On Existence of Solutions to Structured Lyapunov Inequalities

Aivar Sootla and James Anderson

Abstract—In this paper, we derive sufficient conditions on drift matrices under which block-diagonal solutions to Lyapunov inequalities exist. The motivation for the problem comes from a recently proposed basis pursuit algorithm. In particular, this algorithm can provide approximate solutions to optimisation programmes with constraints involving Lyapunov inequalities using linear or second order cone programming. This algorithm requires an initial feasible point, which we aim to provide in this paper. Our existence conditions are based on the so-called $\mathcal{H}$ matrices. We also establish a link between $\mathcal{H}$ matrices and an application of a small gain theorem to the drift matrix. We finally show how to construct these solutions in some cases without solving the full Lyapunov inequality.

I. INTRODUCTION

Lyapunov equations and matrix inequalities play a central role in control theory, since they are used for, e.g., verifying stability of a dynamical systems, optimal control, and model order reduction (cf. [2]). Lyapunov matrix inequalities with sparsity constraints on the decision variables are used in, in the context of distributed control [3], structured model reduction [4] etc. In such applications, a typical constraint on the decision variables is block-diagonality of a matrix. The major bottleneck in solving optimisation programmes with a Lyapunov inequality constraint is scalability, since it is a semidefinite programme (SDP). There exist a number of methods addressing scalability of SDPs (cf. [5], [6], [7]), and in one of them, it was proposed to replace the constraints in the cone of positive semidefinite matrices with conic inner-approximations [8], [9]. There are two main conic approximations: one which results in a linear programme (LP), and another which results in a second order cone programme (SOCP). Since we are dealing with inner approximations of the cone of positive semidefinite matrices, even if the LP or SOCP solution can be computed, this solution is usually conservative with respect to the optimal SDP solution. This limitation was partially addressed using the basis pursuit algorithm [1], which is iterates over LPs or SOCPs and provides a guarantee of improvement with each iteration. This algorithm requires an initial feasible point in order to start the iterations. Hence, major questions still remain concerning existence theorems and scalable computation of block-diagonal solutions to Lyapunov inequalities.

Necessary and sufficient conditions for block-diagonal stability were described more than 20 years ago in [10]. However, these results do not provide a constructive way to build block-diagonal Lyapunov functions. This perhaps explains why these results are relatively unused in the control theory literature. Besides some simple cases, such as, the drift matrix being block-triangular matrix (cf. [11]), it is known that the closed loop interconnection of strictly passive systems has a drift matrix which admits a block-diagonal solution to Lyapunov inequalities [12]. It is also well-known that stable Metzler matrices admit diagonal solutions to the Lyapunov inequality [13]. Additional special cases are covered in [14], [15] and revisited in what follows.

In this paper, we aim at identifying additional cases, when a block-diagonal solution to a Lyapunov inequality can be found using algebraic methods or LPs. We start by studying a generalisation of Metzler matrices known as $\mathcal{H}$ matrices. Stable $\mathcal{H}$ matrices possess many properties of stable Metzler matrices, for example they also admit diagonal solutions to Lyapunov inequalities [16]. We provide another such property, namely we show that for $\mathcal{H}$ matrices, diagonal solutions to Lyapunov inequalities can be computed using algebraic methods and/or LPs. We then investigate conditions on specific blocks in block-partitioned matrices. We establish a link between the $\mathcal{H}$ matrix conditions and a version of the small gain theorem before extending this intuition to block-partitioned case. In the 2 by 2 block partitioned case, we provide an explicit way to construct block-diagonal solutions to the Lyapunov inequalities without the need to solve the full inequality. An extension to $n$ by $n$ block partitioned case is one of the future work directions.

The rest of the paper is organised as follows. In Section II we cover some preliminaries and motivate our problem formulation in Section III. We show how to construct diagonal solutions to Lyapunov inequalities for $\mathcal{H}$ drift matrices in Section IV. We provide stability results for block partitioned matrices and link the condition for $\mathcal{H}$ matrices with the small gain theorem in Section V. In Section VI we provide a large-scale numerical example and we conclude in Section VII where we discuss linear programming solutions to Lyapunov inequalities.

Notation: Our notation is mostly standard: $\rho(A)$ stands for the spectral radius of a matrix $A$, $A \geq 0$ (respectively,
A \gg 0) means that all entries \(a_{ij}\) of \(A\) are nonnegative (respectively, positive), \(A \succeq 0\) (respectively, \(A \succ 0\)) means that \(A\) is positive semidefinite (respectively, positive definite). Let \(S^n\) denote the set of symmetric \(n\) by \(n\) matrices, \(S^n_+\) denotes the cone of positive semidefinite \(n\) by \(n\) matrices. Let \(\|B\|_2\) be the matrix induced norm, that is \(\|B\|_2\) is equal to the maximum singular value of \(B\), and let \(\sigma(B)\) denote the minimum singular value of \(B\). The \(\|\cdot\|_\infty\) norm of a transfer matrix \(G\) is defined as \(\|G\|_\infty = \max_{s \in \mathbb{C} : \text{Re}(s) \geq 0} \|G(s)\|_2\) and \(\|H\|_\infty = \max_{s \in \mathbb{C}} \|H(s)\|_2\) for stable \(G\). For a space \(X\), its dual is denoted as \(X^*\). Finally, \(I\) is the identity matrix of an appropriate dimension.

II. PRELIMINARIES

Consider the linear time invariant dynamical system
\[
\dot{x}(t) = Ax(t), \quad x(0) = x_0
\]
where \(x(t) \in \mathbb{R}^n\). An important concept associated with the system (1) is stability, which is typically verified by solving a linear matrix inequality (LMI).

**Proposition 1:** System (1) is stable if and only if there exists an \(X \succ 0\) that satisfies the LMI
\[
AX + XA^T \prec 0.
\]

A matrix \(X\) which satisfies (2) defines a Lyapunov function of the form \(V(x) = x^T(t)X^{-1}x(t)\) for system (1). In this paper, we aim at describing some sufficient conditions of solvability of the LMI (2) when the decision variable \(X\) satisfies additional sparsity constraints.

