Numerically reliable methods for optimal design of fault detection filters

A. Varga

Abstract— The design problem of fault detection and isolation filters is formulated as a model matching problem and solved using an $\mathcal{H}_2$- or $\mathcal{H}_\infty$-norm optimization approach. A systematic procedure is proposed to choose appropriate filter specifications which guarantee the existence of proper and stable solutions of the model matching problem. This selection is integral part of numerically reliable computational methods to design of $\mathcal{H}_2$- or $\mathcal{H}_\infty$-optimal fault detection filters. The proposed design approach is completely general, being applicable to both continuous- and discrete-time systems, and can easily handle even unstable and/or improper systems.

I. Introduction

Consider the linear time-invariant system described by the input-output relations

$$\mathbf{y}(\lambda) = G_u(\lambda)\mathbf{u}(\lambda) + G_f(\lambda)\mathbf{f}(\lambda) + G_d(\lambda)\mathbf{d}(\lambda), \quad (1)$$

where $\mathbf{y}(\lambda)$, $\mathbf{u}(\lambda)$, $\mathbf{f}(\lambda)$, and $\mathbf{d}(\lambda)$ are Laplace- or Z-transformed vectors of the p-dimensional system output vector $y(t)$, m_u-dimensional plant input vector $u(t)$, m_f-dimensional fault signal vector $f(t)$, and m_d-dimensional disturbance vector $d(t)$, respectively, and where $G_u(\lambda)$, $G_f(\lambda)$ and $G_d(\lambda)$ are the *transfer-function matrices* (TFMs) from the plant control inputs to outputs, fault signals to outputs, and disturbances to outputs, respectively. According to the system type, $\lambda = s$ in the case of a continuous-time system or $\lambda = z$ in the case of a discrete-time system.

The *fault detection and isolation* (FDI) problem can be formulated as follows: determine a physically realizable (i.e., proper and stable) linear residual generator filter (or fault detector) of least dynamical order having the general form

$$\mathbf{r}(\lambda) = R(\lambda)\begin{bmatrix} \mathbf{y}(\lambda) \\ \mathbf{u}(\lambda) \end{bmatrix} \quad (2)$$

such that: (i) $r_i(t) = 0$ when $f_i(t) = 0$; and (ii) $r_i(t) \neq 0$ when $f_i(t) \neq 0$, for $i = 1, \ldots, m_f$. The simpler *fault detection* (FD) problem requires besides (i) above the simpler condition (ii') $r(t) \neq 0$ when any $f_i(t) \neq 0$, for $i = 1, \ldots, m_f$.

One possibility to determine a least order $R(\lambda)$ which solves the FDI problem is to solve the following model matching problem [1], [2]: choose a suitable $M(\lambda)$ (i.e., stable, proper, diagonal and invertible) and find a least McMillan degree solution $R(\lambda)$ of the linear equation with rational matrices

$$R(\lambda)\begin{bmatrix} G_f(\lambda) & G_d(\lambda) & G_u(\lambda) \\ O & O & I_{m_u} \end{bmatrix} = M(\lambda)\begin{bmatrix} I_{m_f} & O & O \end{bmatrix} \quad (3)$$

A. Varga is with the German Aerospace Center, DLR - Oberpfaffenhofen, Institute of Robotics and Mechatronics, D-82234 Wessling, Germany; Email: andras.varga@dlr.de

which is stable and proper. This equation arises by imposing for the filter (2) the specification that $\mathbf{r}(\lambda) = M(\lambda)\mathbf{f}(\lambda)$ for all $\mathbf{d}(\lambda)$ and $\mathbf{u}(\lambda)$, thus achieving an *exact* decoupling of faults from the disturbance and system inputs. The solution of a FD problem corresponds to filter specification $M(\lambda)$ which is stable, proper and full row rank (e.g., a single row matrix).

If for a properly chosen $M(\lambda)$ the compatibility condition

$$\mathrm{rank}\begin{bmatrix} G_f(\lambda) & G_d(\lambda) \\ M(\lambda) & O \end{bmatrix} = \mathrm{rank}\begin{bmatrix} G_f(\lambda) & G_d(\lambda) \end{bmatrix}$$

is fulfilled (for example, if $\begin{bmatrix} G_f(\lambda) & G_d(\lambda) \end{bmatrix}$ is left invertible), then the rational matrix equation (3) can be solved using, for example, the recently developed numerically reliable approach proposed in [3] based on orthogonal pencil reduction methods.

