p}(a)$ for which (3) and (4) are valid. We have for $H$

$$\left[\alpha g(x)^T\right]^T H(x) = -g'(x)\alpha Dx + g'(x)\alpha T^S g(x)$$

$$\leq -g'(x)\alpha Dx = \sum_{i=1}^{n} \alpha_i d_i x_i g_i(x_i)$$

since $-T \in \text{LDSS}$ means that $[\alpha T]^S \leq 0$. For $g \in \{G_m\}, x_i g_i(x_i) \geq 0$ so that by using the $\infty$-norm, we easily get

$$\sum_{i=1}^{n} \alpha_i d_i \|x_i\| g_i(x_i) \leq \sum_{i=1}^{n} \alpha_i d_i \|g_i(x_i)\| \leq \sum_{i=1}^{n} \alpha_i d_i \|H(x)\|$$

$$\leq \text{tr}\|\alpha g(x)\|_{\infty}\|H(x)\|_{\infty}. \quad \text{(8)}$$

By evaluating (8) along points $x_{p}(k)$ and considering (3), we obtain

$$\sum_{i=1}^{n} \alpha_i d_{i} \|x_{i}\|_{i} \|g_{i}(x_{p}(k), i)\| \leq M \text{tr}\|\alpha g(x)\|_{\infty}$$

where $x_{p}(k), i$ denotes the $i$th component of $x_{p}(k)$. For sufficiently large $k$, $\|g(x_{p}(k))\|_{\infty} \neq 0$. Hence,

$$M \geq \sum_{i=1}^{n} \alpha_i d_{i} \|x_{i}\|_{i} \|g_{i}(x_{p}(k), i)\| \|\alpha g(x)\|_{\infty}$$

$$\geq \hat{d} \frac{\alpha}{\text{tr}} \|x_{p}(k), i'(k)\|$$

where $\hat{d} = \min\{d_i\} > 0, \hat{\alpha} = \min\{\alpha_i\} > 0$ and $i'(k)$ is the index that depends in general on $k$, such that $\|g(x_{p}(k))\|_{\infty} = \max_{i=1,\ldots,n} \|g_{i}(x_{p}(k), i)\| = \|g_{i}(x_{p}(k), i'(k))\|$. From (4), $\|g_{i}(x_{p}(k), i'(k))\| \leq \|g_{i}(x_{p}(k), i'(k))\|$ and $\|g_{i}(x_{p}(k), i'(k))\| \rightarrow \infty$. This also implies $\|x_{p}(k), i'(k)\| \rightarrow \infty$ and hence (9) yields a contradiction. The proof of Claim (i) is then completed.

Claim (ii) From Theorem 2 and from the fact that $H$ is proper as shown in the previous Claim (i) in this theorem, we see that Claim (ii) is proved if we show that det $J_H(x) \neq 0$ for each $x \in \mathbb{R}^n$. We have for $g \in C^1$

$$J_H(x) = -D + T \text{ diag}(g_1'(x_1), \ldots, g_n'(x_n)).$$

If (a) holds, then $0 \leq g_i'(x_i) \leq \tilde{G}_i$ and, as seen before, $-T + D\tilde{G}^{-1} \in \mathcal{P}_0$. Hence, det $J_H(x) \neq 0$ is an immediate consequence of Lemma 2 in Appendix A. If (b) holds, we have $g_i'(x_i) \geq 0$ and $-T \in \mathcal{P}_0$. Thus, once more det $J_H(x) \neq 0$ follows from Lemma 2.

IV. GLOBAL ASYMPTOTIC STABILITY OF THE EQUILIBRIUM POINT

Theorem 4, which is the main result in this section, provides a condition ensuring that (1) has a unique equilibrium point that is GAS. This result focuses on the finite sector case and parallels part (a) of Theorem 3. The infinite sector case is also discussed (Theorem 5) and related to previous work.

Conditions for GAS are here provided by using the standard direct method of Lyapunov. The distinguished feature of the present analysis is the class of Lyapunov functions that are employed. Inspired by the work on the absolute stability of nonlinear control systems of the Lur’e type, we make use of Lyapunov functions of the generalized Lur’e-Postnikov type [27], [34]. In particular, we select candidate Lyapunov functions of the following form:

$$V(x) = x^T P x + \sum_{i=1}^{n} \beta_i \int_{0}^{x_i} G_i(\rho) d\rho \quad \text{(10)}$$

where $P$ is a symmetric and positive definite matrix, $\beta_i$ are positive constants, and $G_i(\rho)$ are the nondecreasing functions $G_i(\rho) = g_i(\rho + x_i^0) - g_i(x_i^0)$, $x_i^0$ being the unique equilibrium point of (1). Such functions differ from those currently used in the analysis of neural networks [26] and allows one to cover the case of nonsymmetric interconnection matrices $T$.

Let us consider the following algebraic lemma, which plays a key role in the proof of Theorem 4 for showing the negative definiteness of the time-derivative of the Lyapunov function.

