Robust and Reduced-Order Filtering: New LMI-based Characterizations and Methods

H.D. Tuan[*], **P. Apkarian**[†] and **T. Q. Nguyen**[‡]

Abstract

Several challenging problems of robust filtering are addressed in this paper. First of all, we exploit a new LMI (Linear Matrix Inequality) characterization of minimum variance or of H_2 performance, and demonstrate that it allows the use of parameter-dependent Lyapunov functions while preserving tractability of the problem. The resulting conditions are less conservative than earlier techniques which are restricted to a fixed, that is not depending on parameters, Lyapunov functions. The rest of the paper is focusing on reduced-order filter problems. New LMI-based nonconvex optimization formulations are introduced for the existence of reduced-order filters. Then, several efficient optimization algorithms of local and global optimization are proposed. Nontrivial and less conservative relaxation techniques are discussed as well. The viability and efficiency of the proposed tools are confirmed through computational experiments and also through comparisons with earlier methods.

1 Introduction

The standard robust filter problem can be formulated as follows. Consider the uncertain linear system

$$\begin{array}{rcll} \dot{x} & = & Ax + Bw, & A \in \boldsymbol{R}^{n\times n} \\ y & = & Cx + Dw, & D \in \boldsymbol{R}^{p\times m} \\ z & = & Lx\,, & L \in \boldsymbol{R}^{q\times n} \end{array} \tag{1}$$

where $x \in R^n$ is the state, $y \in R^p$ is the measured output, $z \in R^q$ is the output to be estimated and $w \in R^m$ is the zero mean white noise with identity power spectrum density matrix. The state-space data are subject to uncertainties and obey the polytopic model

$$\begin{bmatrix} A & B \\ C & D \\ L & 0 \end{bmatrix} \in \left\{ \begin{bmatrix} A(\alpha) & B(\alpha) \\ C(\alpha) & D(\alpha) \\ L(\alpha) & 0 \end{bmatrix} = \sum_{i=1}^{s} \alpha_i \begin{bmatrix} A_i & B_i \\ C_i & D_i \\ L_i & 0 \end{bmatrix}, \alpha \in \Gamma \right\}, \tag{2}$$

where Γ is the unit simplex

$$\Gamma := \{(\alpha_1, ..., \alpha_s) : \sum_{i=1}^{s} \alpha_i = 1,\ \alpha_i \geq 0\}\,.$$

[*]Department of Control and Information, Toyota Technological Institute, Hisakata 2-12-1, Tenpaku, Nagoya 468-8511, JAPAN- Email: tuan@toyota-ti.ac.jp

[†]ONERA-CERT, 2 av. Edouard Belin, 31055 Toulouse, FRANCE - Email: apkarian@cert.fr

[‡]Department of Electrical and Computer Engineering, Boston University, 8 St. Mary St., Boston MA 02215, USA- Email: nguyen@bu.edu

The problem consists in constructing an estimator or "filter" in the form

$$\begin{array}{rcll} \dot{x}_F & = & A_F x_F + B_F y, & A_F \in \boldsymbol{R}^{k\times k} \\ z_F & = & L_F x_F, & L_F \in \boldsymbol{R}^{q\times k} \end{array} \tag{3}$$

which provides good robust estimation in the minimum variance sense of the output z in (1). In other words, we want to minimize

$$\max_{\alpha\in\Gamma} \mathbf{E}[(z - z_F)^T (z - z_F)] \tag{4}$$

where $\mathbf{E}$ means the mathematical expectation. Note that the expression (4) involves all possible value of the uncertainty α, hence the term robust filtering problem. When $k = n$ the filter (3) will be referred to as the full-order filter and will be termed reduced-order when $k < n$.

Classically, when all data of the system (1) are exactly known, the optimal value of (4) is $Tr(LPL^T)$ and the optimal full-order solution is the well-known Kalman filter [1] defined as

$$A_F = A - B_F C, \quad B_F = PC^T(DD^T)^{-1}, \quad L_F = L\,,$$

where $P \geq 0$ is the stabilizing solution of the Riccati equation

$$AP + PA^T - PC^T(DD^T)^{-1}CP + BB^T = 0. \tag{5}$$

Note that the existence of the stabilizing solution $P \geq 0$ of Riccati equation (5) implies that matrix A in (1) must be asymptotically stable.

An alternative solution to the full-order filter problem with exact data can be obtained by using LMI characterizations. Indeed, rewrite (1)-(3) in compact form as

$$\begin{array}{rcl} \dot{x}_{cl} & = & \mathcal{A}_{cl} x_{cl} + \mathcal{B}_{cl} w, \\ z_{cl} & = & [\,L \quad -L_F\,]\, x_{cl}, \end{array} \tag{6}$$

where

$$x_{cl} = \begin{bmatrix} x \\ x_F \end{bmatrix}, \quad \mathcal{A}_{cl} = \begin{bmatrix} A & 0 \\ B_F C & A_F \end{bmatrix}, \quad \mathcal{B}_{cl} = \begin{bmatrix} B \\ B_F D \end{bmatrix}, \quad z_{cl} = z - z_F. \tag{7}$$

Then, it has been established (see e.g. [7]) that $\mathbf{E}(z_{cl}^2) < \nu$ if and only if the following matrix inequalities are feasible in the variables $\mathcal{X}$, Z, A_F, B_F and L_F

$$\begin{bmatrix} \mathcal{A}_{cl}^T \mathcal{X} + \mathcal{X}\mathcal{A}_{cl} & \mathcal{X}\mathcal{B}_{cl} \\ \mathcal{B}_{cl}^T \mathcal{X} & -I \end{bmatrix} < 0, \tag{8}$$

$$\begin{bmatrix} \mathcal{X} & \begin{bmatrix} L^T \\ -L_F^T \end{bmatrix} \\ [\,L \quad -L_F\,] & Z \end{bmatrix} > 0, \tag{9}$$

$$\mathrm{Tr}(Z) < \nu\,. \tag{10}$$

Thus, the problem can be formulated alternatively as

$$\min\{\nu : \ (8) - (10)\}\,. \tag{11}$$

A quick justification can be inferred as follows. It is well known that

$$\mathbf{E}(z_{cl}^T z_{cl}) = \mathrm{Tr}(\mathcal{C}_{cl}\mathcal{P}\mathcal{C}_{cl}^T), \tag{12}$$

where $\mathcal{C}_{cl} = [\, L \quad -L_F \,]$ and $\mathcal{P}$ is the solution of the Lyapunov equation

$$\mathcal{P}\mathcal{A}_{cl}^T + \mathcal{A}_{cl}\mathcal{P} + \mathcal{B}_{cl}\mathcal{B}_{cl}^T = 0. \tag{13}$$

From (8), we infer

$$\mathcal{X}^{-1}\mathcal{A}_{cl}^T + \mathcal{A}_{cl}\mathcal{X}^{-1} + \mathcal{B}_{cl}\mathcal{B}_{cl}^T < 0$$

and thus with $\mathcal{A}_{cl}$ assumed stable, we have $\mathcal{P} < \mathcal{X}^{-1}$, which together with (9) and (10) gives

$$\mathbf{E}(z_{cl}^T z_{cl}) = \mathrm{Tr}(\mathcal{C}_{cl}\mathcal{P}\mathcal{C}_{cl}^T) < \mathrm{Tr}(\mathcal{C}_{cl}\mathcal{X}^{-1}\mathcal{C}_{cl}^T) < \mathrm{Tr}(Z) < \nu.$$

Note that (8) is a nonlinear matrix inequality in the variables A_F, B_F and $\mathcal{X}$ because of the product terms $\mathcal{X}\mathcal{A}_{cl}$. However, there are several ways to reduce it to LMIs by linearizing techniques in the spirit of [11, 14] or by using the Projection Lemma in [9]. As a result, the problem can be reduced to the convex optimization problem of minimization of a linear objective over LMI constraints, an easily tractable problem with the help of currently available SemiDefinite Programming solvers.