Unfortunately, in many practical applications the above rank condition is not fulfilled and therefore the equation (3) can not be exactly solved. Instead, we can determine $R(\lambda)$ by solving the $\mathcal{H}_2$- or $\mathcal{H}_\infty$-norm minimization problem

$$\|R(\lambda)\widetilde{G}(\lambda) - M(\lambda)\widetilde{F}(\lambda)\| = \min \quad (4)$$

with

$$\widetilde{G}(\lambda) = \begin{bmatrix} G_f(\lambda) & G_d(\lambda)W_d(\lambda) & G_u(\lambda)W_u(\lambda) \\ O & O & W_u(\lambda) \end{bmatrix}$$

$$\widetilde{F}(\lambda) = \begin{bmatrix} I_{m_f} & O & O \end{bmatrix}$$

where $W_d(\lambda)$ and $W_u(\lambda)$ are optional frequency weightings. For example, by choosing $W_d(\lambda) = \gamma_d I_{m_d}$ and $W_u(\lambda) = \gamma_u I_{m_u}$, with $\gamma_d \gg 1$ and $\gamma_u \gg 1$, we can achieve an increased attenuation of the effects of disturbances and control inputs in the residual signals. Alternatively, by appropriately chosen weights, the attenuation can be achieved only in a certain frequency region of interest.

The $\mathcal{H}_2$- and $\mathcal{H}_\infty$-optimal solution of the FDI problem has been considered by many authors (see [4], [2] and references cited therein). In what follows we discuss shortly the two main approaches used to solve these problems.

The filtering based approaches, pioneered in [5], (see also [6] for recent developments) convert the problem into a standard $\mathcal{H}_2$- or $\mathcal{H}_\infty$-filter synthesis problem to be solved using standard Riccati equation based techniques [7], [8]. The applicability of this approach is conditioned by several technical assumptions, as for example, full row rank of $\widetilde{G}(\lambda)$ and lack of zeros on the extended imaginary axis in the continuous-time or on the unit-circle in the discrete-time. Although these conditions are not necessary for the existence of a solution, the approach still fails when they are not

fulfilled. Since the filtering-based approach provides no clear guidance how to choose appropriate filter specification for successful design, the whole filter design reduces to an *ad-hoc* trial-and-error procedure [9].

The second approach, proposed in [10], is basically a continuous-time $\mathcal{H}_2$-optimal design. It compensates the presence of zeros of $\widetilde{G}(s)$ on the imaginary axis or at infinity by including the same zeros contents in the filter specification $M(s)$. This leads to automatic poles-zeros cancellations when determining the detection filter. However, the computational approach proposed in [10] involves highly sensitive computations like the determination of the Smith-McMillan form of a rational matrix, and thus is not suited as computational procedure for FDI filter design.

In this paper the approach of [11] is extended to solve the underlying model matching problem (3) in an $\mathcal{H}_2$- or $\mathcal{H}_\infty$-optimal sense to obtain a least order stable and proper solution $R(\lambda)$ by choosing a suitable filter specification $M(\lambda)$. For this purpose, we develop methods to compute stable and proper approximate solutions of linear rational equations by adjusting the free term via multiplications with stable and proper factors. The proposed approach relies on the manipulation of rational matrices by using descriptor system representations and is able to address the optimal FDI design problem in the most general setting (i.e., arbitrary rational matrices in the system model (1), no restrictions on poles, zeros or rank of $\widetilde{G}(\lambda)$). Some key computations in the proposed approach are the inner-outer factorization of a rational matrix, the solution of linear rational equations, coprime and spectral factorizations, solution of Nehari's problem. The underlying numerical computations rely on numerically reliable algorithms and are well-suited for robust software implementations.

II. Approximate solution of rational equations

To solve the FDI problem, we can solve the rational equation (3), exploiting the additional freedom of choosing a diagonal $M(\lambda)$ such that the resulting $R(\lambda)$ is proper and stable. Since in general the solution is not unique, we would like to compute a solution of least McMillan degree.