In order to simplify the presentation we say that a matrix \(A \in \mathbb{R}^{N \times N}\) has \(\alpha = \{k_1, \ldots, k_n\}\)-partitioning with \(N = \sum_{i=1}^n k_i\), if the matrix \(A\) is written as follows
\[
A = \begin{pmatrix}
A_{11} & A_{12} & \ldots & A_{1n} \\
A_{21} & A_{22} & \ldots & A_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} & A_{n2} & \ldots & A_{nn}
\end{pmatrix}
\]
where \(A_{ij} \in \mathbb{R}^{k_i \times k_j}\). We say that \(A\) is \(\alpha\)-diagonal if it is \(\alpha\)-partitioned and \(A_{ij} = 0\) if \(i \neq j\), and \(\alpha\)-lower triangular if \(A_{ij} = 0\) if \(i < j\). We aim at characterising \(\alpha\)-diagonally stable matrices \(A \in \mathbb{R}^{N \times N}\), which are such that there exists an \(\alpha\)-diagonal positive definite \(X \in \mathbb{R}^{N \times N}\) satisfying (2). If \(\alpha = \{1, \ldots, 1\}\), we say that an \(\alpha\)-diagonal (respectively, \(\alpha\)-lower triangular, \(\alpha\)-diagonally stable) matrix \(A\) is \(\alpha\)-diagonal (respectively, lower-triangular, diagonally stable).

We will make use of so-called scaled diagonally dominant matrices.

**Definition 1:** A matrix \(A \in \mathbb{R}^{n \times n}\) is called strictly row scaled diagonally dominant if there exist positive scalars \(d_1, \ldots, d_n\) such that
\[
d_i |a_{ii}| > \sum_{j \neq i} d_j |a_{ij}|
\]
for all \(i = 1, \ldots, n\). The matrix \(A\) is strictly row diagonally dominant if \(d_i = 1\) for all \(i\).

A related class to scaled diagonally dominant matrices is the class of \(\mathcal{H}\) matrices. In order to define this class we require the following definitions:

**Definition 2 ([17]):** Given an \(\alpha\)-partitioned matrix \(A\) with nonsingular \(A_{ii}\) for all \(i\), we define the \(\alpha\)-comparison matrix \(\mathcal{M}_\alpha(A)\) as
\[
\mathcal{M}_\alpha(A) = \begin{cases}
\|A_{ii}^{-1}\|_2^{-1} & \text{if } i = j, \\
-\|A_{ij}\|_2 & \text{otherwise},
\end{cases}
\]

When \(\alpha = \{1, \ldots, 1\}\), we will simply write \(\mathcal{M}(A)\).

Note that \(\|A_{ii}^{-1}\|_2^{-1} = \sigma(A_{ii})\). Hence, using a continuity argument we can assume that \(\|A_{ii}^{-1}\|_2^{-1} = 0\) for a singular \(A_{ii}\), and Definition 2 is well-posed. The \(\alpha\)-partitioned matrices allow a version of Gershgorin circle theorem:

**Proposition 2 ([18]):** For an \(\alpha\)-partitioned matrix \(A \in \mathbb{R}^{N \times N}\), where \(\alpha = \{k_1, \ldots, k_n\}\) and \(N = \sum_{i=1}^n k_i\), every eigenvalue of \(A\) satisfies
\[
\|(A - A_{ii})^{-1}\|_2^{-1} \leq \sum_{j=1,j\neq i}^n \|A_{ij}\|_2
\]
for at least one \(i\) where \(i = 1, \ldots, n\).

**Definition 3:** A matrix \(A \in \mathbb{R}^{n \times n}\) is said to be Metzler if all the off-diagonal elements are positive.

**Definition 4:** A matrix \(A \in \mathbb{R}^{n \times n}\) is said to be an \(\mathcal{H}\) matrix, if the minimal real part of the eigenvalues of \(\mathcal{M}(A)\) is greater than or equal to zero.

It is clear that stable Metzler matrices are also \(\mathcal{H}\) matrices. It is also straightforward to show that \(A\) is strictly row and column scaled diagonally dominant if and only if \(\mathcal{M}(A)\) has eigenvalues with positive real part [19]. We also refer the reader to [20], [16] for additional information on \(\mathcal{H}\) matrices.

Let \(\mathcal{DD}_+\) denote the cone of matrices \(A\) such that \(A\) and \(A^T\) are strictly diagonally dominant and the elements on the diagonal of \(A\) are positive (that is, \(A_{ii} > 0\)). Similarly, let \(\mathcal{H}_+\) denote \(\mathcal{H}\) matrices \(A\) with positive elements on the diagonal of \(A\). If \(A\) is a symmetric \(\mathcal{DD}_+\) matrix, then by Proposition 2 with \(\alpha = \{1, \ldots, 1\}\), it is easy to show that \(A \succ 0\). Moreover, the constraint that \(A = A^T \in \mathcal{DD}_+\) can be written as a set of linear constraints
\[
a_{ii} > \sum_{j \neq i} c_{ij} \quad \forall i,
\]
\[-c_{ij} \leq a_{ij} \leq c_{ij}, \quad c_{ij} = c_{ji} \quad \forall i \neq j.
\]
Hence, a constraint \(A \succ 0\) can be replaced by a more restrictive but scalable linear constraints. This approach was proposed in [8], [9] to restrict some sum-of-squares optimisation problems which are naturally SDPs to LPs.

A symmetric \(\mathcal{H}_+\) matrix is also positive semidefinite, and this constraint can be imposed by a number of second order cone constraints [21]. That is \(A = A^T \in \mathcal{H}_+\) if and only if
\[
A = \sum_{i=1}^N E_i^T X_i E_i, \quad \text{with } X \in S_+^n, \quad E_i \in \mathcal{T}_2
\]
where \(E \in \mathcal{T}_2\), if \(E \in \mathbb{R}^{n \times 2}\) and every column of \(E\) has only one non-zero element equal to one, and \(N = |\mathcal{T}_2|\).
To summarise the subsection, we will mention this strong result on diagonal stability of $\mathcal{H}_+$ matrices, which we will revisit in the sequel.

**Proposition 3 ([16]):** Let $-A$ be an $\mathcal{H}_+$ matrix. Then $A$ is diagonally stable if and only if $A$ is nonsingular.