The advantage of the proposed LMI approach, and this is an important contribution of this paper, is that it still works for the problem with unknown data as in (2) and with parameter-dependent functions $\mathcal{X} := \mathcal{X}(\alpha)$. Note that on one hand uncertainties are hardly handled with the Riccati equation approach and, on the other hand, parameter-dependent Lyapunov functions are far less conservative than customary fixed quadratic Lyapunov functions.

Similarly to other results in the vein of robust control for polytopic systems, a variable $\mathcal{X}$, not depending on the uncertainties, has been utilized in [11, 14] for verifying the conditions (8) and (9). The resulting robust estimation may be conservative as the function $\mathcal{X}$ used for verifying (8) and (9) is fixed for all values of the parameter α. This is well-known to be a source of conservatism in applications. The issue of exploiting parameter-dependent functions $\mathcal{X}(\alpha)$ to handle problems with uncertainties is very challenging both in robust control and robust filtering. The latter issue is examined throughout this paper. We extend the results in [5] to the robust filtering problem, and derive specific linearizing transformations which lead to tractable LMI conditions for the full-order robust filtering problem with parameter-dependent functions. These results are naturally less conservative than previous ones. The latter are recovered as a special case by imposing a constant Lyapunov function ($\mathcal{X}(\alpha) := \mathcal{X},\ \forall\alpha$) in the proposed approach.

Another challenging problem is the reduced-order filter problem which is known to be nonconvex even in the exact data case. These problems have been partially addressed e.g. in [14, 18]. Here we will propose several new solution methods for this problem. Namely, we introduce a new LMI characterization which allows us to propose several less conservative relaxations of the original problem. These relaxations are combined with specialized local and global optimization algorithms to construct good practical solutions.

The paper is organized as follows. Robust filter problems are considered in Section 2. Section 3 is devoted to the reduced-order case with known system data and discusses various relaxations and algorithms. Finally, examples illustrating the viability and efficiency of our approach are given in Section 4.

The notation used throughout the paper is standard. M^T is the transpose of a matrix M while $\mathrm{Tr}(M)$ is its trace. The notation $M < 0$ ($M \leq 0$, resp.) means that M is negative definite (negative semi-definite, resp.). In symmetric block matrices or long matrix expressions, we use $*$ as an ellipsis for terms that can be induced by symmetry. A basis of the nullspace of a matrix A will be denoted $\mathcal{N}_A$.

Some technical results in the paper make use of the Projection Lemma.

Lemma 1 (Projection Lemma) *[9] Given a symmetric matrix $\Psi \in \boldsymbol{R}^{m\times m}$ and two matrices P, Q of column dimension m, the following problem*

$$\Psi + P^T X^T Q + Q^T X P < 0.$$

is solvable in a matrix X of compatible dimension if and only if

$$\mathcal{N}_P^T \Psi \mathcal{N}_P < 0,\ \mathcal{N}_Q^T \Psi \mathcal{N}_Q < 0$$

where $\mathcal{N}_P$ and $\mathcal{N}_Q$ are any basis of the nullspace of P and Q, respectively.

2 Robust minimum variance filter problem

A drawback of the standard matrix inequality characterization (8)-(9) is that the function $\mathcal{X}$ used for checking the filter performance is closely interrelated with the state-space variables A_F and B_F. This makes the problem difficult to solve and causes unnecessary restrictions on the filter variables. This is particularly critical when uncertainties come into play as for polytopic systems (2). To overcome this difficulty, we exploit a reciprocal variant of the Projection Lemma 1 to alleviate the interrelation between $\mathcal{X}$ and filter variables. This technique introduces an extra slack variable V which brings additional flexibility in the robust filtering problem. An important consequence is that the apparent nonconvexity of filter synthesis problems with parameter-dependent Lyapunov functions can be bypassed. Moreover, because the function $\mathcal{X}$ is depending on uncertain parameters the resulting characterizations are generally far less conservative than customary single quadratic approaches. The following lemma will be useful in that respect. Note that for simplicity of the presentation, we shall drop the dependence of variables and data on uncertainties α for a while.

Lemma 2 *The constraints (8)-(10) are feasible in A_F, B_F, C_F,$\mathcal{X}$, Z and ν if and only if the following conditions are feasible in A_F, B_F, L_F, $\mathcal{X}$, Z, ν and V*

$$\begin{bmatrix} -(V+V^T) & V^T\mathcal{A}_{cl}+\mathcal{X} & V^T\mathcal{B}_{cl} & V^T \\ \mathcal{A}_{cl}^T V+\mathcal{X} & -\mathcal{X} & 0 & 0 \\ \mathcal{B}_{cl}^T V & 0 & -I & 0 \\ V & 0 & 0 & -\mathcal{X} \end{bmatrix} < 0, \tag{14}$$

$$\begin{bmatrix} \mathcal{X} & \begin{bmatrix} L^T \\ -L_F^T \end{bmatrix} \\ \begin{bmatrix} L & -L_F \end{bmatrix} & Z \end{bmatrix} > 0, \tag{15}$$

$$\mathit{Tr}(Z) < \nu. \tag{16}$$

Proof: Rewrite (14) as

$$\begin{bmatrix} 0 & \mathcal{X} & 0 & 0 \\ \mathcal{X} & -\mathcal{X} & 0 & 0 \\ 0 & 0 & -I & 0 \\ 0 & 0 & 0 & -\mathcal{X} \end{bmatrix} + \mathcal{P}^T V \mathcal{Q} + \mathcal{Q}^T V^T \mathcal{P} < 0 \tag{17}$$

with

$$\mathcal{P} = [\, -I \quad \mathcal{A}_{cl} \quad \mathcal{B}_{cl} \quad I \,], \ \mathcal{Q} = [\, I \quad 0 \quad 0 \quad 0 \,].$$

Noting that

$$\mathcal{N}_{\mathcal{P}} = \begin{bmatrix} \mathcal{A}_{cl} & \mathcal{B}_{cl} & I \\ I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}, \ \mathcal{N}_{\mathcal{Q}} = \begin{bmatrix} 0 & 0 & 0 \\ I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

and aplying the Projection lemma 1 with respect to variable V in (17), the existence of V and $\mathcal{X}$ satisfying (17) is equivalent to the existence of $\mathcal{X}$ satisfying the inequality

$$\begin{bmatrix} \mathcal{A}_{cl}^T \mathcal{X} + \mathcal{X}\mathcal{A}_{cl} - \mathcal{X} & \mathcal{X}\mathcal{B}_{cl} & \mathcal{X} \\ \mathcal{B}_{cl}^T \mathcal{X} & -I & 0 \\ \mathcal{X} & 0 & -\mathcal{X} \end{bmatrix} < 0$$

which is equivalent to (8) by a Schur's complement argument.

See also [5] for more details on other equivalent LMI characterizations to (8)-(10).