For convenience, we consider the more general *dual* problem to solve a linear rational system of the form

$$G(\lambda)X(\lambda) = F(\lambda)M(\lambda) \tag{5}$$

where $G(\lambda)$ and $F(\lambda)$ are given $l \times m$ and $l \times q$ rational TFMs, respectively. Equation (5) corresponds to the transposed equation (3) with $M(\lambda)$ redefined. To solve (5), we need additionally to choose an $q \times q$ invertible diagonal, stable and proper $M(\lambda)$ such that the resulting $m \times q$ solution $X(\lambda)$ is proper, stable and has the least possible McMillan degree.

The system (5) has a solution provided the rank condition

$$\text{rank}\, G(\lambda) = \text{rank}\, [\, G(\lambda)\ F(\lambda)\,] \tag{6}$$

is fulfilled. When this condition is not fulfilled, we want to compute a stable and proper $X(\lambda)$ which minimizes the $\mathcal{H}_2$- or $\mathcal{H}_\infty$-norm of the residual

$$\mathcal{R}(\lambda) := G(\lambda)X(\lambda) - F(\lambda)M(\lambda)$$

Note that in this setting, there are no restrictions of any kind on $G(\lambda)$ and $F(\lambda)$ (they are arbitrary and can be even polynomial matrices), but we assume that an $M(\lambda)$ can be chosen such that a proper and stable solution $X(\lambda)$ exists and the corresponding norm $\|\mathcal{R}(\lambda)\|$ is finite.

The approach we propose has two main computational stages. The first stage is common to both the $\mathcal{H}_2$- or $\mathcal{H}_\infty$-norm minimization and basically achieves the reduction of the original problem to a simpler one for which, in the second step, either the exact algorithm of [11] is used to solve the $\mathcal{H}_2$-norm minimization problem or an $\mathcal{H}_\infty$-model matching approach can be applied (see [12]).

The main computation in the first stage is the determination of the *quasi* inner-outer of factorization

$$G(\lambda) = G_i(\lambda)G_o(\lambda),$$

where $G_i(\lambda)$ is square and inner and $G_o(\lambda)$ has the form

$$G_o(\lambda) = \left[\begin{array}{c} G_{o,1}(\lambda) \\ O \end{array} \right] \tag{7}$$

with $G_{o,1}(\lambda)$ full row rank. Recall that $G_i(\lambda)$ is *inner* if it has only stable poles and satisfies $G_i^*(\lambda)G_i(\lambda) = I_l$, where $G_i^*(s) := G_i^T(-s)$ in a continuous-time setting and $G_i^*(z) := G_i^T(1/z)$ in a discrete-time setting. The full row rank part $G_{o,1}(\lambda)$ is *quasi* outer, having no zeros in the open-right half complex plane in a continuous-time setting or outside the unit circle in a discrete-time setting. Note that in the standard inner-outer factorization of a stable $G(\lambda)$ without zeros on the extended imaginary axis in a continuous-time setting or on the unit circle in a discrete-time setting (see for example [8]), the full row rank part $G_{o,1}(\lambda)$ is an *outer* TFM (i.e., stable and minimum-phase).

We partition the inner factor column-wise in accordance with the row structure of the factor $G_o(\lambda)$

$$G(\lambda) = [\, G_{i,1}(\lambda) \mid G_{i,2}(\lambda)\,] \left[\begin{array}{c} G_{o,1}(\lambda) \\ O \end{array} \right] \tag{8}$$

It follows that

$$\begin{aligned} \|\mathcal{R}(\lambda)\| &= \|G_o(\lambda)X(\lambda) - G_i^*(\lambda)F(\lambda)M(\lambda)\| \\ &= \left\| \left[\begin{array}{c} G_{o,1}(\lambda) \\ O \end{array} \right] X(\lambda) - \left[\begin{array}{c} G_{i,1}^*(\lambda) \\ G_{i,2}^*(\lambda) \end{array} \right] F(\lambda)M(\lambda) \right\| \\ &= \left\| \left[\begin{array}{c} G_{o,1}(\lambda)X(\lambda) - \widehat{F}_1(\lambda)M(\lambda) \\ -\widehat{F}_2(\lambda)M(\lambda) \end{array} \right] \right\| \end{aligned}$$

where $\widehat{F}_1(\lambda) = G_{i,1}^*(\lambda)F(\lambda)$ and $\widehat{F}_2(\lambda) = G_{i,2}^*(\lambda)F(\lambda)$.