### III. Motivation and Problem Formulation

#### A. Conic Programming

Conic optimisation problems take the generic form of optimising a linear functional over the intersection of an affine subspace and a proper cone. Typically conic programmes have the following primal and dual formulations:

$$\begin{align*}
\text{minimise} & \quad c^T x \\
\text{maximise} & \quad b^T y \\
\text{s.t.} & \quad Ax = b \\
& \quad c - A^T y = s \\
& \quad x \in \mathcal{K} \\
& \quad (y, s) \in (\mathbb{R}^m, \mathcal{K}^*)
\end{align*}$$

where $\mathcal{K}$ is a proper cone (i.e. closed, non-empty, pointed, convex) and $\mathcal{K}^*$ is the dual cone of $\mathcal{K}$ defined as

$$\mathcal{K}^* := \{ y \mid \langle y, x \rangle \geq 0, \forall x \in \mathcal{K} \}.$$ 

It is well known, that the cone of positive semidefinite matrices $\mathcal{S}_+^n$ is self dual meaning that $(\mathcal{S}_+^n)^* = \mathcal{S}_+^n$. The cone of symmetric $\mathcal{D} \mathcal{D}_+$ matrices

$$\mathcal{K}_{\text{LP}} = \{ X \in \mathcal{S}_+^n \cap \mathcal{D} \mathcal{D}_+ \},$$

however, is not self-dual and it is larger than the cone $\mathcal{S}_+^n$. More specifically:

$$\mathcal{K}_{\text{LP}}^* = \{ X \in \mathcal{S}_+^n \mid v_i^T X v_i \geq 0, \forall v_i \in \mathcal{T}_i \},$$

where $\mathcal{T}_i$ is the set of all vectors in $\mathbb{R}^n$ with a maximum of two non-zero elements, each of which is $\pm 1$. The cone of symmetric $\mathcal{H}_+$ matrices defined as

$$\mathcal{K}_{\text{SOCP}} = \{ X \in \mathcal{S}_+^n \cap \mathcal{H}_+ \},$$

has the dual

$$\mathcal{K}_{\text{SOCP}}^* = \{ X \in \mathcal{S}_+^n \mid E_i^T X E_i \geq 0, \forall E_i \in \mathcal{T}_2 \}.$$ 

We will make use of these cones and their duals in the remainder of the paper.

#### B. Structured Gramians via Basis Pursuit

The standard from primal SDP [22] is written as

$$\begin{align*}
\min \quad & \langle C, X \rangle \\
\text{s.t.} & \quad X \in \mathcal{S}_+^n, \quad \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, m
\end{align*}$$

where $\mathcal{S}_+^n$ is the cone of $n \times n$ positive semidefinite matrices. The basis pursuit algorithm proceeds as follows: At each iteration the algorithm re-parameterizes the simpler cone that approximates $\mathcal{S}_+^n$ and then solves an optimization problem over this cone, the solution of which is then used to update the cone for the next iteration. In particular, the algorithm specifies for a fixed matrix $L$, the cone

$$\mathcal{K}(L) = \{ X \mid X = L^T Q L, \quad Q = Q^T \in \mathcal{D} \mathcal{D}_+ \}.$$ 

Note that $Z \in \mathcal{K}(L) \Rightarrow Z \succeq 0$. The algorithm in [1] solves a sequence of optimization problems of the form (5) but with the conic constraint replaced by $X \in \mathcal{K}(L_k)$ where the sequence $\{L_k\}$ is given by

$$L_0 = I$$

$$L_{k+1} = \text{decomp}(X_k)$$

where $\text{decomp}(X_k)$ is a Cholesky decomposition of $X_k$, the optimal solution decision variable from iteration $k$. In some cases, $X_k$ can have singular values close to zero, thus creating numerical problems in the iterative procedure. In order to avoid such cases, we can remove a $k$-th column of $L$ with the $k$-th entry close to zero (we assume that $L$ is lower triangular). One can also use LDL decomposition to avoid dealing with negative eigenvalues of $X_k$. This approach also slightly improves numerical complexity of the conic programme by lowering the number of decision variables and constraints. We finally note that the method relies on the fact that at the first iteration a feasible solution $X_0$ exists.

In many applications, it is desirable to solve the following problem using an LP or SOCP rather than the more natural SDP:

$$\begin{align*}
\min \quad & \text{trace}(X) \\
\text{such that:} & \quad AX + X A^T < 0 \\
& \quad X = X^T \text{ is } \alpha\text{-diagonal},
\end{align*}$$

where $A$ is Hurwitz, and $\alpha$ is a given partitioning. Note that since $A$ is Hurwitz, then the condition $X > 0$ is implied by the solvability of (6). The basis pursuit can be applied given an $X$ satisfying the constraints of the programme (6).

Therefore, we set up our problem: find $X$ satisfying the constraints of (6) with algebraic or linear programming methods.

There are many practical applications, where the problem of the form (6) appears and one of them is structured model reduction. For example, consider the boiler-header system depicted in Figure 1. The state space can be partitioned according to dimensions of the subsystems, which are $\{3, 3, 1\}$. The system always admits diagonal generalised Gramians, since the drift matrix of the closed loop system is a stable $\mathcal{H}_+$ matrix.

In order to perform structured model reduction [12] the authors computed a $\{3, 3, 1\}$-diagonal generalised controllability Gramian $P = \text{diag}\{P_1, P_2, P_3\}$ such that $P_1, P_2 \in \mathbb{R}^{3 \times 3}$ and $P_3 \in \mathbb{R}$. The optimal trace of such a Gramian computed using semidefinite programming is equal to $1.2817 \times 10^4$. The figure shows the block-diagram of the boiler-header system.
Using linear programming, we found the minimum trace of $2.4132 \cdot 10^4$ signifying a loss of quality of almost 100%. After solving one iteration of the basis pursuit algorithm we obtain objective equal to $1.5893 \cdot 10^4$, an additional iteration of the algorithm gives $1.3172 \cdot 10^4$, and one more provides a value equal to $1.3093 \cdot 10^4$, which comes really close to the optimal value. Naturally, on this example we do not need basis pursuit or linear programming to obtain an optimal solution due to the low complexity of the problem. However, this example indicates that the basis pursuit algorithm can be beneficial to obtain an approximate solution of large scale Lyapunov inequalities using linear programmes.

IV. $\mathcal{H}$ Matrices and Diagonal Stability

The main result of this section concerns diagonal stability of $\mathcal{H}$ matrices, where we sharpen the results from [16] by providing an explicit diagonal Lyapunov function for a class of $\mathcal{H}_+$ matrices.