2.1 Robust full-order filters

By exploiting Lemma 2, it is possible to derive tractable synthesis conditions for the robust filter problem in the full-order case. Hereafter, we shall need partitions of V and $\mathcal{X}$ in (14)-(15) in the form

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}, \qquad \mathcal{X} = \begin{bmatrix} X_1 & X_3^T \\ X_3 & X_2 \end{bmatrix}, \tag{18}$$

where for the full-order case, all submatrices appearing in (18) are of size $n \times n$.

Note that (14) implies that V is nonsingular and by using a small perturbation if necessary, we can also assume, without loss of generality, that V_{21}, V_{22} are nonsingular as well [5].

Now, using the notations (7) and (18), the general matrix inequality (14) becomes

$$\begin{bmatrix} -(V_{11}+V_{11}^T) & * & * & * & * & * & * \\ -(V_{21}+V_{12}^T) & -(V_{22}+V_{22}^T) & * & * & * & * & * \\ A^T V_{11} + C^T B_F^T V_{21} + X_1 & A^T V_{12} + C^T B_F^T V_{22} + X_3^T & -X_1 & * & * & * & * \\ A_F^T V_{21} + X_3 & A_F^T V_{22} + X_2 & -X_3 & -X_2 & * & * & * \\ B^T V_{11} + D^T B_F^T V_{21} & B^T V_{12} + D^T B_F^T V_{22} & 0 & 0 & -I & * & * \\ V_{11} & V_{12} & 0 & 0 & 0 & -X_1 & * \\ V_{21} & V_{22} & 0 & 0 & 0 & -X_3 & -X_2 \end{bmatrix} < 0 \tag{19}$$

This condition can then be turned into an LMI constraint in two steps:

1. perform in (19) the congruence transformation

$$\operatorname{diag}\left[\, I \quad V_{22}^{-1}V_{21} \quad I \quad V_{22}^{-1}V_{21} \quad I \quad I \quad V_{22}^{-1}V_{21} \,\right] \tag{20}$$

2. introduce the (new) linearization transformations

$$\begin{aligned}
\widehat{A}_F &= V_{21}^T A_F V_{22}^{-1} V_{21} && (21)\\
\widehat{B}_F &= V_{21}^T B_F && (22)\\
S_1 &= V_{21}^T V_{22}^{-T} V_{21} && (23)\\
S_2 &= V_{21}^T V_{22}^{-T} V_{12}^T && (24)\\
\widehat{X} &= \begin{bmatrix} \widehat{X}_1 & \widehat{X}_3^T \\ \widehat{X}_3 & \widehat{X}_2 \end{bmatrix} := \begin{bmatrix} I & 0 \\ 0 & V_{21}^T V_{22}^{-T} \end{bmatrix} \mathcal{X} \begin{bmatrix} I & 0 \\ 0 & V_{22}^{-1} V_{21} \end{bmatrix}. && (25)
\end{aligned}$$

Note that because of the invertibility of the matrices involved, the linearizing transformations are back and forth, thus the resulting conditions are equivalent. We end up with the following LMI condition:

$$\begin{bmatrix}
-(V_{11}+V_{11}^T) & * & * & * & * & * & * \\
-(S_1+S_2) & -(S_1+S_1^T) & * & * & * & * & * \\
A^T V_{11} + C^T \widehat{B}_F^T + \widehat{X}_1 & A^T S_2^T + C^T \widehat{B}_F^T + \widehat{X}_3^T & -\widehat{X}_1 & * & * & * & * \\
\widehat{A}_F^T + \widehat{X}_3 & \widehat{A}_F^T + \widehat{X}_2 & -\widehat{X}_3 & -\widehat{X}_2 & * & * & * \\
B^T V_{11} + D^T \widehat{B}_F^T & B^T S_2^T + D^T \widehat{B}_F^T & 0 & 0 & -I & * & * \\
V_{11} & S_2^T & 0 & 0 & 0 & -\widehat{X}_1 & * \\
S_1 & S_1 & 0 & 0 & 0 & -\widehat{X}_3 & -\widehat{X}_2
\end{bmatrix} < 0\,. \tag{26}$$

Similarly, again with the notations (7) and (18), LMI (15) becomes

$$\begin{bmatrix} X_1 & X_3^T & L^T \\ X_3 & X_2 & -L_F^T \\ L & -L_F & Z \end{bmatrix} > 0\,. \tag{27}$$

The congruent transformation

$$\operatorname{diag}\left[\, I \quad V_{22}^{-1}V_{21} \quad I \,\right] \tag{28}$$

then yields

$$\begin{bmatrix} \widehat{X}_1 & \widehat{X_3}^T & L^T \\ \widehat{X}_3 & \widehat{X}_2 & -\widehat{L}_F^T \\ L & -\widehat{L}_F & Z \end{bmatrix} > 0\,, \tag{29}$$

where $\widehat{L}_F$ is defined as

$$\widehat{L}_F := L_F V_{22}^{-1} V_{21}\,. \tag{30}$$

Summing up , we have derived the following intermediate result.

Lemma 3 *The nonlinear matrix inequalities (14)-(16) are feasible in*

$$\begin{bmatrix} A_F & B_F \\ L_F & 0 \end{bmatrix}, \quad \mathcal{X}, \quad Z, \quad V, \quad \nu\,,$$

if and only if LMIs (16), (26) and (29) are feasible with respect to

$$\begin{bmatrix} \widehat{A}_F & \widehat{B}_F \\ \widehat{L}_F & 0 \end{bmatrix}, \quad V_{11}, \quad \widehat{X}, \quad S_1, \quad S_2, \quad \nu\,. \tag{31}$$

The triple (A_F, B_F, L_F) defining the full-order filter (3) is then readily derived from the variables in (31) solution to LMIs (16), (26) and (29) according to the following steps:

(i) *compute V_{22}, V_{21} by solving the factorization problem*

$$S_1 = V_{21}^T V_{22}^{-1} V_{21}\,.$$

(ii) *compute (A_F, B_F, L_F) using*

$$\begin{bmatrix} A_F & B_F \\ L_F & 0 \end{bmatrix} := \begin{bmatrix} V_{21}^{-T} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \widehat{A}_F & \widehat{B}_F \\ \widehat{L}_F & 0 \end{bmatrix} \begin{bmatrix} V_{21}^{-1}V_{22} & 0 \\ 0 & I \end{bmatrix}. \tag{32}$$

Compared with the linearization techniques used in [5, 11, 14], the advantage of the transformations (21)-(25) and (30) is that the intermediate variables $(\widehat{A}_F, \widehat{B}_F, \widehat{L}_F)$ are independent of the system data so that the very same approach is still valid for systems depending on uncertain parameters. The link between these entities and the variables $\widehat{X}$ is, as in [5], via the slack variable V. These features are crucial in dealing with uncertain systems of the class (2). Indeed, from (26) and (29), and the system data satisfying (2), we are allowed to use parameter-dependent function $\widehat{X}(\alpha)$ in the form

$$\mathcal{X}(\alpha) = \sum_{i=1}^{s} \alpha_i \mathcal{X}_i := \sum_{i=1}^{s} \alpha_i \begin{bmatrix} X_{1,i} & X_{3,i}^T \\ X_{3,i} & X_{2,i} \end{bmatrix}, \quad \alpha \in \Gamma \tag{33}$$

for enforcing conditions (14)-(15) while still preserving the problem tractability. The main result of this section is Theorem 1 which characterizes robust estimation in the minimum variance sense with the help of such "polytopic" Lyapunov functions. Note also that $\mathcal{X}(\alpha)$ is positive definite for all admissible values of α if and only if this holds for the $\mathcal{X}_i$'s.