In the next two sections we address the second stage of the proposed approach and give in terms of TFMs high-level algorithms to solve the $\mathcal{H}_2$- and $\mathcal{H}_\infty$-norm minimization problems. In Section V, we discuss numerically reliable state space algorithms for the solution of the key computational problems.

III. Computation of the $\mathcal{H}_2$-solution

The approach to solve the $\mathcal{H}_2$-norm minimization problem (5) extends the exact solution method proposed in [11]. If $X(\lambda)$ is an *exact* solution of the equation

$$G_{o,1}(\lambda)X(\lambda) = \widehat{F}_1(\lambda)M(\lambda) \tag{9}$$

then

$$\|\mathcal{R}(\lambda)\|_2 = \|\widehat{F}_2(\lambda)M(\lambda)\|_2$$

Note that the computed solution $X(\lambda)$ is *exact* for the original linear system provided $\widehat{F}_2(\lambda) = 0$. Since $G_{o,1}(\lambda)$ has full row rank, the corresponding compatibility condition (6) for the equation (9) is fulfilled, and thus the system (9) has a solution which can be made proper and stable by appropriately selecting $M(\lambda)$. The general solution of (9) can be expressed as

$$\widehat{X}(\lambda) = X_0(\lambda) + X_N(\lambda)Y(\lambda), \tag{10}$$

where $X_0(\lambda)$ is a particular solution of (9) and $X_N(\lambda)$ is a rational basis matrix for the right nullspace of $G_{o,1}(\lambda)$. The parametrization (10) of all $\mathcal{H}_2$-optimal solutions allows to determine suitable $Y(\lambda)$ leading to a solution of least McMillan degree. The choice of $M(\lambda)$ must guarantee that $X(\lambda) := \widehat{X}(\lambda)M(\lambda)$ is proper and stable and the residual norm is finite. Therefore, we need to choose $M(\lambda)$ to additionally ensure that $\widehat{\mathcal{R}}(\lambda) := \widehat{F}_2(\lambda)M(\lambda)$ is stable and strictly proper in the continuous-time case, or stable and only proper in the discrete-time case.

IV. COMPUTATION OF THE $\mathcal{H}_\infty$-SOLUTION

To compute the $\mathcal{H}_\infty$-solution we have to solve the two-blocks minimal distance problem

$$\gamma_{opt} = \inf \left\| \begin{bmatrix} Y(\lambda) - \widehat{F}_1(\lambda)M(\lambda) \\ -\widehat{F}_2(\lambda)M(\lambda) \end{bmatrix} \right\|_\infty$$

where we denoted $Y(\lambda) := G_{o,1}(\lambda)X(\lambda)$. In this phase we assume that $M(\lambda)$ has been chosen to ensure that the above infimum exists. This implies that $\widehat{F}_1(\lambda)M(\lambda)$ and $\widehat{F}_2(\lambda)M(\lambda)$ must be proper and have no poles on the imaginary axis in a continuous-time setting or on the unit circle in a discrete-time setting. Note that γ_{opt} can be easily bounded as $\gamma_l \leq \gamma_{opt} \leq \gamma_u$, where

$$\gamma_l = \|\widehat{F}_2(\lambda)M(\lambda)\|_\infty, \quad \gamma_u = \left\| \begin{bmatrix} \widehat{F}_1(\lambda)M(\lambda) \\ \widehat{F}_2(\lambda)M(\lambda) \end{bmatrix} \right\|_\infty.$$

A standard approach to solve the above norm-minimization problem is the well-known γ-iteration [12], which allows to compute suboptimal solutions which are arbitrarily close to the optimal one. For a given $\gamma > \|\widehat{F}_2(\lambda)M(\lambda)\|_\infty$ (e.g., $\gamma = (\gamma_l + \gamma_u)/2$), consider the solution of the suboptimal problem

$$\left\| \begin{bmatrix} Y(\lambda) - \widehat{F}_1(\lambda)M(\lambda) \\ -\widehat{F}_2(\lambda)M(\lambda) \end{bmatrix} \right\|_\infty \leq \gamma \tag{11}$$