**Lemma 1:** Let $T$ be an $n \times n$ matrix and $D, G$ be $n \times n$ positive definite diagonal matrices such that $-T + D\tilde{G}^{-1} \in \mathcal{P}_0$, i.e., $[\alpha(-T + DG^{-1})]^S = 0$ for some $n \times n$ diagonal matrix $\alpha > 0$. Then, it is possible to find a scalar $k > 0$, a nonsingular $n \times n$ matrix $Q_1$, an $n \times n$ matrix $\Gamma$ and three positive definite $n \times n$ matrices $\gamma_0 = \text{diag}(\gamma_0, \ldots, \gamma_0)$, $\gamma_0 = P^0, \Gamma = \Pi^0$ such that

$$\begin{cases} P(-D) + (-D)P = -Qr^4 \\ P(-T) = -Q + k\alpha(-T) + \gamma_0 \\ \Gamma + \Pi = 2k[\alpha(-T)]^S + 2\gamma_0 G^{-1} \end{cases} \quad \text{(11)}$$

**Proof:** See Appendix B. \hspace{1cm} \square

**Theorem 4:** Suppose that $g \in \{G_m\}$ in the finite sector case and $-T + D\tilde{G}^{-1} \in \mathcal{P}_0$. Then, for each $I \in \mathbb{R}^n$, (1) has a unique equilibrium point which is GAS.

**Proof:** Since $-T + D\tilde{G}^{-1} \in \mathcal{P}_0$, from Corollary 1, (1) has a unique equilibrium point $x^e$. By means of the coordinate translation $z = x - x^e$, (1) can be put into the form

$$\dot{z} = -T z + G(z) \quad \text{(12)}$$

where $G(z) = \{G_1(z_1), \ldots, G_n(z_n)\}$ and $G_i(z_i) = z_i + x_i^0 - g_i(x_i^0)$. We have $G(0) = 0$ and $G \in \{G_m\}$ in the finite sector case with the same $G$ as the function $g$. System (12) has a unique equilibrium at $z = 0$.

Furthermore, $-T + D\tilde{G}^{-1} \in \mathcal{P}_0$ means that there exists $\alpha = \text{diag}(\alpha_1, \ldots, \alpha_n) > 0$ such that $[\alpha(-T + DG^{-1})]^S = 0$. To show that $z = 0$ is GAS for (12), consider the candidate Lyapunov function of the generalized Lur’e-Postnikov type

$$V(z) = z^T P z + 2k \sum_{i=1}^{n} \alpha_i \int_{0}^{x_i} G_i(\rho) d\rho \quad \text{(13)}$$

where $P$ is an $n \times n$ matrix with $P = P^0 > 0$ and $k$ is a positive constant. It is clear that $V$ is positive definite and radially unbounded, so that to prove GAS it suffices to show...
that the time derivative $\dot{V}$ of $V$ along solutions of (12) is negative definite for a suitable choice of $P$ and $k$ [27]. We have

$$\begin{align*}
\dot{V}(z) &= [\nabla V(z)]^T \dot{z} \\
&= [2Pz + 2k\alpha z \theta(z)]^T(-Dz + TG(z)) \\
&= 2z^TP(-D)z + 2z^TPG(z)z + 2kG(z) \alpha(-D)z \\
&+ 2kG(z) \alpha(TG(z)) \\
&= z^TP(-D)z + (-D)Pz - 2z^TP(-T) \\
&- k\alpha(-D)z - 2kG(z) \alpha(-D)z + 2kG(z) \alpha(TG(z)).
\end{align*}$$

Recall that $|G_i(z_i)| \leq \bar{G}_i |z_i|$ and $z_i G_i(z_i) \geq 0$. Hence, for any choice of the constants $\gamma_i > 0, i = 1, \ldots, n$, we have

$$G^T(z) \gamma_0 G^{-1}(z) \leq z^T \gamma_0 G(z).$$

The first three terms in $\dot{V}$ can be made to form a perfect square plus a nonpositive term by using the previous Lemma 1. Indeed, by Lemma 1, there exist matrices $P = P^T > 0, P = P^T > 0, \Pi > 0, \Gamma, Q$ (with $Q \neq 0$), and a constant $k > 0$ such that the first three terms can be written as

$$\begin{align*}
\dot{V}(z) &= z^T[P(-D) + (-D)P]z \\
&- 2z^T[P(-T) - k\alpha(-D) - \gamma_0 G(z)] \\
&- G^T(z) (2k(\alpha(-T))^T + 2\gamma_0 G^{-1}) G(z) \\
&- 2z^T \gamma_0 G(z) - G^T(z) \gamma_0 G^{-1}(z).
\end{align*}$$

Therefore

$$\begin{align*}
\dot{V}(z) &= -[Q^T z - \Gamma G(z)]^T Q z - \Gamma G(z) ] - G^T(z) \Pi G(z).
\end{align*}$$

Notice that the last term in $\dot{V}$ in (15) is nonpositive as a consequence of (14). To show that $\dot{V}$ is negative definite, consider that $G(z) \neq 0$, then $G(z) \Pi G(z) < 0$ since $\Pi > 0$, so that $\dot{V}(z) < 0$. Otherwise, if for some $z \neq 0$, we have $G(z) = 0$, then $\dot{V}(z)$ reduces to $-z^T Q Q z < 0$ since $Q$ is nonsingular, and hence $Q Q^T$ is positive definite. Therefore, $\dot{V}(z) < 0$ for every $z \neq 0$ and also $\dot{V}(0) = 0$, which means that $\dot{V}$ is negative definite. This fact implies that $z = 0$ is GAS for (12) and hence that $z^2$ is GAS for (1). The proof of Theorem 4 is thus complete.

Theorem 4 represents a generalization of the results on GAS in [21]. Indeed in [21], where use is made of Lyapunov functions of the form $\Sigma_{i=1}^{n} \beta_i \int_{0}^{G_i(\rho)} G_i(\rho) d\rho$ (the second term in $V(z)$ in (10)), only sigmoidal (bounded and strictly increasing) activations are considered. On the contrary, Theorem 4 covers also the case of unbounded activations and/or those with infinite intervals with zero slope, which are of special interest, e.g., for pwl models.