Theorem 1 (robust full-order) *There exists a (full-order) filter such that the worst-case condition*

$$\max_{\alpha \in \Gamma} \mathbf{E}[(z - z_F)^T (z - z_F)] < \nu\,,$$

holds true, that is for all admissible systems described in (2), whenever the following (vertex) conditions hold simultaneously:

$$\begin{bmatrix} -(V_{11}+V_{11}^T) & * & * & * & * & * & * \\ -(S_1+S_2) & -(S_1+S_1^T) & * & * & * & * & * \\ A_i^T V_{11} + C_i^T \widehat{B}_F^T + \widehat{X}_{1,i} & A_i^T S_2^T + C_i^T \widehat{B}_F^T + \widehat{X}_{3,i}^T & -\widehat{X}_{1,i} & * & * & * & * \\ \widehat{A}_F^T + \widehat{X}_{3,i} & \widehat{A}_F^T + \widehat{X}_{2,i} & -\widehat{X}_{3,i} & -\widehat{X}_{2,i} & * & * & * \\ B_i^T V_{11} + D_i^T \widehat{B}_F^T & B_i^T S_2^T + D_i^T \widehat{B}_F^T & 0 & 0 & -I & * & * \\ V_{11} & S_2^T & 0 & 0 & 0 & -\widehat{X}_{1,i} & * \\ S_1 & S_1 & 0 & 0 & 0 & -\widehat{X}_{3,i} & -\widehat{X}_{2,i} \end{bmatrix} < 0 \tag{34}$$

$$\begin{bmatrix} \widehat{X}_{1,i} & \widehat{X}_{3,i}^T & L_i^T \\ \widehat{X}_{3,i} & \widehat{X}_{2,i} & -\widehat{L}_F^T \\ L_i & -\widehat{L}_F & Z \end{bmatrix} > 0 \tag{35}$$

$$i = 1, \ldots, s,$$

together with (16), with the notation

$$
\begin{bmatrix} \widehat{X}_{1,i} & \widehat{X}_{3,i}^T \\ \widehat{X}_{3,i} & \widehat{X}_{2,i} \end{bmatrix} := \begin{bmatrix} I & 0 \\ 0 & V_{21}^T V_{22}^{-T} \end{bmatrix} \begin{bmatrix} X_{1,i} & X_{3,i}^T \\ X_{3,i} & X_{2,i} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & V_{22}^{-1} V_{21} \end{bmatrix} . \tag{36}
$$

The sought triple (A_F, B_F, L_F) defining the full-order filter (3) can then be computed as in Lemma 3. The polytopic function establishing robust estimation is given by (33) and (36).

Consequently, a best upper bound of the minimum of (4) is provided by the optimization problem

$$
\min_{V_{11}, S_1, S_2, \widehat{A}_F, \widehat{B}_F, \widehat{L}_F, \widehat{X}_i, \nu} \{\nu : \ (16), (34), (35), \ i = 1, 2, ..., s.\} . \tag{37}
$$

Proof: The proof is immediate from Lemma 3 and the properties of convex combinations.

2.2 LMI relaxation for robust reduced-order filters

Hereafter, we consider the case when the order of the filter is set to $k < n$. Then, of course (14)-(16) with $\mathcal{A}_{cl}, \mathcal{B}_{cl}, \mathcal{C}_{cl}$ defined by (7) are still in force. However, with the partition (18) the matrix V_{21} becomes rectangular of dimension $k \times n$. This makes the change of variable (32) no longer valid. One can get rid of this difficulty by imposing some (possibly conservative) special structure on the slack variable V_{21}. With such a restriction, similar linearizations are possible. Indeed, take

$$
V_{21} = [\,\widetilde{V}_{21} \quad 0_{k\times(n-k)}\,] \tag{38}
$$

where $\widetilde{V}_{21}$ is a square matrix of dimension $k \times k$, which is supposed to be regular. Then, performing the congruent transformation

$$
\mathrm{diag}\,[\,I \quad V_{22}^{-1}\widetilde{V}_{21} \quad I \quad V_{22}^{-1}\widetilde{V}_{21} \quad I \quad I \quad V_{22}^{-1}\widetilde{V}_{21}\,] \tag{39}
$$

in (19), yields the following LMI

$$
\begin{bmatrix}
-(V_{11}+V_{11}^T) & * & * & * & * & * & * \\
-(\widetilde{S}_1+S_2) & -(S_1+S_1^T) & * & * & * & * & * \\
A^T V_{11} + C^T \tilde{B}_F^T + \widehat{X}_1 & A^T S_2^T + C^T \widehat{B}_F^T + \widehat{X}_3^T & -\widehat{X}_1 & * & * & * & * \\
\tilde{A}_F + \widehat{X}_3 & \widehat{A}_F + \widehat{X}_2 & -\widehat{X}_3 & -\widehat{X}_2 & * & * & * \\
B^T V_{11} + D^T \tilde{B}_F^T & B^T S_2^T + D^T \widehat{B}_F^T & 0 & 0 & -I & * & * \\
V_{11} & S_2^T & 0 & 0 & 0 & -\widehat{X}_1 & * \\
\widetilde{S}_1 & S_1 & 0 & 0 & 0 & -\widehat{X}_3 & -\widehat{X}_2
\end{bmatrix} < 0 \tag{40}
$$

with the definitions

$$\begin{aligned}
\widetilde{B}_F &= \begin{bmatrix} \widehat{B}_F \\ 0_{(n-k)\times p} \end{bmatrix}, \quad \widehat{B}_F = \widetilde{V}_{21}^T B_F \\
\widetilde{S}_1 &= [\, S_1 \quad 0_{k\times(n-k)} \,], \quad S_1 = \widetilde{V}_{21}^T V_{22}^{-T} \widetilde{V}_{21}, \\
S_2 &= \widetilde{V}_{21}^T V_{22}^{-T} \widetilde{V}_{12}^T, \\
\widetilde{A}_F &= [\, \widehat{A}_F \quad 0_{k\times(n-k)} \,], \quad \widehat{A}_F = \widetilde{V}_{21}^T A_F \widetilde{V}_{22}^{-1} \widetilde{V}_{21}, \\
\begin{bmatrix} \widehat{X}_1 & \widehat{X}_3^T \\ \widehat{X}_3 & \widehat{X}_2 \end{bmatrix} &= \begin{bmatrix} I & 0 \\ 0 & \widetilde{V}_{21}^T V_{22}^{-T} \end{bmatrix} \mathcal{X} \begin{bmatrix} I & 0 \\ 0 & V_{22}^{-1} \widetilde{V}_{21} \end{bmatrix}
\end{aligned} \tag{41}$$

Analogously, the congruent transformation

$$\operatorname{diag}[\, I \quad V_{22}^{-1}\widetilde{V}_{21} \quad I \,] \tag{42}$$

allow to turn (27) into (29) with

$$\widehat{L}_F = L_F V_{22}^{-1} \widetilde{V}_{21}. \tag{43}$$

Thus, we are now in a position to state the following reduced-order counterpart of Theorem 1.