First we compute the left spectral factorization (see [8])

$$\gamma^2 I - M^*(\lambda)\widehat{F}_2^*(\lambda)\widehat{F}_2(\lambda)M(\lambda) = W^*(\lambda)W(\lambda) \tag{12}$$

where by construction, $W(\lambda)$ is biproper, stable and minimum-phase. Further, we compute the stable-unstable additive decomposition

$$L_s(\lambda) + L_u(\lambda) = \widehat{F}_1(\lambda)M(\lambda)W^{-1}(\lambda) \tag{13}$$

If $\gamma > \gamma_{opt}$, the two-blocks problem (11) is equivalent to the one-block problem (see [12, Theorem 1, page 106])

$$\left\| \left(Y(\lambda) - \widehat{F}_1(\lambda)M(\lambda) \right) W^{-1}(\lambda) \right\|_\infty \leq 1 \tag{14}$$

and $\gamma_H := \|L_u^*(\lambda)\|_H < 1$ ($\|\cdot\|_H$ denotes the Hankel norm of a stable TFM). In this case we readjust $\gamma_u = \gamma$. Otherwise (i.e., $\gamma_H \geq 1$), we readjust $\gamma_l = \gamma$. Then, for $\gamma = (\gamma_l + \gamma_u)/2$ we redo the factorization (12) and decomposition (13). This process is repeated until $\gamma_u - \gamma_l \leq \varepsilon$ (a given tolerance).

If $\gamma_u \geq \gamma > \gamma_{opt}$, the stable solution of (14) can be expressed as

$$Y(\lambda) = (L_s(\lambda) + Y_s(\lambda))W(\lambda),$$

where, for $1 \geq \gamma_1 \geq \gamma_H$, $Y_s(\lambda)$ is a stable solution of the γ_1-suboptimal Nehari problem

$$\|Y_s(\lambda) - L_u(\lambda)\|_\infty \leq \gamma_1 \tag{15}$$

The $\mathcal{H}_\infty$-solution $X(\lambda)$ is the exact least McMillan degree solution of the linear rational equation

$$G_{o,1}(\lambda)X(\lambda) = Y(\lambda) \tag{16}$$

Since $G_{o,1}(\lambda)$ is only a quasi-outer factor, it can still have zeros on the extended imaginary axis in a continuous-time setting or on the unit circle in a discrete-time setting. In the case when these zeros are not cancelled in the solution, the resulting $X(\lambda)$ can be replaced by $X(\lambda)\widetilde{M}(\lambda)$, where $\widetilde{M}(\lambda)$ is chosen such that $X(\lambda)\widetilde{M}(\lambda)$ is proper and stable, and the norm condition (11) is still fulfilled when replacing $Y(\lambda)$ by $Y(\lambda)\widetilde{M}(\lambda)$. For example, to ensure properness, $\widetilde{M}(\lambda)$ can be chosen diagonal with the diagonal terms of the form

$$\widetilde{M}_i(s) = \frac{1}{(\tau s + 1)^{k_i}} \quad \text{or} \quad \widetilde{M}_i(z) = \frac{1}{z^{k_i}}$$

for continuous- or discrete-time settings, respectively. Note that these factors have unit $\mathcal{H}_\infty$-norm.

V. NUMERICAL ISSUES

The high-level computations in terms of TFMs in the proposed approaches can be performed via state-space models based reliable numerical computations. In what follows we shortly discuss some of these techniques giving more details on the basic computational step common to both $\mathcal{H}_2$ and $\mathcal{H}_\infty$-approaches, namely, the computation of the exact solution of a linear rational equation $G(\lambda)X(\lambda) = F(\lambda)$.