In the infinite sector case, we have proved GAS if $-T \in$ LDSS only with respect to the smaller class of strictly increasing activations. The following result holds:

**Theorem 5:** Suppose that $q \in \{G_m\}$ in the infinite sector case and that $\alpha_i, i = 1, \ldots, n$, are strictly increasing. If $-T \in$ LDSS, then for each $I \in \mathbb{R}^n$, (1) has a unique equilibrium point which is GAS.

For $-T \in$ LDSS, the existence of a unique equilibrium is ensured by part (b) of Theorem 3, while boundedness of solutions (and hence, no finite forward escape time) and GAS can be proved by using radially unbounded and positive definite Lyapunov functions of the form $\Sigma_{i=1}^{n} \alpha_i \int_{0}^{G_i(\rho)} G_i(\rho) d\rho$ as in [21], where $\alpha = \text{diag}(\alpha_1, \ldots, \alpha_n) > 0$ is such that $[\alpha(-T)]^T > 0$. The detailed proof of Theorem 5 is omitted for brevity.

**Remark 4:** As mentioned above, Lyapunov functions of the Lur’e-Postnikov type have been extensively used for the absolute stability problem of Lur’e control systems [27], [34]. Indeed, it can be easily shown that system (1) can be written as a standard multivariable Lur’e system, so that the classical matrix Popov stability criterion (see [27], pp. 216–217) can be applied to determine conditions ensuring GAS of the equilibrium point of (1). However, this application is not satisfactory because it requires to check at each frequency if a certain matrix is positive definite and more importantly, to assume the nonsingularity of $T$. This last assumption derives from the use of the matrix extension of the Popov-Yakubovich-Kalman Lemma due to Anderson in [36], which requires that $(-D,T)$ is a controllable pair. Such an unnecessary hypothesis is avoided in our proof because we use Lemma 1, which relies on rearranging in a different way the terms of the time-derivative of the Lyapunov function. It is worthwhile to notice that $T$ may be singular in some important practical cases as shown in Section VI.

**Remark 5:** The use of frequency domain methods for assessing GAS of nonlinear systems of the Lur’e type has been employed since a long time (see, e.g., [35], [27]). Recently, a thorough investigation of the application of such methods to a general class of neural circuits has been given in [17]. More precisely, the multivariable version of the small gain and circle criteria [27], [34] is used to provide conditions for GAS of these general circuits. However, taking into account that the Popov criterion generally gives better results than the small gain and circle criteria [27], [34] is used to provide conditions for GAS of these general circuits. However, taking into account that the Popov criterion may become hard from a computational point of view, and therefore, the results in [17] should be preferable.

**V. CHECKING CONDITIONS FOR GAS AND SOME ROBUSTNESS ISSUES**

In the previous section, we have given conditions for ensuring GAS of the neural network (1). These conditions require to check if either the matrix $-T + DG^{-1} \in$ LDS or $-T \in$ LDSS. Thus, two important issues must be discussed for practical applications of these results: 1) The computational burden of checking these conditions; 2) The robustness of...
these conditions against uncertainty on the system matrices $T$, $D$, and $G$.

A. Checking LDS and LDSS Conditions

Let us consider the problem of verifying if a certain matrix is LDS or LDSS. This problem has been investigated since a long time by following two different ways: 1) Finding classes of matrices that can be shown to be LDS or LDSS by checking simple algebraic conditions; 2) Developing algorithms that provide a numerical answer to the problem.

It is well known that a $P$ matrix is not necessarily LDS. However, for several classes of matrices, these two conditions are equivalent (see, e.g., [37]), so that checking LDS amounts to test a finite number of simple algebraic conditions. Two of these classes are of special interest in neural network models, namely, symmetric matrices (Definition 5) and matrices of class $Z$, i.e., matrices whose off-diagonal entries are nonpositive. In neural models, the case where $-T \in Z$ corresponds to the important class of cooperative neural networks [20], [38], [24]. The following result concerning the finite sector case is an immediate consequence of Theorem 4.

**Corollary 2:** Suppose that $g \in \{G_m\}$ in the finite sector case. Then, for each $I \in \mathbb{R}^n$, (1) has a unique equilibrium point that is GAS if one of the following conditions is satisfied: (i) The neural network is symmetric ($T = T^T$) and $-T + DG^{-1} > 0$; (ii) The neural network is cooperative ($-T \in Z$) and $-T + DG^{-1}$ is an M-matrix.

**Proof:** The result follows from Theorem 4 and from the fact that condition (i) and (ii) imply that $-T + DG^{-1} \in \text{LDS}$.

For a general matrix, it is only possible to have a numerical answer to the problem of verifying LDS or LDSS conditions. The distinguished feature of this problem is that it requires the solution of a convex optimization problem and hence, efficient numerical algorithms can be used (see, e.g., [39], [40]). In order to justify this fact, we first recall a basic property of symmetric matrices.

**Property 1:** The set $S$ of positive definite symmetric matrices is convex, i.e., if $S_1, S_2 \in S$ then $(1-\lambda)S_1 + \lambda S_2 \in S$ for each $\lambda \in [0, 1]$.