Theorem 2 (relaxation for robust reduced-order filter) *There exists a reduced kth-order filter such that the worst-case condition*

$$\max_{\alpha\in\Gamma} \mathbf{E}[(z-z_F)^T(z-z_F)] < \nu \,,$$

holds true, that is for the uncertain system described in (2), whenever LMIs (16), (29) and (40) are satisfied for all vertex indices $i = 1, \ldots, s$, i.e. (16), (35), and

$$\begin{bmatrix}
-(V_{11}+V_{11}^T) & * & * & * & * & * & * \\
-(\widetilde{S}_1+S_2) & -(S_1+S_1^T) & * & * & * & * & * \\
A_i^T V_{11} + C_i^T \tilde{B}_F^T + \widehat{X}_{1,i} & A_i^T S_2^T + C_i^T \widehat{B}_F^T + \widehat{X}_{3i}^T & -\widehat{X}_{1,i} & * & * & * & * \\
\tilde{A}_F + \widehat{X}_{3,i} & \widehat{A}_F + \widehat{X}_{2,i} & -\widehat{X}_{3,i} & -\widehat{X}_{2,i} & * & * & * \\
B_i^T V_{11} + D_i^T \tilde{B}_F^T & B_i^T S_2^T + D_i^T \widehat{B}_F^T & 0 & 0 & -I & * & * \\
V_{11} & S_2^T & 0 & 0 & 0 & -\widehat{X}_{1,i} & * \\
\widetilde{S}_1 & S_1 & 0 & 0 & 0 & -\widehat{X}_{3,i} & -\widehat{X}_{2,i}
\end{bmatrix} < 0 \tag{44}$$

$$i = 1, 2, ..., s.$$

The triple (A_F, B_F, L_F) defining the kth-order filter (3) is obtained from the solution $V_{11}, S_1, S_2, \widehat{A}_F, \widehat{B}_F, \widehat{L}_F$, of these LMIs by steps analogous to those in Theorem 1 with V_{21} replaced with $\widetilde{V}_{21}$ while $\mathcal{X}(\alpha)$ verifying (14)-(16) is defined as

$$\mathcal{X}(\alpha) = \begin{bmatrix} I & 0 \\ 0 & V_{22}^T \widetilde{V}_{21}^{-T} \end{bmatrix} \sum_{i=1}^{s} \alpha_i \widehat{X}_i \begin{bmatrix} I & 0 \\ 0 & \widetilde{V}_{21}^{-1} V_{22} \end{bmatrix}. \tag{45}$$

Consequently, an upper bound of the minimum of (4) for reduced-order ($k < n$) filters is provided by the LMI optimization problem

$$\min_{V_{11}, S_1, S_2, \widehat{A}_F, \widehat{B}_F, \widehat{L}_F, \nu, \widehat{X}_i} \{\nu : \; (16), (35), (44)\}. \tag{46}$$

3 Optimal reduced-order filter

An obvious advantage of the linearization techniques in the previous section is that it provides an accurate and practically tractable approach for the full-order robust filtering problems. It also gives a relaxation for the reduced-order case. This relaxation may, however, be arbitrarily conservative. The purpose of this section is to derive a convenient optimization formulation for the synthesis of reduced-order filters with exact data in (2), that is $A_i := A$, $B_i := B$, etc. Even with this simplification in force, the reduced-order case is very hard because of its inherent nonconvexity. To tackle this problem, the number of complicating or nonconvex variables is reduced as much as possible. The nonconvex constraints are reformulated in such a way that they are easily handled by optimization algorithms. Local and global optimization techniques are considered at different stages of a general optimization process to improve efficiency. Conjointly, new relaxation techniques are introduced.

Note that in [18] a problem related to the reduced-order filter has also been considered. Instead of the optimal kth-order filter problem, it is concerned with the kth-order filter which approximates the full-order Kalman filter. This approach provides only more or less accurate solutions to the problem with unknown degree of optimality. In this setting, an iterative algorithm has been proposed in [18] that generates a convergent sequence to a stationary point. This algorithm requires either solving multidimensional differential equations or computations involving the exponential function of matrices which are computationally costly.

3.1 Rank-constrained LMI formulation and relaxation

When all matrices A, B, C and D in (1) are exactly known, we shall use an alternative characterization derived below.

Set $K := [\, A_F \quad B_F \,]$ and re-express $\mathcal{A}_{cl}, \mathcal{B}_{cl}$ in (7) as

$$\mathcal{A}_{cl} = A_a + B_a K C_a,\ \mathcal{B}_{cl} = B_{1,a} + B_a K D_{21,a}, \tag{47}$$

with the notations

$$\begin{aligned} &A_a = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, B_a = \begin{bmatrix} 0 \\ I \end{bmatrix},\ C_a = \begin{bmatrix} 0 & I \\ C & 0 \end{bmatrix},\ B_{1,a} = \begin{bmatrix} B \\ 0 \end{bmatrix},\ D_{21,a} = \begin{bmatrix} 0 \\ D \end{bmatrix}, \\ &C_{1,a} = [\, L \quad 0 \,],\ D_a = [\, 0 \quad -I \,]\ . \end{aligned} \tag{48}$$

Then rewrite (8)-(9) as

$$\begin{bmatrix} A_a^T\mathcal{X} + \mathcal{X}A_a & \mathcal{X}B_{1,a} \\ B_{1,a}^T\mathcal{X} & -I \end{bmatrix} + \begin{bmatrix} C_a^T \\ D_{21,a}^T \end{bmatrix} K^T [\, B_a^T\mathcal{X} \quad 0 \,] + \begin{bmatrix} \mathcal{X}B_a \\ 0 \end{bmatrix} K [\, C_a \quad D_{21,a} \,] < 0, \tag{49}$$

$$\begin{bmatrix} \mathcal{X} & C_{1,a}^T \\ C_{1,a} & Q \end{bmatrix} + \begin{bmatrix} 0 \\ D_a^T \end{bmatrix} L_F^T [\, 0 \quad I \,] + \begin{bmatrix} 0 \\ I \end{bmatrix} L_F [\, D_a \quad 0 \,] > 0\,. \tag{50}$$

Using the Projection Lemma 1, the existence of K in (49) is equivalent to

$$\mathcal{N}_E^T \begin{bmatrix} A_a^T\mathcal{X} + \mathcal{X}A_a & \mathcal{P}B_{1,a} \\ B_{1,a}^T\mathcal{X} & -I \end{bmatrix} \mathcal{N}_E < 0 \tag{51}$$

$$\mathcal{N}_G^T \begin{bmatrix} A_a^T\mathcal{X} + \mathcal{X}A_a & \mathcal{X}B_{1,a} \\ B_{1,a}^T\mathcal{X} & -I \end{bmatrix} \mathcal{N}_G < 0 \tag{52}$$

with

$$E := [\,B_a^T\mathcal{X} \quad 0\,] = [\,B_a^T \quad 0\,]\begin{bmatrix}\mathcal{X} & 0\\ 0 & I\end{bmatrix}, \quad G := [\,C_a \quad D_{21,a}\,]\,.$$

It is readily seen that

$$\mathcal{N}_E = \begin{bmatrix}\mathcal{X}^{-1} & 0\\ 0 & I\end{bmatrix}\mathcal{N}_{[\,B_a^T \quad 0\,]}\,.$$

Therefore, we get the equivalences

$$(51) \quad \Leftrightarrow \quad \mathcal{N}_{[\,B_a^T \quad 0\,]}^T\begin{bmatrix}\mathcal{X}^{-1}A_a^T + A_a\mathcal{X}^{-1} & B_{1,a}\\ B_{1,a}^T & -I\end{bmatrix}\mathcal{N}_{[\,B_a^T \quad 0\,]} < 0\,, \tag{53}$$

$$(52) \quad \Leftrightarrow \quad \mathcal{N}_G^T\begin{bmatrix}A_a^T\mathcal{X} + \mathcal{X}A_a & \mathcal{P}B_{1,a}\\ B_{1,a}^T\mathcal{X} & -I\end{bmatrix}\mathcal{N}_G < 0\,. \tag{54}$$