1) Computation of inner-outer factorization: For the computation of the quasi inner-outer factorization in continuous-time, the algorithm developed in [13] can be employed. This algorithm achieves basically a row compression of the underlying $G(s)$ and moves all unstable zeros into symmetric positions with respect to the imaginary axis. In the discrete-time case, a similar algorithm can be employed, with obvious modifications to include the infinite zeros among the unstable ones. In this case, the unstable zeros are reflected with respect to the unit circle. For the determination of the full discrete-time inner factor special formulas (see for example [8]) are available. An implementation of both continuous- and discrete-time algorithms is available in the Descriptor Systems Toolbox for MATLAB [14].

2) Computation of a particular solution $X_0(\lambda)$: Let assume that the compound TFM $[\,G(\lambda)\,F(\lambda)\,]$ has a minimal descriptor realization of the form

$$
\begin{aligned}
E\lambda x(t) &= Ax(t) + B_G u(t) + B_F \nu(t) \\
\xi(t) &= Cx(t) + D_G u(t) + D_F \nu(t)
\end{aligned} \tag{17}
$$

satisfying

$$
[\,G(\lambda)\,F(\lambda)\,] = C(\lambda E - A)^{-1}[\,B_G\ B_F\,] + [\,D_G\ D_F\,] \tag{18}
$$

According to the system type, λ represents here either the differential operator $\lambda x(t) = \dot{x}(t)$ in the case of a continuous-time system or the advance operator $\lambda x(t) = x(t+1)$ in the case of a discrete-time system.

It is easy to see that $X(\lambda) = [\,O\ I_m\,]Y(\lambda)$ is a solution of $G(\lambda)X(\lambda) = F(\lambda)$ if and only if $Y(\lambda)$ satisfies

$$
\begin{bmatrix} A - \lambda E & B_G \\ C & D_G \end{bmatrix} Y(\lambda) = \begin{bmatrix} B_F \\ D_F \end{bmatrix} \tag{19}
$$

To solve (19), we isolate a full rank part of the pencil

$$
S_G(\lambda) := \begin{bmatrix} A - \lambda E & B_G \\ C & D_G \end{bmatrix}
$$

by reducing it to a particular Kronecker-like form. Let Q and Z be orthogonal matrices to reduce $S_G(\lambda)$ to the Kronecker-like form

$$
\overline{S}_G(\lambda) = Q S_G(\lambda) Z = \begin{bmatrix} B_r & A_r - \lambda E_r & A_{r,reg} - \lambda E_{r,reg} \\ 0 & 0 & A_{reg} - \lambda E_{reg} \end{bmatrix},
$$

where $A_{reg} - \lambda E_{reg}$ is a regular subpencil and the pair $(A_r - \lambda E_r, B_r)$ is controllable with E_r nonsingular. The above reduction can be computed by employing numerically stable algorithms as those proposed in [15], [16].

If $\overline{Y}(\lambda)$ is a solution of the reduced equation

$$
\overline{S}_G(\lambda)\overline{Y}(\lambda) = Q \begin{bmatrix} B_F \\ D_F \end{bmatrix} \tag{20}
$$

then $Y(\lambda) = Z\overline{Y}(\lambda)$ and thus

$$
X(\lambda) = \begin{bmatrix} O & I_m \end{bmatrix} Z\overline{Y}(\lambda)
$$

is a solution of the equation $G(\lambda)X(\lambda) = F(\lambda)$. Partition

$$
Q \begin{bmatrix} -B_F \\ -D_F \end{bmatrix} = \begin{bmatrix} \overline{B}_1 \\ \overline{B}_2 \end{bmatrix}
$$

in accordance with the row structure of $\overline{S}_G(\lambda)$.

In general we can determine $\overline{Y}(\lambda)$ of the form

$$
\overline{Y}(\lambda) = \begin{bmatrix} O \\ \overline{Y}_2(\lambda) \\ \overline{Y}_3(\lambda) \end{bmatrix},
$$

where the partitioning of $\overline{Y}(\lambda)$ corresponds to the column partitioning of $\overline{S}_G(\lambda)$. We obtain

$$
\begin{bmatrix} \overline{Y}_2(\lambda) \\ \overline{Y}_3(\lambda) \end{bmatrix} = (\overline{A} - \lambda\overline{E})^{-1} \begin{bmatrix} \overline{B}_1 \\ \overline{B}_2 \end{bmatrix}
$$

where

$$
\overline{A} - \lambda\overline{E} = \begin{bmatrix} A_r - \lambda E_r & A_{r,reg} - \lambda E_{r,reg} \\ 0 & A_{reg} - \lambda E_{reg} \end{bmatrix}
$$