**Theorem 1:** Let $-A$ be an $\mathcal{H}_+$ matrix with a nonsingular $M(A)$. Then the following conditions hold

1. There exist positive vectors $v = (v_1 \ldots v_n)^T$, $w = (w_1 \ldots w_n)^T$ such that $M(A) v, w^T M(A)$ are also positive.
2. There exists a diagonal $X$ such that $-(AX + XAT)$ is an $\mathcal{H}_+$ matrix. Moreover, we can choose it as $X = P_w P_w^{-1}$, where $P_v = \text{diag}(v_1, \ldots, v_n)$, $P_w = \text{diag}(w_1, \ldots, w_n)$, and $v, w$ satisfy point 1.
3. There exists a diagonal positive definite matrix $Y$ such that

$$-P_w A P_w^{-1} Y - Y P_w^{-1} A^T P_w \in \mathcal{D} \mathcal{D}_+$$

**Proof:**

1. By definition $-M(A)$ is a Metzler matrix with all eigenvalues $\lambda_i (M(A)) \leq 0$, since $M(A)$ is nonsingular by the premise, $-M(A)$ is a Hurwitz Metzler matrix. Hence the claim follows by applying the results from [23].
2. Let $X = P_w P_w^{-1}$, then

$$(M(A) X + X M(A^T)) w = (M(A) v + X M(A)^T) w \gg 0,$$

where the inequality follows since $M(A) v$ and $M(A)^T w$ are positive and $X$ is nonnegative. Hence $S = -M(A) X - X M(A^T)$ is a Metzler matrix and there exists a positive vector such that $S w$ is negative. This implies that $S$ is a symmetric Hurwitz and Metzler matrix, which means that $M(A) X + X M(A^T)$ is positive definite.

Note that $a_{ii} < 0$ for all $i$, let

$$(-AX - X A^T)_{ij} = -a_{ij} x_j - a_{ji} x_i$$

$$(M(A) X + X M(A^T))_{ij} = \begin{cases} -a_{ij} x_j - a_{ji} x_i & i = j \\ -|a_{ij}| x_j - |a_{ji}| x_i & i \neq j \end{cases}$$

It is straightforward to show that $M(A) X + X M(A^T) \leq M(-AX - X A^T)$, moreover the elements on the diagonal are equal. This means that we can write $M(A) X + X M(A^T) = s I - R_1$, $M(-AX - X A^T) = s I - R_2$, where the matrices $R_1$ and $R_2$ satisfy $R_1 \geq R_2 \geq 0$. According to Weilandt’s theorem $\rho(R_1) \geq \rho(R_2)$ (cf. [16]). Therefore the minimal eigenvalue of $M(A) X + X M(A^T)$ is smaller or equal to the minimal eigenvalue of $M(-AX - X A^T)$. This implies that $M(-AX - X A^T)$ has eigenvalues with positive real part, hence $-AX - X A^T$ is an $\mathcal{H}_+$ matrix.

3. Consider the matrix $R = P_w M(A) P_v$, and $\epsilon$ the vector of ones. Now it is easy to see that $R \epsilon \gg 0$:

$$P_w M(A) P_v \epsilon = P_w M(A) v \gg 0.$$  

This implies that the matrix $P_w M(A) P_v$ is row strictly diagonal dominant. Similarly, we can show that $P_w M(A) P_v$ is column strictly diagonal dominant. This by definition implies that the matrix $-P_w M(A) P_v$ is a row and column diagonal dominant matrix with positive elements on the diagonal or a $\mathcal{D} \mathcal{D}_+$ matrix. Hence the matrix $-P_w A P_v - P_v A^T P_w$ is positive definite. Since we can set $Y = P_v P_w$, the result follows.

We showed that there exists a diagonal $X$ matrix such that the matrix $Z = -P_w A P_w^{-1} X - X P_w^{-1} A^T P_w$ is a $\mathcal{D} \mathcal{D}_+$ matrix and hence positive definite. Note that the constraint $Z = Z^T \in \mathcal{D} \mathcal{D}_+$ is linear and if needed we can relax the sparsity constraints on $X$. This implies that given an $\mathcal{H}$ drift matrix, we can compute an $\alpha$-diagonal Lyapunov function with an arbitrary $\alpha$ using linear programming.

If the entries of the $A$ matrix are poorly scaled then solving a linear programme can be numerically challenging. Using our methods, this can be avoided if we compute an initial point using the right and left eigenvectors of $M(A)$, instead of the positive vectors $v$ and $w$ satisfying point 1). Having an initial point re-scales the optimisation programme and can provide feasible points as shown on a specific example in [24].

**Theorem 1** is a direct generalisation of the similar result for Metzer matrices (cf. [23]), but our result can be applied to a broader class of matrices including lower-triangular matrices. Using Theorem 1 another results for Metzer matrices can be extended to problems such as construction of sum- and max-separable Lyapunov functions (cf. [23]).

The state-space transformation $P_w$ is essential in order to guarantee the diagonal dominance of the inequality. Consider an asymptotically stable matrix

$$A = \begin{pmatrix} -1 & -2 \\ 2 & -5 \end{pmatrix}.$$

and a positive definite $X = \text{diag}(x_1, x_2)$. The matrix $-A$ is an $\mathcal{H}_+$ matrix and it is stable. The diagonal dominance of $AX + X A^T$ requires the following inequalities to be fulfilled

$$2 \cdot x_1 > 2x_2 + 2x_1, \quad 2 \cdot 5x_2 > 2x_2 + 2x_1$$

for some positive $x_1, x_2$. The first inequality is equivalent to $0 > 2x_2$, which is impossible to fulfil.