Let us now focus on the LDS condition. From Definition 5, it follows that $-T + DG^{-1} \in \text{LDS}$ if and only if the set

$$A = \{\alpha = \text{diag}(\alpha_1, \ldots, \alpha_n) > 0 : \alpha(-T + DG^{-1})S > 0\}$$

is not empty. Now, exploiting Property 1, it is easily realized that the set $A$ is convex. Thus, checking if $-T + DG^{-1} \in \text{LDS}$ is equivalent to look for the existence of a feasible solution of a convex optimization problem of the form

$$\min_{\alpha \in A} 1$$

Recently, interior point polynomial methods for convex programming have been developed [41]. These methods have been shown to be very effective in practice especially when the constraints are given as linear matrix inequalities (see [42]), as it happens for the set $A$.

B. Robustness of Conditions for GAS

Our goal is to give a measure on how the matrix $-T + DG^{-1}$ (resp., $-T$) can differ from some given nominal value without losing the LDS (resp., LDSS) condition. To this end, we employ methods typically used for the robustness analysis of matrices with respect to parametric variations, a problem that has recently been addressed in several contributions (see, e.g., [43]–[45]). We focus on the significant case where the interconnection matrix $T$ is perturbed and is modeled as

$$T = T^0 + \Delta T$$

where $T^0$ denotes the nominal matrix and $\Delta T$ represents its deviation, while the other two matrices are not affected by uncertainty, i.e., $D = D^0$ and $G = G^0$.

Thus, the problem is to determine a measure of the admissible deviations $\Delta T$ preserving either the LDS or the LDSS condition. Obviously, the results strongly depend on the structure assumed for the matrix $T$. Here, we consider two different classes of matrices. The first one is characterized by a bound on the Frobenius norm of $\Delta T$, i.e.,

$$\Delta T_F(\rho) = \{\Delta T = [\delta t_{ij}] \in \mathbb{R}^{n \times n} : \|\Delta T\|_F \leq \rho, \rho > 0\}$$

while the second is the well-known interval model of uncertain matrices (see, e.g., [44], pp. 242–244)

$$\Delta T_I(\rho) = \{\Delta T = [\delta t_{ij}] \in \mathbb{R}^{n \times n} : \Delta T = \bar{\rho} \Sigma; \rho > 0; \Sigma = [\bar{\sigma}_{ij}, \sigma_{ij}] \in \{\bar{\sigma}_{ij}, \sigma_{ij} \leq 0, \sigma_{ij} \geq 0\}. \quad (16)$$

Let us concentrate on the LDS condition and first consider the class $\Delta T_I(\rho)$. By employing an idea similar to one in [46], we derive the following:

**Proposition 2:** Let $T^0$ be an $n \times n$ matrix and $D^0, G^0$ be $n \times n$ positive definite diagonal matrices such that $-T^0 + D^0(G^0)^{-1} \in \text{LDS}$, i.e., $[\alpha(-T^0 + D^0(G^0)^{-1})]S > 0$ for some $n \times n$ diagonal matrix $\alpha > 0$. Then, $[\alpha(-T^0 + D^0(G^0)^{-1})]S > 0$ for each $n \times n$ matrix $\Delta T \in \Delta T_I(\rho)$ if $\rho < \rho_F = \frac{1}{||F||_2}$ \quad (17)

where $F \in \mathbb{R}^{n^2 \times n}$ is the matrix

$$F = [(L^1)^{-1}[\alpha e_{(i)}e_{(i)}^T]S L^{-1}] \cdots [(L^1)^{-1}[\alpha e_{(i)}e_{(i)}^T]S L^{-1}] \cdots [(L^1)^{-1}[\alpha e_{(i)}e_{(i)}^T]S L^{-1}]^T.$$ \quad (18)

Here, $L \in \mathbb{R}^{n \times n}$ is any factor of $[\alpha(-T^0 + D^0(G^0)^{-1})]S$, i.e.,

$$L^T L = [\alpha(-T^0 + D^0(G^0)^{-1})]S \quad (19)$$

and $e_{(i)}$ is the $i$th standard basis vector.
FORTI AND TESI: NEW CONDITIONS FOR GLOBAL STABILITY OF NEURAL NETWORKS

Proof: Observing that
\[ \begin{align*}
&\left[\alpha(-T^0 - \Delta T + D^0(G^0)^{-1})\right]S \\
&= \left[\alpha(-T^0 + D^0(G^0)^{-1})\right]S - [\alpha\Delta T]S
\end{align*} \]
and employing (19), it turns out that the condition \([\alpha(-T^0 - \Delta T + D^0(G^0)^{-1})]S > 0\) is equivalent to
\[ E_n > (L')^{-1}\left[\alpha\Delta T\right]^S S^{-1} \]
where the last equality follows from (18). Now, employing (20), we can rewrite relation (22) in the simplified form
\[ ||[\delta t_{ij} E_n]||_2 > ||\delta T||_F \]
which in turn completes the proof.

Notice that evaluating \(\rho_F\) only requires the computation of the \(\|\cdot\|_2\) of a rectangular matrix, a task that can be efficiently performed in a numerical way (see, e.g., [47]). However, the bound is not tight in general.²

Conversely, a tight bound can be obtained for the interval matrices \(AT \in A'\&(PF)\) such that
\[ \left[\alpha(-T^0 - \Delta T + D^0(G^0)^{-1})\right]S > 0 \]
for some diagonal matrix \(\alpha > 0\). Then, the result in Theorem 4 continues to hold for the perturbed matrix \(T + \Delta T\), provided \(||\Delta T||_F < \rho_F\) with \(\rho_F\) given in (17), or \(\Delta T \in \Delta T(\rho)\) with \(\rho < \rho_1\) and \(\rho_1\) given in (21).