Partitioning $\mathcal{X}$ and its inverse as

$$\mathcal{X} = \begin{bmatrix}X & N\\ N^T & *\end{bmatrix} > 0, \qquad \mathcal{X}^{-1} = \begin{bmatrix}Y & M\\ M^T & *\end{bmatrix} > 0 \tag{55}$$

and using the relationships

$$\mathcal{N}_{[\,B_a^T \quad 0\,]} = \begin{bmatrix}I & 0\\ 0 & 0\\ 0 & I\end{bmatrix}, \quad \mathcal{N}_G = \begin{bmatrix}W_1\\ 0\\ W_2\end{bmatrix}, \text{ where } \begin{bmatrix}W_1\\ W_2\end{bmatrix} = \mathcal{N}_{[\,C \quad D\,]},$$

it is readily checked that (53) and (54) reduce to

$$\begin{bmatrix}YA^T + AY & B\\ B^T & -I\end{bmatrix} < 0\,, \tag{56}$$

$$\mathcal{N}_{[\,C \quad D\,]}^T\begin{bmatrix}XA + A^TX & XB\\ B^TX & -I\end{bmatrix}\mathcal{N}_{[\,C \quad D\,]} < 0\,. \tag{57}$$

We note that (53) and (54) are LMIs in (X, Y). Similarly, by virtue of the Projection Lemma, the existence of L_F in (50) is equivalent to $\mathcal{X} > 0$ and the feasibility of the following LMI

$$\begin{bmatrix}X & L^T\\ L & Z\end{bmatrix} > 0\,. \tag{58}$$

The last point now is the condition imposed on X and Y that makes the completion (55) possible. It is known [9] that this completion is indeed possible if and only if

$$X - Q = Y^{-1} \tag{59}$$

where Q is a symmetric matrix of size $n \times n$ satisfying

$$Q \geq 0, \qquad \text{rank}(Q) \leq k\,. \tag{60}$$

From (59), in order to reduce the number of complicating variables in (56)-(57), we perform in (56) the congruent transformation:

$$\begin{bmatrix}Y^{-1} & 0\\ 0 & I\end{bmatrix}\,. \tag{61}$$

This yields the equivalent inequality

$$\begin{bmatrix} A^T Y^{-1} + Y^{-1} A & Y^{-1} B \\ B^T Y^{-1} & -I \end{bmatrix} < 0 \,, \tag{62}$$

which by virtue of (59) can be written as

$$\begin{bmatrix} A^T (X-Q) + (X-Q) A & (X-Q) B \\ B^T (X-Q) & -I \end{bmatrix} < 0 \,. \tag{63}$$

Thus, the optimal $k-$th-order filter can be formulated as

$$\min_{X,Q,Z,\nu} \{\nu : \; (10), (57), (58), (60), (63)\} \,, \tag{64}$$

where only (60) is the source of nonconvexity. This difficulty is our main focus hereafter. When the optimal solution of (64) has been found, the optimal k-order filter (3) is easily derived by solving (49), (50) which for a given $\mathcal{X}$ become LMIs with respect to $K = [\, A_F \quad B_F \,]$ and L_F.

Some convex relaxations of (60) are considered first which are based on the following result.

Lemma 4 *A positive semi-definite matrix Q of dimension $n \times n$ has a rank less than $k \leq n$ if it has at least $(n-k)$ zero diagonal entries, i.e. there are indexes $1 \leq i_1 < i_2 < \ldots < i_{n-k} \leq n$ such that $Q_{i_j i_j} = 0, \; j = 1, 2, \ldots, (n-k)$.*

From the above lemma, it follows that for any $1 \leq i_1 < i_2 < \ldots < i_{n-k} \leq n$ an upper bound for (64) is provided by the following (convex) LMI optimization problem

$$(\mathrm{RL}) \quad \min_{X,Q,Z,\nu} \left\{\nu : \; (10), (57), (58), (63), Q \geq 0, \; Q_{i_j i_j} = 0, \; j = 1, 2, \ldots, (n-k)\right\}. \tag{65}$$

Clearly, when either $k = n$ (full-order case) or $k = 0$ (static case) then (64) is equivalent to (65), i.e. (64) becomes a (convex) LMI optimization problem.

3.2 Tailored optimization algorithms

Note that Q satisfies (60) if and only if

$$Q = RR^T \tag{66}$$

with some new slack matrix variable R of dimension $n \times k$. Therefore, (64) can be regarded as an LMI program subject to the additional quadratic constraint (66). In this setting, various optimization techniques, local or global, can be used. See [2, 3, 4, 8, 15, 16] for a sample. For efficiency reasons, it is necessary that these algorithms are specifically tailored to the problem properties and exploit structural informations. This is considered from in the sequel.

3.2.1 Penalty/conditional gradient method

Trivially, (66) is equivalent to

$$\begin{bmatrix} Q & R \\ R^T & I \end{bmatrix} \geq 0 \tag{67}$$

$$\mathrm{Tr}(Q - RR^T) = 0 \tag{68}$$

where (67) is an LMI constraint. Thus, the optimal kth-order filter problem can be recast as

$$\min_{X,Q,Z,\nu,R} \{\nu : \ (10), (57), (58), (63), (67), (68)\}. \tag{69}$$

Note that (67) implies that $\mathrm{Tr}(Q - RR^T) \geq 0$, thus a most natural method to handle the nonconvex constraint (68) is to use a penalty term $\mu\mathrm{Tr}(Q - RR^T)$ combined with the cost function ν. This penalty term prescribes a high cost to the violation of the constraint (68), hence will force this condition if μ is chosen sufficiently large. The original problem (69) is then replaced with:

$$\min \ \left\{\nu + \mu\mathrm{Tr}(Q - RR^T) : \ (10), (57), (58), (63), (67)\right\} \tag{70}$$

It is classically known [6] that the global optimal value of (70) tends to that of (69) as $\mu \to +\infty$. However, increasing the penalty parameter μ renders the problem more and more ill-conditioned, so a standard implementation of the penalty technique follows the following iterative scheme:

1. *select an initial feasible value of the variables, and a penalty parameter* $\mu^0 > 0$*,*
2. *solve the subproblem (70)*
3. *update the penalty parameter*

$$\mu^{\kappa+1} := \begin{cases} \beta\mu^\kappa & \text{if} \quad Tr(Q^\kappa - R^\kappa R^{\kappa T}) > \gamma\, Tr(Q^{\kappa-1} - R^{\kappa-1} R^{\kappa-1 T}) \\ \mu^\kappa & \text{if} \quad Tr(Q^\kappa - R^\kappa R^{\kappa T}) \leq \gamma\, Tr(Q^{\kappa-1} - R^{\kappa-1} R^{\kappa-1 T}) \end{cases} \tag{71}$$

4. *if* $Tr(Q^\kappa - R^\kappa R^{\kappa T})$ *is small enough stop, else go to 2.*

Typical values of the parameters are $\beta = 5$ and $\gamma = 0.25$. Hence, the penalty parameter is increased when the observed violation of the constraint does not show sufficient decrease over the previous minimization.