Let partition $[\,O\ I_m\,]Z$ in accordance with the column structure of $S_G(\lambda)$ as

$$
[\,O\ I_m\,]Z = [\,D_r\ C_r\ C_{reg}\,] \tag{21}
$$

and denote

$$
\overline{B} = \begin{bmatrix} \overline{B}_1 \\ \overline{B}_2 \end{bmatrix}, \quad \overline{C} = [\,C_r\ C_{reg}\,]
$$

Then a particular solution $X_0(\lambda)$ of the equation $G(\lambda)X(\lambda) = F(\lambda)$ can be expressed in the form

$$
X_0(\lambda) = \overline{C}(\lambda\overline{E} - \overline{A})^{-1}\overline{B}
$$

This descriptor realization is generally non-minimal, since poles-zeros cancellations can take place in the case $G(\lambda)$ and $F(\lambda)$ share some common zeros. For a non-square $G(\lambda)$, the poles of $X_0(\lambda)$ also contains a set of freely assignable poles (so called "spurious" poles) which originate from the column singularity of $G(\lambda)$. For more details on the pole structure of $X_0(\lambda)$ see [11].

3) Computation of the nullspace basis $X_N(\lambda)$: A right nullspace basis $X_N(\lambda)$ of $G(\lambda)$ can be computed from a right nullspace basis $\overline{Y}_N(\lambda)$ of $\overline{S}_G(\lambda)$ as

$$
X_N(\lambda) = [\,O\ I_m\,]Z\overline{Y}_N(\lambda)
$$

From the Kronecker-like form $\overline{S}_G(\lambda)$, we can determine $\overline{Y}_N(\lambda)$ in the form

$$
\overline{Y}_N(\lambda) = \begin{bmatrix} I \\ (\lambda E_r - A_r)^{-1}B_r \\ O \end{bmatrix}.
$$

With C_r and D_r defined in (21), we obtain a descriptor realization of $X_N(\lambda)$ as

$$
X_N(\lambda) = C_r(\lambda E_r - A_r)^{-1}B_r + D_r.
$$

Note that $X_N(\lambda)$ is a proper TFM which has least McMillan degree [17]. Moreover, the poles of $X_N(\lambda)$ are freely assignable by appropriately choosing the transformation matrices Q and Z to reduce the system pencil $S_G(\lambda)$.

4) Computation of least McMillan degree solution $\widehat{X}(\lambda)$: $X_0(\lambda)$ and $X_N(\lambda)$ can be set up to share the same state and output matrices and have very particular input and feedtrough matrices. To determine a least McMillan degree solution $\widehat{X}(\lambda)$ in the form (10), a suitable $Y(\lambda)$ can be computed by employing the technique proposed in [3] which extends the approach of [18] to possibly non-proper particular solutions $X_0(\lambda)$. The key computational ingredient is the minimal cover algorithm for proper descriptor systems recently proposed in [19].

5) Computation of filter specification $M(\lambda)$: As a last step, usually a diagonal filter specification $M(\lambda)$ is determined such that $X(\lambda) := \widehat{X}(\lambda)M(\lambda)$ is proper and stable. The diagonal structure can be enforced by computing for each column of $\widehat{X}(\lambda)$ a stable and proper right coprime factorization. Suitable algorithms for this purpose have been proposed in [20]. Note that, we can determine $M(\lambda)$ in a factored form $M(\lambda) = M_f(\lambda)M_s(\lambda)$, where $M_f(\lambda)$ is chosen to compensate the finite zero-pole excess in the

solution $\widehat{X}(\lambda)$, while $M_s(\lambda)$ must cancel unstable poles in each columns of $\widehat{X}(\lambda)$. To have finite residuals in the continuous-time case, $M_f(\lambda)$ must be chosen to additionally ensure that $\mathcal{R}(\lambda)M(s)$ is strictly proper. Note that this condition is not required to be fulfilled in the discrete-time case.