V. $\alpha$-Diagonal Stability and $\mathcal{H}_+$ Matrices

A. A Motivating Example

In this section, we cover two main classes of results for diagonal stability and compare them to a classical example from [14] for cyclic systems. These classes stem from two arguments based on the passivity and the small gain theorem. In this section, we will argue that the $\mathcal{H}_+$ matrix condition
is an implicit constraint in these stability proofs. In order to explain our motivation consider an example studied in [14] and let:

\[
A_n^0 = \begin{pmatrix}
\beta_1 & \cdots & \beta_n \\
0_{n-1 \times 1} & \cdots & 0_{n-1 \times 1}
\end{pmatrix}
\]

where \(\beta_1, \ldots, \beta_n\) are positive scalars. This matrix represents the dynamics of a negative feedback of a cascade of transfer functions \(G_i(s) = \frac{\beta_i}{s + \alpha_i}\). First, let us consider the 2 by 2 case, which gives

\[
A_2^0 = \begin{pmatrix}
-\alpha_1 & -\beta_1 \\
-\beta_2 & -\alpha_2
\end{pmatrix}
\]

and two transfer functions \(G_1 = \frac{\beta_1}{s + \alpha_1}\) and \(G_2 = \frac{\beta_2}{s + \alpha_2}\). According to the small gain theorem, the system is stable if

\[
\left\| \frac{\beta_1}{s + \alpha_1} \right\|_{\mathcal{H}_\infty} \cdot \left\| \frac{\beta_2}{s + \alpha_2} \right\|_{\mathcal{H}_\infty} = \frac{\beta_1 \beta_2}{\alpha_1 \alpha_2} < 1.
\]

This argument can be extended to an arbitrary size matrix resulting in the condition

\[
\left(\frac{\beta_1 \cdots \beta_n}{\alpha_1 \cdots \alpha_n}\right) < 1. \quad (8)
\]

Surprisingly, it is straightforward to verify by definition that \(A_n^0\) is an \(\mathcal{H}_+\) matrix if and only if (8) holds. Hence on this loop the \(\mathcal{H}_+\) matrix condition is a small gain condition.

Alternatively, using passivity arguments it was shown in [14], that \(A_n^0\) is asymptotically stable if and only if

\[
\left(\frac{\beta_1 \cdots \beta_n}{\alpha_1 \cdots \alpha_n}\right) < (\sec(\pi/n))^n. \quad (9)
\]

This in particular means that for \(n = 2\) all matrices in the form \(A_2^0\) are not only stable, but also diagonally stable, however, they may not be \(\mathcal{H}_+\) matrices. This analysis is based on passivity arguments and has been extended to less restrictive classes of systems in [15]. It is easy to verify that with \(n \to \infty\) the limit \((\sec(\pi/n))^n\) converges to one. Hence, it appears (for this class of system) that for large dimensions, \(\mathcal{H}_+\) matrices constitute a large subset of diagonally stable matrices. We will pursue the relation between \(\mathcal{H}_+\) matrices and small gain argument in the \(\alpha\)-diagonal case in the remainder of the paper.

**B. Passivity and Small Gain Conditions for \(\alpha\)-Diagonal Stability**

Let \(G^c\) be the closed loop transfer function depicted in Figure 2, which is an interconnection of two Linear Time Invariant (LTI) subsystems

\[
G_i = \begin{bmatrix}
A_i & B_i \\
C_i & D_i
\end{bmatrix}
\]

where \(A_i \in \mathbb{R}^{k_i \times k_i}, B_i \in \mathbb{R}^{k_i \times m_i}, C_i \in \mathbb{R}^{l_i \times k_i}, D_i \in \mathbb{R}^{l_i \times m_i}\) and \(m_1 = l_2, m_2 = l_1\). The closed loop transfer function from \([u_1, u_2]\) to \([y_1, y_2]\) has the following state-space realisation

\[
G^c = \begin{bmatrix}
A_{11}^c & A_{12}^c & B_{11}^c & B_{12}^c \\
A_{21}^c & A_{22}^c & B_{21}^c & B_{22}^c \\
C_{11}^c & C_{12}^c & D_{11}^c & D_{12}^c \\
C_{21}^c & C_{22}^c & D_{21}^c & D_{22}^c
\end{bmatrix},
\]

where

\[
A_{11}^c = A_1 - B_1 R_2 D_2 C_1, \quad A_{12}^c = -B_1 R_2 C_2 \quad (10)
\]

\[
A_{21}^c = B_2 R_1 D_1 C_2, \quad A_{22}^c = A_2 - B_2 R_1 D_1 C_2
\]

and \(R_{12} = (I + D_1 D_2)^{-1}, R_{21} = (I + D_2 D_1)^{-1}\) and the rest of the matrices are computed accordingly. For the sake of simplicity we assume that this realisation is minimal. Passivity and small gain arguments can both be used to determine if the closed loop system is stable but we will focus on the small gain condition. Passivity results in this direction will be addressed in future work, similar ideas were pursued in [25], [26].

It is straightforward to verify that stability of the system with inputs \(u_1, u_2\) and outputs \(y_1, y_2\) depends on stability of the transfer function \(L = (I - G_2 G_1)^{-1}\).

**Proposition 4 (Small Gain Theorem):** Suppose \(B\) is a Banach-algebra and \(Q \in B\). If \(|Q| < 1\), then \((I - Q)^{-1}\) exists and

\[
(I - Q)^{-1} = \sum_{k=0}^{\infty} Q^k.
\]

Applying Proposition 4 we can verify that if \(|Q| := \|G_2 G_1\|_{\mathcal{H}_\infty} \leq \|G_2\|_{\mathcal{H}_\infty} \|G_1\|_{\mathcal{H}_\infty} < 1\) then the function \(L\) and hence the closed loop are stable (cf. [27]). We can apply the small gain condition to the closed transfer function, which would result in a condition on \(\alpha\)-diagonal stability of the matrix \(A^c\). However, given only a partitioning \(\alpha = \{k_1, k_2\}\) and a realisation of the closed loop transfer function \(G^c\), these conditions again would be hard to verify. We can apply a small gain theorem in another way, namely apply it to the matrix \(A^\alpha\) directly. In this case, we do not need to know the realisation of transfer functions \(G_1\) and \(G_2\), all we need to know is the matrix \(A^\alpha\) and the partitioning \(\alpha\).

The conditions on \(\alpha\)-diagonal stability of \(A^\alpha\) are established in the following proposition.