We remark that bounds similar to (17) and (21) and a result analogous to that in Theorem 6 can be obtained also for the LDSS condition. Moreover, by employing similar techniques, robustness bounds can also be provided in cases where the matrices \(D\) and \(G\) are perturbed.

VI. APPLICATION TO LINEAR AND QUADRATIC PROGRAMMING NEURAL NETWORKS

The previous results are applied here to analyze GAS for linear and quadratic programming neural networks [10] in the general case of practical importance where the constraint amplifiers are dynamical.

A. GAS of Programming Neural Networks

A quadratic programming problem can be formulated as follows: Minimize the scalar function
\[ \phi(v) = a^T v + \frac{1}{2} v^T Q v \]
subject to the affine constraints (intended componentwise)
\[ f(v) = Bv - c \geq 0 \]
where \(f: \mathbb{R}^q \to \mathbb{R}^d\), \(a, v \in \mathbb{R}^d, c \in \mathbb{R}^d, B \in \mathbb{R}^{p \times q}, Q = Q^T \in \mathbb{R}^{q \times q}\) and \(Q > 0\) for quadratic programming or \(Q = 0\) in the special case of linear programming.

The neural circuit proposed in [10] for solving such problems is composed of the interconnection of two sets of amplifiers: Variable amplifiers obeying \((i = 1, \cdots, q)\)
\[ C_i \frac{dv_i}{dt} + v_i = -a_i - \sum_{j=1}^q c_{ij} v_j - \sum_{j=1}^p b_{ij} v_j^2 \]
and constraint amplifiers satisfying \((j = 1, \cdots, p)\)
\[ f_j(v) = -c_j + \sum_{i=1}^q b_{ji} v_i \]
The variable amplifiers are linear and dynamical (the dynamics is modeled by the input capacitance $C_i > 0$). A finite amplifier input resistance $R_i > 0$ is considered, as in Section IV in [10]. On the contrary, the constraint amplifiers are nonlinear since their input-output characteristic $g_j$ corresponds to the unbounded pwl diode-like function

$$g_j(\rho) = \begin{cases} 0 & \text{if } \rho \geq 0 \\ R_j^0 \rho & \text{if } \rho < 0 \end{cases} \tag{26}$$

where $R_j^0 > 0$ is the gain for negative $\rho$.

In (24) and (25) and in the dynamical analysis in [10], the constraint amplifiers are assumed memoryless (i.e., it is supposed that the constraints are updated instantaneously). However, as noticed also in [10], this is preserved if the variable amplifier input capacitances $C_i$ are chosen much larger than the perturbing parasitic capacitances of the constraint amplifiers. However, the problem at hand is to prove GAS also for the $(p+q)$th order model.

By means of Theorem 4, we now prove not only that the higher order system is GAS for small perturbing capacitances of the constraint amplifiers but, more importantly, we prove that GAS holds for any (positive) value of the input capacitances of the constraint amplifiers. This rather unexpected result has important consequences. (a) The network proposed in [10] shows a robustness degree with respect to constraint amplifier parasitics. (b) It is not necessary to introduce very high $C_i$ for the variable amplifiers, which would render unpractically slow the circuit dynamics. Instead, the $C_i$ of the variable amplifiers can be of the same order as the System (31) is of the form of system (1). The function $g$ coincides with the nonlinear function modeled by a capacitance $C_j^0 > 0$. This model coincides with the memoryless model in [10], by letting $C_{ji}^0 = 0$ (see Fig. 2(a)). The constraint amplifier in Fig. 2(b) obeys

$$\frac{C_j^0}{R_j^0} \frac{dv^c_j}{dt} + \frac{v^c_j}{R_j^0} = -I$$

$$v^c_j = h(v^c_j) \tag{27}$$

where from (26)

$$h(\rho) = \begin{cases} 0 & \text{if } \rho \geq 0 \\ \rho & \text{if } \rho < 0 \end{cases} \tag{28}$$

Then, we get the following $(p+q)$th order system for the neural circuit (see (23) and (27)–(29)).

$$\begin{align*}
C_i \frac{dv_i}{dt} + \frac{v_i}{R_i} &= -a_i - \sum_{j=1}^{q} \sum_{m=1}^{p} b_{ij} h(v^c_j) \\
&= i = 1, \cdots, q
\end{align*} \tag{30}$$

$$C_j^0 \frac{dv^c_j}{dt} + \frac{v^c_j}{R_j^0} = -e_j + \sum_{i=1}^{p} b_{ji} v_i$$

In matrix-vector notation (30) reads as follows:

$$\begin{bmatrix}
\dot{v} \\
\dot{v}^c
\end{bmatrix} = -\begin{bmatrix}
R_i^{-1} C_i^{-1} & 0 \\
0 & (R_i^{-1} C_i)^{-1}
\end{bmatrix}
\begin{bmatrix}
v \\
v^c
\end{bmatrix}$$

$$+ \begin{bmatrix}
C_i^{-1} & 0 \\
0 & (C_i)^{-1}
\end{bmatrix}
\begin{bmatrix}
-Q & -B^T \\
B & 0
\end{bmatrix}
\begin{bmatrix}
h(v) \\
h(v^c)
\end{bmatrix}$$

$$+ \begin{bmatrix}
-C_i^{-1} a \\
-(C_i)^{-1} e
\end{bmatrix} \tag{31}$$

where we have let $v = (v_1, \cdots, v_q)^T$, $v^c = (v^c_1, \cdots, v^c_q)^T$, $R_i^{-1} = \text{diag}(1/R_1, \cdots, 1/R_q)$, $C_i^{-1} = \text{diag}(1/C_1, \cdots, 1/C_q)$ $(R_i^{-1} C_i)^{-1}$ and $(C_i)^{-1}$ are defined similarly) and $h(v) = (h(v_1), \cdots, h(v_q))^T$.