Note that the subproblem (70) is solved locally using its important concave feature. As the penalized $\nu + \mu\mathrm{Tr}(Q - RR^T)$ is concave [2, 3], its linear approximation at the current iterate R^κ is also its global majorant, i.e.

$$\nu + \mu[\mathrm{Tr}(Q) - \mathrm{Tr}(RR^T)] \leq l_\mu(\nu, Q, R) := \nu + \mu[\mathrm{Tr}(Q) - Tr(R^\kappa R^{\kappa T}) - 2\mathrm{Tr}((R - R^\kappa)R^{\kappa T})],$$

for all ν, Q, R. It should be emphasized that this property is not satisfied in general nonlinear problems.

Thus, the subproblem (70) can be solved by conditional gradient steps which use successive linear approximations of the penalized function according to the sequence of iterates:

$$(\nu^{\kappa+1}, Q^{\kappa+1}, R^{\kappa+1}) := \text{argmin}\,\{l_{\mu^\kappa}(\nu, Q, R) :\ (10), (57), (58), (63), (67)\}\,. \tag{72}$$

A stationary point is obtained when $(\nu^{\kappa+1}, Q^{\kappa+1}, R^{\kappa+1}) = (\nu^\kappa, Q^\kappa, R^\kappa)$, the linear model cannot be decreased further and the inner steps can be stopped. Note also that (72) is readily solved as an LMI program. This provides a feasible descent segment $[R^\kappa, R^{\kappa+1}]$ in the set defined by the LMI constraints. Again, invoking the concavity of the penalized cost, the best next iterate is given by $R^{\kappa+1}$ since the minimum value of a concave function is attained at the extreme points (unit descent step size).

3.2.2 Augmented Lagrangian technique

The advantage of the penalty/conditional gradient method introduced previously lies in the simplicity of its implementation and also in the fact that good upper bounds are usually attained in a few iterations. However, like most first-order methods it may be very slow in the neighborhood of a stationary point. Also, large penalty parameters are a source of ill-conditioning in the conditional gradient scheme. These difficulties can be overpassed by using a more sophisticated augmented Lagrangian/Newton method in which the penalized cost in (70) is replaced with

$$\nu + \text{Tr}(\Lambda(RR^T - Q)) + \mu \text{Tr}\left((RR^T - Q)(RR^T - Q)^T\right)\,, \tag{73}$$

where the Lagrange multipliers Λ and the penalty μ must be updated at each outer iteration. It is also generally recommended to use conditional Newton steps instead of conditional gradient steps in the inner iteration to achieve good rates of convergence. Also, the multiplier update must obey at least a first-order rule. More details on this technique can be found in [8] for robust control problems. The Lagrangian technique, however, is superfluous when the global methods in Section 3.2.3 are implemented.

3.2.3 Branch and bound technique

As mentioned above, the LMI condition (67) implies $\text{Tr}(Q - RR^T) \geq 0$, we can, therefore replace (68) with

$$\text{Tr}(Q - RR^T) = \text{Tr}(Q) - \sum_{i=1}^{n}\sum_{j=1}^{k} R_{i,j}^2 \leq 0 \tag{74}$$

and instead of (69), we consider the equivalent problem

$$\min_{X,Q,Z,\nu,R} \{\nu :\ (10), (57), (58), (63), (67), (74)\}. \tag{75}$$

The difficulty of (75) is concentrated in the nonconvex constraint (74) which nevertheless has the following special structures useful for branching and bounding in Branch and Bound (BB) resolution methods of global optimization.

- As mentioned, the left hand side of (74) is a concave function on Q, R. Hence (74) is actually an inverse convex constraint, i.e. (75) is a convex program with additional inverse convex constraint. Such class of nonconvex problems have been studied intensively in global optimization [12, 17].

- Problem (75) becomes convex when variable R is held fixed, i.e. only R can be regarded as a "complicating variable" causing the problem difficulty [13, 17, 15]. Therefore it is sufficient to perform the branching process in R-space instead of the whole space of all variables (X, Q, Z, ν, R). This alleviates the computational burden.

- From (74), the function $-\mathrm{Tr}(RR^T)$ is separately concave in each variable $R_{i,j}$ [17]. Therefore for every rectangle

$$\mathcal{M} = \{R: \ m_{ij} \leq R_{i,j} \leq M_{ij}, \ i = 1, 2, ..., n; \ j = 1, 2, ..., k\}$$

with given $M_{ij} > m_{ij}$, the best convex relaxation of the inverse convex constraint (74) is [17, Prop. 5.7]

$$\mathrm{Tr}(Q) - \mathrm{Tr}[(\mathcal{M}_{lw} + \mathcal{M}_{up})R^T - \mathcal{M}_{lw}\mathcal{M}_{up}^T] \leq 0, \ \mathcal{M}_{lw} = [m]_{ij}, \ \mathcal{M}_{up} = [M]_{ij}. \quad (76)$$

Accordingly, a good lower bound of the optimal value of (75) with $R \in \mathcal{M}$ is provided by the following LMI optimization problem

$$\beta(\mathcal{M}) = \min_{X,Q,Z,\nu,R} \{\nu : \ (10), (57), (58), (63), (67), (76)\}. \quad (77)$$

- With the optimal solution $(X(\mathcal{M}), Q(\mathcal{M}), Z(\mathcal{M}), \nu(\mathcal{M}), R(\mathcal{M}))$ of (77), an upper bound of the value of (75) can be easily computed by the following LMI program

$$\gamma(\mathcal{M}) = \min_{X,Z,\nu} \left\{\nu : \ (10), (57), (58), \ Q = R(\mathcal{M})R(\mathcal{M})^T\right\}. \quad (78)$$

Based on the above observations, a suitable BB algorithm solving the global optimal solution of (75) can be implemented (see [17, 15] for more details).

4 Illustrative Examples

This section discusses some examples and provide comparison results with earlier techniques both for robust and reduced-order filtering problems..

4.1 Robust filter examples

We consider the following example borrowed from [11, (68)-(70)]

$$\begin{array}{rcl} \dot{x} & = & \begin{bmatrix} 0 & -1+0.3\alpha \\ 1 & -0.5 \end{bmatrix} x + \begin{bmatrix} -2 & 0 \\ 1 & 0 \end{bmatrix} w \\ y & = & [-100+10\beta \quad 100]\, x + [0 \quad 1]\, w \\ z & = & [1 \quad 0]\, x \end{array} \quad (79)$$

with two alternative uncertainty set descriptions, either

$$|\alpha| \leq 1, \quad |\beta| \leq 1, \tag{80}$$

or

$$|\alpha| \leq 1, \quad \alpha = \beta. \tag{81}$$

The comparison between results obtained using Theorem 1 and 2 and those of [11, 14] are provided in Table 1. LMI computations were performed using the Matlab LMI Control Toolbox [10].

method	system	filter order	best upper bound
[11]	(79),(80)	full	5.728
[14]	(79),(80)	full	4.867
Th. 1	(79),(80)	full	2.382
[11]	(79),(81)	full	4.819
[14]	(79),(81)	full	4.373
Th. 1	(79),(81)	full	2.382
[14]	(79),(80)	1	4.946
Th. 2	(79),(80)	1	3.001
[14]	(79),(81)	1	4.556
Th. 2	(79),(81)	1	3.079

Table 1: computational comparisons
robust full- and reduced-order filters

From this Table, the advantage of the proposed method appears clearly. Note that with all α satisfying (81), the asymptotic stability of $A(\alpha)$ in (79) can be checked by a single Lyapunov function $V(x) = x^T X x$. However, if we replace (81) with

$$|\alpha| \leq 3, \; |\beta| \leq 1 \tag{82}$$

then a single Lyapunov function is not satisfactory for checking the asymptotic stability, eventhough $A(\alpha)$ in (79) is asymptotically for all $|\alpha| \leq 3.2$. As a result, the approaches of [11, 14] with parameter-independent Lyapunov functions fails (LMI constraints are infeasible). In contrast, the techniques of Theorems 1 and 2 are still operational in this case. The computational results for problem (79), (82) and also problem (79) with

$$|\alpha| \leq 3, \; \beta = \alpha \tag{83}$$

are sketched in Table 2.