**Proposition 5:** Let \(A^\alpha\) be \(\alpha\) partitioned with \(\alpha = \{k_1, k_2\}\)

\[
A^\alpha = \begin{pmatrix}
A_{11}^\alpha & A_{12}^\alpha \\
A_{21}^\alpha & A_{22}^\alpha
\end{pmatrix}.
\]

Let \(K_1(s) = -(sI - A_{11}^\alpha)^{-1} A_{12}^\alpha, K_2(s) = (sI - A_{22}^\alpha)^{-1} A_{21}^\alpha\) with Hurwitz \(A_{11}^\alpha, A_{21}^\alpha\). If there exists a \(\gamma > 0\) such that

\[
\|K_1\|_{\mathcal{H}_\infty} < 1/\gamma \quad \text{and} \quad \|K_2\|_{\mathcal{H}_\infty} < \gamma,
\]

then the matrix \(A^\alpha\) is \(\alpha\)-diagonally stable.

**Proof:** We need to show that there exists an \(\alpha\)-diagonal Lyapunov function for the system \(x^c = A^\alpha x^c\). For the sake of clarity we drop the superscript \(c\) from \(A^\alpha_{ij}\) and simply write \(A_{ij}\). The inequality \(|K_1|_{\mathcal{H}_\infty} < 1/\gamma\) and the Bounded
Real Lemma imply that $X_1 > 0$ solves the Riccati equation

$$0 = A_{11}X_1 + X_1A_{11}^T + \gamma^2X_1X_1 + A_{12}A_{21} = (11)$$

$$A_{11}X_1 + X_1A_{11}^T + (X_1A_{12}) (\begin{array}{cc} \gamma^2 I_{k_1} & 0 \\ 0 & I_{k_2} \end{array}) (X_1A_{12}^T),$$

which has always has a solution since $(I, A_{11})$ is a controllable pair (cf. [2]), since the control matrix is equal to $I$ and we can control every state independently.

Again, due to the Bounded Real Lemma the inequality $\|K_2\|_{\infty} < \mu$ is equivalent to

$$A_{22}Y_2 + Y_2A_{22}^T + \mu^2 Y_2 = A_{21}A_{21} = 0,$$  

where $Y_2 > 0$ since $(I, A_{22})$ is a controllable pair (cf. [2]). Let $\mu = \gamma - \epsilon$ for some $\epsilon > 0$ such that $\mu > \|K_2\|_{\infty}$, which implies that $\mu^2 Y_2 > \gamma^2 Y_2$ and consequently:

$$A_{22}Y_2 + Y_2A_{22}^T + \gamma^2 Y_2 + A_{21}A_{21} < 0.$$  

By multiplying the equation by $\gamma^{-2}$ setting $X_2 = Y_2^{-1}$

$$A_{22}X_2 + X_2A_{22}^T + X_2X_2$$

$$\gamma^{-2} A_{21}A_{21}^T < 0 \iff (\begin{array}{cc} \gamma^2 I_{k_1} & 0 \\ 0 & I_{k_2} \end{array}) \left(A_{21}^T X_2 + X_2A_{22} \right)^{-1} (A_{21} X_2) > 0$$

Combining the inequalities (11) and (13) yields

$$0 > A_{11}X_1 + X_1A_{11}^T - (X_1A_{12}) (A_{21}^T X_2$$

$$\cdot (A_{22}X_2 + X_2A_{22}^T) (A_{21} X_2) = A_{11}X_1 + X_1A_{12}$$

$$- (X_1A_{21} + A_{12}X_2) (X_2A_{22} + A_{22}^T X_{21}) (A_{21} X_2)^{-1} (A_{21} X_2) > 0.$$  

Applying the Schur complement properties to (14) yields

$$\begin{pmatrix} A_{11}X_1 + X_1A_{11}^T & A_{12}X_2 + X_1A_{12} \\ A_{21}X_1 + X_2A_{21}^T & A_{22}X_2 + X_2A_{22}^T \end{pmatrix} < 0,$$

thus the blocks on the diagonal are negative definite which completes the proof.

Our proof is constructive, and shows how to build an $\alpha$-diagonal Lyapunov function by solving two Riccati equations (11) and (12) instead of solving an LMI. We link a simplified version of these conditions with $\alpha$-partitioned and $\mathcal{H}_+$ matrices.

C. Conditions for $\alpha$-Diagonal Stability via $\mathcal{H}_+$ Matrices

The authors in [18] showed that $A$ is Hurwitz if it is an $\alpha$-partitioned matrix with matrix $\mathcal{M}_\alpha(A)$ being $\mathcal{H}_+$ matrix. In particular, this result shows that $A$ is implied by stability of all the blocks $A_{ii}$. We provide a generalisation of this result.

Lemma 1: Let $A$ be $\alpha$-partitioned matrix and $\mathcal{M}_\alpha(A)$ be an $\mathcal{H}_+$ matrix. Let also $A_{ii}$ be Hurwitz matrices, and the Hamiltonian matrices

$$H_i = \begin{pmatrix} A_{ii} & \gamma_i^{-1}I \\ -I & -A_{ii}^{-1} \end{pmatrix}$$

have no purely imaginary eigenvalues with $\gamma_i = \|A_{ii}^{-1}\|_2 + \epsilon$ for all $\epsilon > 0$. Then $A$ is a Hurwitz matrix.

Proof: We prove the result by contradiction. Let $A$ have eigenvalues with a positive real part. Since $\mathcal{M}_\alpha(A)$ is an $\mathcal{H}_+$ matrix, there exists positive scalars $d_i$ such that for every $i$

$$\|A_{ii}^{-1}\|_2 > \sum_{i \neq j} \|A_{ij}\|_2 d_j \frac{d_j}{d_i}. (16)$$

The matrix $A$ is unstable if and only if $D^{-1}AD$ is unstable with $D = \text{diag}\{d_1, d_2, \ldots, d_n\}$. Let $\lambda$ be the eigenvalue of $D^{-1}AD$ with a positive real part. By Proposition 2 there exists an index $i$ such that

$$\|(\lambda I - A_{ii})^{-1}\|_2^{-1} \leq \sum_{i \neq j} \|A_{ij}\|_2 \frac{d_j}{d_i} = \sum_{i \neq j} \|A_{ij}\|_2 d_j \frac{d_j}{d_i}. (17)$$

Now since the Hamiltonian matrix $H_i$ has no purely imaginary eigenvalues for all $\epsilon > 0$ and $A_{ii}$ is Hurwitz, this implies that $\|(\lambda I - A_{ii})^{-1}\|_2 = |\|A_{ii}^{-1}\|_2|$. Therefore the maximum of $\|(zI - A_{ii})^{-1}\|_2$ over $z$ with $\text{Re}(z) \geq 0$ is equal to $\|A_{ii}^{-1}\|_2$, and $\|(\lambda I - A_{ii})^{-1}\|_2 \leq |\|A_{ii}^{-1}\|_2|$. Hence due to (16)

$$\|(\lambda I - A_{ii})^{-1}\|_2^{-1} \geq \|A_{ii}^{-1}\|_2 > \sum_{i \neq j} \|A_{ij}\|_2 d_j \frac{d_j}{d_i}.$$

We arrive at the contradiction with (17), which completes the proof.