System (31) is of the form of system (1). The function $g$ coincides with the nonlinear function

$$g = \begin{bmatrix}
v \\
h(v^c)
\end{bmatrix} \tag{32}$$

and is such that $g \in \{G_m\}$ in the finite sector case. In particular, from (29), $G = E_m$. The matrix $T$ is given by

$$T = \begin{bmatrix}
C_i^{-1} & 0 \\
0 & (C_i)^{-1}
\end{bmatrix}
\begin{bmatrix}
-Q & -B^T \\
B & 0
\end{bmatrix} \tag{33}$$

Notice that $T$ is necessarily nonsymmetric due to the presence of $B$ and $-B^T$, so that the energy function approach as in a gradient-type method [10], [26], [48] is not applicable. Moreover, notice that $T$ is singular in the most interesting case where $p > q$ (more constraints than variables). Hence, the classical matrix Popov stability criterion (see [27, pp.
216–217]) cannot be directly applied as discussed in Remark 4. Being \( G = E_n \), we have
\[
-T + DG^{-1} = \begin{bmatrix} C^{-1} & 0 \\ 0 & (Cc)^{-1} \end{bmatrix} \begin{bmatrix} Q & B^t \\ -B & 0 \end{bmatrix} \\
+ \begin{bmatrix} R^{-1}C^{-1} & 0 \\ 0 & (R^t)^{-1}(Cc)^{-1} \end{bmatrix}.
\]
Let us verify that \( -T + DG^{-1} \in \text{LDS} \), so that Theorem 4 yields the desired result that (31) is GAS. Indeed, by choosing
\[
\alpha = \begin{bmatrix} C & 0 \\ 0 & C' \end{bmatrix} > 0
\]
we get
\[
\left[ \alpha (-T + DG^{-1}) \right]^{S} = \begin{bmatrix} Q & B^t \\ -B & 0 \end{bmatrix}^{S} + \begin{bmatrix} R^{-1} & 0 \\ 0 & (R^t)^{-1} \end{bmatrix}
\begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} R^{-1} & 0 \\ 0 & (R^t)^{-1} \end{bmatrix}
\]
which is positive definite since \( Q > 0 \) for a quadratic programming problem or \( Q = 0 \) for a linear programming problem and \( R^{-1} > 0, (R^t)^{-1} > 0 \).

Let us now consider the robustness of GAS against variations of the system parameters. We analyze this problem for the linear programming problem case. As the extension to quadratic programming is straightforward. To this purpose, let us model the matrices \( B \) and \( -B^t \) in (33) as
\[
B = B^0 + \Delta B(D) \\
-B^t = -(B^0)^t + (\Delta B(2))^t
\]
where \( B^0 \) is the nominal matrix and \( \Delta B = [\Delta B(D)] \in R^{p \times 2q} \) describes the deviations with respect to the nominal value. All the other matrices are assumed to be fixed at their nominal values and, to simplify notation, are still denoted with the same symbols.

By employing the results in Section V, our goal is to give a measure of the deviations \( \Delta B \) preserving GAS. From the above analysis, we observe that this happens if the matrix
\[
\begin{bmatrix} R^{-1} & -(B^0)^t + (\Delta B(2))^t \\ B^0 & (R^-)^{-1} \end{bmatrix} - \begin{bmatrix} 0 & (\Delta B(2))^t \\ \Delta B(D) & 0 \end{bmatrix}
\]
is LDS. Thus, identifying the matrices \( \alpha(-T + DG^0)^{-1} \) and \( \alpha \Delta T \) in Proposition 2 with the first and second matrix of (34), respectively, and exploiting the special structure of (34), it is straightforward though tedious to show that GAS is ensured for each \( \Delta B \) such that
\[
\|\Delta B\|_F < \rho_F = \min \left\{ \sqrt{\frac{2}{\max_{i=1, \ldots, q} R_i} \left( \frac{p}{\sum_{j=1}^q R_j^2} \right)} \right\}.
\]
Let us now consider the following interval model:
\[
\Delta B_f(\rho) = \Delta B \in R^{p \times 2q}; \Delta b_{ij}^{(1)} \in [-\rho |b_{ij}^0|, \rho |b_{ij}^0|], \\
\Delta b_{ij}^{(2)} \in [-\rho |b_{ij}^0|, \rho |b_{ij}^0|, \rho > 0].
\]
Notice that \( \rho \) gives here a measure of the admissible tolerance for each element. For instance, \( \rho = 0.1 \) means that each component implementing an element of \( -B \) or \( B^t \) is allowed for a variation of 10%. Using Proposition 3 and again exploiting the special structure of (34), it turns out that GAS is preserved for each \( \Delta B \in \Delta B_f(\rho) \) if
\[
\rho < \rho_1 = \frac{1}{\|R^{1/2}B^0(R)^{-1/2}\|_2}
\]
where \( R^{1/2} = \text{diag}(1/R_1^{1/2}, \ldots, 1/R_q^{1/2}) \), \( (R)^{-1/2} = \text{diag}(1/(R_1)^{1/2}, \ldots, 1/(R_q)^{1/2}) \) and \( |B^0| \) is the absolute-value matrix.