4.2 Reduced-order examples with exact data

Remind that we can use (40) or (65) for relaxing the optimal reduced-order filter problem with exact system data. Our experiments show that the rank relaxation (65) often gives

method	system	filter order	best upper bound
[11] or [14]	(79),(82)	full	$+\infty$
Th. 1	(79),(82)	full	93.365
Th. 2	(79),(82)	1	106.493
[11] or [14]	(79),(83)	full	$+\infty$
Th. 1	(79),(83)	full	100.963
Th. 2	(79),(83)	1	106.517

Table 2: Further computational comparisons
robust full- and reduced-order filters

much less conservative results than (40). Consider the following system borrowed from [18, (5.4)-(5.6)]

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & 1.0 & 0.5 \\ -5.0 & -0.02 & 0 \\ 1.5 & 0 & -0.1 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} w \\ y &= \begin{bmatrix} 1 & 1 & -2 \end{bmatrix} x + \begin{bmatrix} 0 & 1 \end{bmatrix} w \\ z &= \begin{bmatrix} 1 & 1 & -2 \end{bmatrix} x \,. \end{aligned} \tag{84}$$

For the reduced 2nd-order case, the best value given in [18], is $\sqrt{\nu} = 4.74$. After a few iterations, the penalty/conditional gradient method in Subsection 3.2 achieves the much better value $\sqrt{\nu} = 2.4503$ corresponding to

$$Q = \begin{bmatrix} 2.3521 & 0.4489 & -1.4617 \\ 0.4489 & 0.3310 & -0.5034 \\ -1.4617 & -0.5034 & 1.1137 \end{bmatrix} .$$

The later value is very close to the true global optimal value $\sqrt{\nu} = 2.4253$ found by BB method in Subsection 3.2.3 which corresponds to

$$Q = \begin{bmatrix} 2.5425 & 0.4700 & -1.6467 \\ 0.4700 & 0.3211 & -0.5210 \\ -1.6467 & -0.5210 & 1.2668 \end{bmatrix} .$$

Consider a different example from [18, (5.4)-(5.6)]

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & -0.1 & 0 & 0 & 0 \\ 1 & -0.3 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 & 0.016 \\ 0 & -0.3 & 1 & 0 & 0.06 \\ 0 & -0.1 & 0.1 & -1.5 & -0.9 \end{bmatrix} x + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} w \\ y &= \begin{bmatrix} 0.1 & 0 & 0 & -0.5 & 1.6 \\ 0.1 & 0.2 & 0 & -0.3 & 0.12 \end{bmatrix} x + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} w \\ z &= \begin{bmatrix} 0.1 & 0 & 0 & -0.5 & 1.6 \\ 0.1 & 0.2 & 0 & -0.3 & 0.12 \end{bmatrix} x \,. \end{aligned} \tag{85}$$

The best result of [18] gives $\sqrt{\nu} = 2.06$ for reduced 3rd-order filters. The relaxation method of(65) yields the improved value $\sqrt{\nu} = 1.9120$ corresponding to

$$Q = \begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 \\ 0 & 0.1183 & 0 & 0.0563 & 0.0763 \\ 0 & 0 & \mathbf{0} & 0 & 0 \\ 0 & 0.0563 & 0 & 0.6451 & 0.5659 \\ 0 & 0.0763 & 0 & 0.5659 & 2.1510 \end{bmatrix}.$$

Note that this value is almost globally optimal for the nonconvex problem (64) since it is very close to the full-order case, $\sqrt{\nu} = 1.77$.

5 Concluding remarks

In this paper, different techniques and tools for robust and/or reduced-order minimum variance filter problems have been developed. For the synthesis of robust filters, we introduce a new LMI representation which allows the use of parameter-dependent Lyapunov functions while preserving tractability of the problem. This approach generalizes and improves on earlier techniques.

For the reduced-order synthesis, we have introduced more or less conservative relaxations. These relaxed formulations are readily solved as LMI programs but might fail to achieve satisfactory performance levels. In such case, one can either use a penalty/conditional gradient algorithm to get a better local solution or a combination of the penalty/conditional gradient method and the BB method if global optimality is practically required.

References

[1] B.D.O. Anderson, J.B. Moore, *Optimal filtering*, Prentice-Hall, 1979.

[2] P. Apkarian, H.D. Tuan, Robust control via concave minimization: local and global algorithms, *IEEE Trans. Automatic Control* 45(2000), 299-305.

[3] P. Apkarian, H.D. Tuan, Concave programming in control theory, *J. of Global Optimization* 15(1999), 343-370.

[4] P. Apkarian, H.D. Tuan, A sequence SDP Gauss/Newton algorithms for rank LMI constrained problem, *Proc. of CDC 1999, Phoenix*, pp. 2328-2334.

[5] P. Apkarian, H.D. Tuan, J. Bernussou, Flexible LMI representations for analysis and synthesis in continuous-time, to appear in *Proc. of CDC 2000, Sydney.*

[6] D.P. Bertsekas, *Nonlinear programming*, Athena Scientific, USA, Belmont, Mass., 1995.

[7] S. Boyd, L. ElGhaoui, E. Feron, V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, SIAM Studies in Applied Mathematics, Philadelphia, 1994.

[8] B. Fares, P. Apkarian, D. Noll, An augmented Lagrangian method for a class of LMI-constrained problems in robust control theory, to appear in *Int. J. of Control.*

[9] P. Gahinet, P. Apkarian, A linear matrix inequality approach to H_∞ control, *Inter. J. of Nonlinear Robust Control* 4(1994), 421-448.

[10] P. Gahinet, A. Nemirovski, A. Laub, M. Chilali, *LMI control toolbox,* The Math. Works Inc., 1995.

[11] J. C. Geromel, Optimal linear filtering under parameter uncertainty, *IEEE Trans. Signal Processing* 47(1999), 168-175.

[12] R. Horst, H. Tuy, *Global optimization: deterministic approaches* (3rd edition), Springer, 1996.

[13] H. Konno, P.T. Thach, H. Tuy, *Optimization on Low Rank Nonconvex Structures,* Kluwer Academic, 1997.

[14] C.E. de Souza, A. Trofino, An LMI approach to the design of robust $\mathcal{H}_2$ filters, in *Recent Advances on Linear Matrix Inequality Methods in Control*, L. El Ghaoui and S. Niculescu (Eds.), SIAM, 1999.

[15] H.D. Tuan, P. Apkarian, S. Hosoe, H. Tuy, D.C. optimization in robust control: the feasibility problems, *Int. J. of Control* 73(2000), 89-104.

[16] H.D. Tuan, P. Apkarian, Y. Nakashima, A new Lagrangian dual global optimization algorithm for solving bilinear matrix inequalities, *International J. of Nonlinear Robust Controls* 10(2000), 561-578.

[17] H. Tuy, *Convex analysis and global optimization,* Kluwer Academic, 1998.

[18] L. Xie, W. Yan, Y. Soh, L_2 optimal filter reduction: a closed-loop approach, *IEEE Trans. Signal Processing* 46(1998), 11-20.