Lemma 1 allows us to determine stability of $A$ by verifying $\alpha$-partition of the blocks $A_{ii}$ subject to the condition (15) and $\mathcal{M}_\alpha(A)$ being a $\mathcal{H}_+$ matrix. This, however, does not directly imply that there exists an $\alpha$-diagonal Lyapunov function. In what follows, we only present the result for $\alpha = \{k_1, k_2\}$ partitioning.

Theorem 2: Let $A$ be $\alpha$ partitioned with $\alpha = \{k_1, \ldots, k_2\}$, then under the premise of Lemma 1 the matrix $A$ is $\alpha$-diagonally stable.

Proof: The proof is using the small gain argument for the systems $G_1(s) = (sI - A_{ii})^{-1}A_{12}$, $G_2(s) = (sI - A_{ii})^{-1}A_{21}$. We have that $\|G_1\|_{\infty} < |\|G_2\|_{\infty}|$, $\|G_1\|_{\infty} \leq \Delta$ where

$$\Delta := \|A_{21}\|_2 (\|sI - A_{i1}\|_2) \|A_{12}\|_2 \|sI - A_{22}\|_2^{-1} \|A_{i2}\|_2 < \gamma.$$  

Under the premise of Lemma 1 we have that $\|A_{21}\|_2 < |\|A_{ii}\|_2|^{-1}$ and $\gamma^{-1} A_{21} < |\|A_{ii}\|_2|^{-1}$. Hence $\|G_1\|_{\infty} < \gamma^{-1}$, while $\|G_2\|_{\infty} < \gamma$. Proposition 5 proves the claim.

Note that if $A$ is such that $\mathcal{M}_\alpha(A)$, $\mathcal{M}_\alpha(A^T)$ is in $\mathcal{D}_D$, it is not generally true that $\mathcal{M}_\alpha(A^T) = \mathcal{D}_D$. This property holds for $\alpha = \{1, \ldots, 1\}$ and was used in the proof of Theorem 1. Hence the absence of this property for a general $\alpha$ is the major obstacle for extending Theorem 1 to the $\alpha$-diagonal case.

VI. NUMERICAL EXAMPLE

Consider the one-dimensional heat equation in the form

$$\frac{\partial T(x, t)}{\partial t} = \alpha \frac{\partial^2 T(x, t)}{\partial x^2} + u(x, t) \quad x \in (0, 1), t > 0$$

$$T(0, t) = T(1, t) = 0, \quad t \geq 0$$

$$T(x, 0) = 0, \quad x \in [0, 1]$$
with $\alpha = -0.01$, where $T(t, x)$ denotes the temperature at time $t$ at $x$. Assume, we want to heat (i.e. apply an input) at a point of the rod located at $1/3$ of its length, and observe the temperature at a point on the rod located at $2/3$ of its length. Then as in [28], we can obtain the following spatially discretised model:

$$
\begin{align*}
X'(t) &= AX(t) + Bu(t),
Y(t) &= CX(t),
\end{align*}
$$

where $X(t) \in \mathbb{R}^n$ is the temperature at time $t$ at each of the $n$ spatial discretisation points, and

$$
A = \alpha(n + 1)^2 \begin{pmatrix}
2 & -1 \\
-1 & 2 & -1 \\
& \ddots & \ddots & \ddots \\
& & -1 & 2 & -1 \\
& & & -1 & 2
\end{pmatrix} \in \mathbb{R}^{n \times n},
$$

(18)

The matrices $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$ are equal to zero except for the entries $[n/3]$ and $[2n/3]$, respectively, which are equal to one.

Our goal is to compute the diagonal controllability Gramians $P$ for various $n$ with a minimal trace, which we will do in the dual form:

$$
\begin{align*}
\max_Y \text{trace}(BB^TY),
\text{s.t. } &
\text{diag}\{YA + A^TY + I\} = 0 \\
& Y > 0.
\end{align*}
$$

In the dual form, we have an LP relaxation where $-Y$ belongs to the dual to the cone of symmetric $DD^+_n$ matrices, and an SOCP relaxation, $-Y$ belongs to the dual to the cone of symmetric $\mathcal{H}_+$ matrices. We solve only the dual SDP formulation and the corresponding relaxation. Due to the structure of the system, the trace of its Gramians does not change much with dimensions and we always get the optimal values in the range between 6.5 to 6.6 for the SDP programme. Remarkably the results for the SOCP relaxation are only slightly higher, but in the same range of values. This however, is due to structure of the system, where the drift is Metzler and the matrix $BB^T$ has only one non-zero entry on the diagonal. The optimal solutions for the LP relaxation are in the range between 10.7 – 10.9, hence there is a drop in quality when using this relaxation.

In Table I we provide the computational times for various systems sizes $n$. The entry ‘NA’ means that the programme terminated due to running out of memory. Note however, that we do not take into account the time for parsing the constraints (that is, we plot only the variable “solvetime” in Yalmip [29]). Since $A$ is a Metzler matrix it is straightforward to find a transformation $T$ such that $TAT^{-1}$ becomes a diagonally dominant matrix (see Theorem I). We have implemented the LP relaxation while transforming the $A$, $B$ matrices with such a transformation $T$. The optimal solutions for the trace vary between 7.1 and 7.3, thus drastically improving the quality of the relaxation with a mild loss (if any) in computational time.