B. Simulation Results and Robustness

To verify the results in this section, we have considered a prototype linear programming problem [26], which amounts to minimize the linear function \( $(w)$ \) subject to the four affine constraints described by \( f(v) = Bv - e \geq 0 \), where
\[
a = [-1, 0]^t; \quad B = [-5/12 \ -5/2 \ 1 \ 0]^t \\
e = [-35/12 \ -35/2 \ -5 \ -5]^t.
\]
The feasibility region defined by the constraints is depicted in Fig. 3(a). The constrained minimum is \( v^* = (5, 5) \). For this problem, we have solved system (23)–(25), which corresponds to memoryless constraints, by using a fourth-order standard Runge-Kutta numerical integration algorithm, and the result is reported in Fig. 3(a). This result is compared with that obtained in the case where also the constraints are dynamical (system (30)), see Fig. 3(b). As expected, in the latter case, the trajectory may slightly enter the region where constraints are violated. However, after some quickly decaying oscillations, the trajectory eventually converges towards the constrained minimum also for dynamical constraints, thus confirming the property of GAS predicted by the theory.

Other simulations performed also for quadratic programming problems, all were in agreement with GAS. Although GAS is true for any choice of the ratio \( C/C^0 \) between the variable amplifier capacitance and the constraint amplifier capacitance, it is not convenient, as intuitively reasonable, to slow down too much the constraints with respect to the variables, since in this case, the trajectory may exceedingly enter the region where constraints are violated prior to converge to the minimum. Our simulations showed that a ratio \( C/C^0 \) on the order of 5+20 is a good compromise between convergence speed and the need to obtain trajectories that remain close to the feasibility region boundaries and do not oscillate too much before converging.

Finally, let us consider the robustness of the studied linear programming problem by computing the two bounds (35) and (36), which measure the admissible deviations \( \Delta B = [\Delta B^{(1)} \Delta B^{(2)}] \in R^{p \times 4} \) with respect to the nominal value
Thus including among others the important case of pwl neural models.

The conditions for GAS have been developed via the Lyapunov direct method by employing a class of Lyapunov functions different from those currently used in the analysis of neural networks. More precisely, we have considered Lyapunov functions of the generalized Lur’e-Postnikov type, which enabled us to cover the case of nonsymmetric $T$.

The applicability of the results has been demonstrated by solving an open problem on GAS for the class of neural circuits introduced in [10] for solving linear and quadratic programming problems. Such a class is characterized by pwl unbounded activations (constraint neurons) and by nonsymmetric and singular $T$ (for dynamical constraint amplifiers). The principal conclusion here obtained is that such circuits are GAS also in case of dynamical constraints and, as a consequence, their analogical convergence speed can be exploited to the fullest extent for on-line optimization. It is worth pointing out that the standard approach based on the energy function concept cannot be employed for the general case of dynamical constraints since $T$ is not symmetric.

**APPENDIX A**

We prove Proposition 1 in two steps. The first step (Lemma 2) shows that the matrix $-TK + D$ is nonsingular for all values that the incremental ratio for the neuron activations $g_i$ can attain (such values are the diagonal entries of $K$). This is proved by exploiting some basic algebraic properties of matrices in the classes $P$ and $P_0$. The second step is a standard argument for proving injectivity under the nonsingularity of $-TK + D$.

**Lemma 2**: Let $T$ be an $n \times n$ matrix and $D, \overline{G}$ be $n \times n$ positive definite diagonal matrices. If $-TK + D \overline{G}^{-1} \in P$, then for each $n \times n$ diagonal matrix $K$ such that $0 \leq K \leq \overline{G}$, we have

$$\det(-TK + D) \neq 0.$$  

(38)

If $-T \in P_0$, then (38) holds for each $n \times n$ diagonal matrix $K \geq 0$.

**Proof of Lemma 2.** Assume that $-T + D \overline{G}^{-1} \in P$. Let $K = \text{diag}(k_1, \cdots, k_n)$ and consider three distinct cases.

1. $K = 0$. In this case, the result is obvious, since $D > 0$.
2. $k_i > 0$, $i = 1, \cdots, n$. This implies det $K \neq 0$. Notice that $-TK + D = (-TK + D K^{-1})K$. Then, it suffices to show that $\det(-TK + D K^{-1}) \neq 0$. To this end, let us prove that $-TK + D K^{-1} \in P$. We have

$$-TK + D K^{-1} = -T + D \overline{G}^{-1} + D(K^{-1} - \overline{G}^{-1})$$

where $K^{-1} - \overline{G}^{-1} \geq 0$, since $0 < K \leq \overline{G}$. By hypothesis, $-T + D \overline{G}^{-1} \in P$. The sum of a matrix of class $P$ and a diagonal matrix with nonnegative entries is also a matrix of class $P$, as a consequence of property 2 in Theorem 1.1
in [32]. Therefore, \(-T + DK^{-1} \in P\) and then \(\det(-T + DK^{-1}) > 0\).

(3) \(k_i = 0\) for some (but not all) \(i = 1, \ldots, n\). Without loss of generality, we can assume \(k_i > 0, i = 1, \ldots, m\) and \(k_j = 0, j = m + 1, \ldots, n\), where \(1 < m < n\) (we are always brought back to this case by a suitable permutation of rows and columns). Then, \(-TK + D\) can be partitioned as
\[
-TK + D = \begin{bmatrix}
T_{11} & T_{1\Pi} \\
T_{\Pi 1} & T_{\Pi \Pi}
\end{bmatrix}
\begin{bmatrix}
K_{11} & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
D_{11} & 0 \\
0 & D_{\Pi \Pi}
\end{bmatrix}
\]
\[
-\begin{bmatrix}
T_{11}K_{11} + D_{11} & 0 \\
-D_{\Pi 1}K_{\Pi 1} & -D_{\Pi \Pi}
\end{bmatrix}
\]
where the subscript \(I, I\) denotes the principal submatrix composed of the first \(m\) rows and columns (the submatrices denoted by the other subscripts are defined similarly). Hence,
\[
\det(-TK + D) = \det D_{\Pi \Pi} \det(-T_{11}K_{11} + D_{11}).
\]
We have \(\det D_{\Pi \Pi} \neq 0\). Furthermore, since \(K_{11} > 0\) and since \(-T + DG^{-1} \in P\) implies \(-T_{11} + D_{11}G_{11} \in P\), we get by the same argument as that in the previous point (2) that \(\det(-T_{11}K_{11} + D_{11}) \neq 0\). Therefore, \(\det(-TK + D) \neq 0\).

If \(-T \in P_0\), then \(-TK \in P_0\) for any \(K \geq 0\) [32]. Hence, \(-TK + D \in P\) and once more \(\det(-TK + D) \neq 0\).

Proof of Proposition 1. Suppose, for purposes of contradiction, that there exist \(x, \bar{x} \in R^n\) with \(x \neq \bar{x}\) such that \(H(x) = H(\bar{x})\). From (2), we get \(-D(x - y) + T(g(x) - g(y)) = 0\), or
\[
(-TK + D)(x - y) = 0
\]
where we have let \(g(x) - g(y) = K(x - y)\), with \(K = \text{diag}(k_1, \ldots, k_n)\). Since \(g \in \{G_m\}\), the mean value theorem yields \(0 \leq k_i \leq G_i\) in the finite sector case and \(k_i \geq 0\) in the infinite sector case. However, from Lemma 2, \(\det(-TK + D) \neq 0\). Hence, from (39), \(x = y\), which is a contradiction. \(\square\)

APPENDIX B

Proof of Lemma 1. Choose a positive number \(k\) such that
\[
k > \frac{1}{\varepsilon} \frac{\|D^{-1}T\|_2^2}{\varepsilon}
\]
where \(\varepsilon = \lambda_m\{[\alpha(-T + DG^{-1})]S\} > 0\). Then, choose \(\varepsilon_0 = 2 \kappa D > 0\).

With these choices, system (11) reduces to
\[
\begin{align*}
P(-D) + (-D)P &= -QQ^T \\
P(-T) &= -Q\Gamma \\
\Gamma^\top\Gamma + \Pi &= 2k[\alpha(-T + DG^{-1})]S
\end{align*}
\]
(41)
Let us verify that we can construct matrices \(P = P^* > 0, \Pi = \Pi^* > 0, Q\) with \(\det Q \neq 0\) and \(\Gamma\) such that (41) is satisfied. Choose \(Q = E_n\) (\(E_n\) is the identity matrix) so that the first equation in (41) yields
\[
P = P^* = \frac{1}{2}D^{-1} > 0
\]
From the second equation in (41), \(\Gamma\) is therefore given by
\[
\Gamma = \frac{1}{2}D^{-1}T
\]
and then, by substituting in the third equation, we get
\[
\Pi = \Pi^* = 2k[\alpha(-T + DG^{-1})]S - \frac{1}{4}(D^{-1}T)^\top D^{-1}T
\]
\[
\geq 2k\mu E_n - \frac{1}{4}(D^{-1}T)^\top D^{-1}T
\]
(42)
since \(\mu = \lambda_m\{[\alpha(-T + DG^{-1})]S\}\). By considering that 
\[
\|D^{-1}T\|_2^2 = \lambda_m\{(D^{-1}T)^\top D^{-1}T\}
\]
we get from (40) and (42)
\[
\Pi \geq (2k\mu - \frac{1}{4}\|D^{-1}T\|_2^2)E_n > 0
\]
\(\square\)

REFERENCES


Mauro Forti received the electronic engineering degree (cum laude) from the University of Florence, Italy, in 1988 with a thesis on electromagnetic compatibility problems in the power distribution system. Since 1991, he has been a researcher on applied mathematics and network theory with the Electronic Engineering Department of the University of Florence. His main research interests are in the field of nonlinear circuits and systems with emphasis on stability analysis of artificial neural networks. His research activity also includes some aspects of applied electromagnetism (mainly electromagnetic compatibility).

Dr. Forti was awarded the 1988 Francini prize and the SIP (Italian Public Telephone Network) prize in telecommunications.

Alberto Tesi (M'92) received the Laurea in electronic engineering from the Universita di Firenze, Italy, in 1984. In 1989, he received the Ph.D. degree from the Universita di Bologna, Italy.

Upon completion of the first degree, he worked as a system engineer at Autostrade s.p.a (Italian Highway Agency) until 1986. In 1990, he joined the Dipartimento di Sistemi e Informatica of the Universita di Firenze, where he is currently an Associate Professor of automatic control. His research interests are mainly in robust control of uncertain linear systems and analysis of nonlinear dynamics of complex systems including chaos and neural networks.