### Table I

| Size of the system | 50  | 100 | 150 | 200 |
|-------------------|-----|-----|-----|-----|
| SDP solution      | 0.94| 22.7| 310.7| NA  |
| SOCP relaxation   | 0.74| 4.11| 11.9| 31.2|
| LP relaxation     | 0.01| 0.02| 0.05| 0.10|
| LP relaxation w scaling | 0.01 | 0.03 | 0.05 | 0.10 |

We have provided some sufficient conditions on $A$, which guarantee the existence of feasible points in $\mathcal{G}$ and interpreted these results as small gain like conditions. Moreover, our sufficient conditions also provide computationally cheap solutions, for example Proposition [2] replaces an LMI constraint with two Riccati Equation solutions. If we drop the “$X$ is $\alpha$-diagonal” constraint and set $Q = AX + XA^T$, then the LMI [2] has a solution for any $Q < 0$ if and only if $\lambda_1(A) + \lambda_1(A) \neq 0$. Since $Q$ is arbitrarily negative definite, we can replace the constraint $AX + XA^T < 0$ with $-AX - XA^T \in DD^+_n$. Thus our solvability LMI becomes a linear program. Finally we show how our constructive proofs can be used to initiate a recently developed basis pursuit algorithm for solving large scale optimization problems.

### References

[1] A. A. Ahmadi and G. Hall, “Sum of squares basis pursuit with linear and second order cone programming,” 2015, to appear in Contemporary Mathematics.

[2] K. Zhou, J. C. Doyle, and K. Glover, Robust and optimal control. Prentice Hall New Jersey, 1996, vol. 40.

[3] F. Lin, M. Fardad, and M. R. Jovanović, “Design of optimal sparse feedback gains via the alternating direction method of multipliers,” *IEEE Trans Automatic Control*, vol. 58, no. 9, pp. 2426–2431, Sep 2013.

[4] H. Sandberg and R. M. Murray, “Model reduction of interconnected linear systems,” Optimal control applications & methods, vol. 30, no. 3, pp. 225–245, 2009.

[5] R. P. Mason and A. Papachristodoulou, “Chordal sparsity, decomposing SDPs and the Lyapunov equation,” in *Proc Am Control Conf*, 2014, pp. 531–537.

[6] S. Kim, M. Kojima, M. Mevissen, and M. Yamashita, “Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion,” *Mathematical programming*, vol. 129, no. 1, pp. 33–68, 2011.

[7] M. Yamashita, K. Fujisawa, and M. Kojima, “SDPARA: Semidefinite programming algorithm parallel version,” *Parallel Computing*, vol. 29, no. 8, pp. 1053–1067, 2003.

[8] A. A. Ahmadi and A. Majumdar, “DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization,” in *Proc Conf Inform Sci Syst*. Princeton University, 2014.

[9] A. Majumdar, A. A. Ahmadi, and R. Tedrake, “Control and verification of high-dimensional systems with DSOS and SDSOS programming,” in *IEEE Conf Decision Control*, 2014, pp. 394–401.

[10] D. Carlson, D. Hershkowitz, and D. Shasha, “Block diagonal semistability factors and lyapunov semistability of block triangular matrices,” *Linear Algebra Appl.*, vol. 172, pp. 1–25, 1992.

[11] J. Anderson and A. Sootta, “Decentralised H2-norm estimation and guaranteed error bounds using structured gramians,” in *Proc Sym Math Theory New Syst*, Groningen, Netherlands, July, 2014.

[12] P. Trnka, C. Stirn, H. Sandberg, V. Havlena, and J. Rehor, “Structured model order reduction of parallel models in feedback,” *IEEE Trans Control Systems Technology*, vol. 21, no. 3, pp. 739–752, 2013.

[13] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. SIAM, 1994, vol. 9.
[14] M. Arcak and E. D. Sontag, “Diagonal stability of a class of cyclic systems and its connection with the secant criterion,” *Automatica*, vol. 42, no. 9, pp. 1531–1537, 2006.

[15] M. Arcak, “Diagonal stability on cactus graphs and application to network stability analysis,” *IEEE Trans Autom Control*, vol. 56, no. 12, pp. 2766–2777, 2011.

[16] D. Hershkowitz and H. Schneider, “Lyapunov diagonal semistability of real H-matrices,” *Linear Algebra Appl.*, vol. 71, pp. 119–149, 1985.

[17] S.-h. Xiang and Z.-y. You, “Weak block diagonally dominant matrices, weak block h-matrix and their applications,” *Linear Algebra Appl.*, vol. 282, no. 1, pp. 263–274, 1998.

[18] D. G. Feingold, R. S. Varga, et al., “Block diagonally dominant matrices and generalizations of the gerschgorin circle theorem,” *Pacific J. Math.*, vol. 12, no. 4, pp. 1241–1250, 1962.

[19] R. S. Varga, “On recurring theorems on diagonal dominance,” *Linear Algebra Appl.*, vol. 13, no. 1, pp. 1–9, 1976.

[20] J. Liu and Y. Huang, “Some properties on schur complements of h-matrices and diagonally dominant matrices,” *Linear Algebra Appl.*, vol. 389, pp. 365–380, 2004.

[21] E. G. Boman, D. Chen, O. Parekh, and S. Toledo, “On factor width and symmetric H-matrices,” *Linear Algebra Appl.*, vol. 405, pp. 239–248, 2005.

[22] S. Boyd, L. Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.

[23] A. Rantzer, “Scalable control of positive systems,” *European Journal of Control*, vol. 24, pp. 72–80, 2015.

[24] A. Sootla and J. Anderson, “Structured projection-based model reduction with application to stochastic biochemical networks,” *Submitted to IEEE Trans. Autom. Control*, Oct. 2015, http://arxiv.org/abs/1510.05784.

[25] C. Sturk, H. Sandberg, P. Trnka, V. Havlena, and J. Rehor, “Structured model order reduction of boiler-header models,” in *Proceedings of the 18th IFAC World Congress*, vol. 18, 2011, pp. 3341–3347.

[26] J. Anderson, A. Teixeira, H. Sandberg, and A. Papachristodoulou, “Dynamical system decomposition using dissipation inequalities,” in *IEEE Conf Decision Control*, 2011, pp. 211–216.

[27] H. K. Khalil, *Nonlinear systems*. Prentice Hall, 2002.

[28] Y. Chahlaoui and P. V. Dooren, “A collection of benchmark examples for model reduction of linear time invariant dynamical systems,” University of Louvain,” SLICOT Working Note 2002-2, February 2002.

[29] J. Lofberg, “YALMIP: A toolbox for modeling and optimization in MATLAB,” in *Proceedings of the 2004 IEEE International Symposium on Computer Aided Control Systems Design*. IEEE, 2004, pp. 284–289.