6) Left spectral factorization: For a stable and proper $F(\lambda)$, state-space formulas for both continuous- and discrete-time settings are provided in [8] to compute a left spectral factor $W(\lambda)$ satisfying

$$\gamma^2 I - F^*(\lambda)F(\lambda) = W^*(\lambda)W(\lambda)$$

The underlying algorithm relies on solving appropriate Riccati equations. If $F(\lambda)$ is unstable but has no poles on the imaginary axis in continuous-time or on the unit circle in the discrete-time, a preliminary left coprime factorization with inner denominator must be computed as $F(\lambda) = M^{-1}(\lambda)N(\lambda)$, where both $M(\lambda)$ and $N(\lambda)$ are stable, and $M(\lambda)$ is inner (i.e., $M^*(\lambda)M(\lambda) = I$). Then, $W(\lambda)$ can be computed as above using $N(\lambda)$ instead of $G(\lambda)$. Suitable algorithms to compute this factorization are proposed in [20]. Software implementations to compute spectral factorizations are available in the HTOOLS Toolbox for MATLAB [21].

7) Solution of Nehari problem: To solve the suboptimal Nehari problem, a state-space approach for a continuous-time setting has been developed in [22]. For the discrete-time setting, bilinear transformation can be employed by converting the problem into a continuous-time one [22]. Software implementations are available in the HTOOLS Toolbox for MATLAB [21].

VI. NUMERICAL EXAMPLE

The example used in this section is only intended to illustrate the basic algorithms to solve linear rational equations in a *least-squares* sense. All computations have been performed using tools available in the Descriptor Toolbox[1] for MATLAB elaborated by the author [14]. Although state-space representation based computations have been used in all steps, we converted all intermediary results to a nicer TFM form to allow a compact presentation.

Consider the rational system $G(s)X(s) = F(s)M(s)$ with

$$G(s) = \begin{bmatrix} \dfrac{s+2}{(s+1)^2} & \dfrac{1}{s+1} \\ \dfrac{1}{s+1} & \dfrac{1}{s+2} \end{bmatrix}, \quad F(s) = \begin{bmatrix} \dfrac{s}{s+1} \\ \dfrac{s}{s+3} \end{bmatrix}$$

where a proper and stable solution $X(s)$ has to be determined by suitably choosing $M(s)$. The matrix $G(s)$ has rank 1 and the system is not compatible. Therefore we compute solutions which minimize the $\mathcal{H}_2$- and $\mathcal{H}_\infty$-norms of the residual $\mathcal{R}(s) = G(s)X(s) - F(s)M(s)$. First we determine

[1]See http://www.robotic.dlr.de/~varga/num/desctool.html for the contents of current version V1.04c

the inner-outer factorization of $G(s)$ in the form (8), where

$$[\,G_{i,1}(s) \mid G_{i,2}(s)\,] = \frac{\sqrt{2}}{2} \left[\begin{array}{cc|cc} -\dfrac{s+2}{s+\sqrt{5/2}} & \dfrac{-s+1}{s+\sqrt{5/2}} \\ -\dfrac{s+1}{s+\sqrt{5/2}} & \dfrac{s-2}{s+\sqrt{5/2}} \end{array} \right]$$

$$G_{o,1}(s) = \sqrt{2} \left[\begin{array}{cc} -\dfrac{s+\sqrt{5/2}}{(s+1)^2} & -\dfrac{s+\sqrt{5/2}}{(s+1)(s+2)} \end{array} \right]$$

and compute

$$\widehat{F}_1(s) = G_{i,1}^*(s)F(s) = -\sqrt{2}\,\frac{s(s+2.137)(s-1.637)}{(s+3)(s+1)(s-\sqrt{5/2})}$$

$$\widehat{F}_2(s) = G_{i,2}^*(s)F(s) = -\frac{\sqrt{2}}{2}\,\frac{s}{(s+3)(s-\sqrt{5/2})}$$

The $\mathcal{H}_2$-solution: A particular solution $X_0(s)$ of the equation $G_{o,1}(s)X(s) = \widehat{F}_1(\lambda)$ is

$$X_0(s) = \frac{0.5s(s+1)(s+2)(s+2.137)(s-1.637)}{(s-\sqrt{5/2})(s+\sqrt{5/2})(s+1.5)(s+3)} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and a rational right nullspace basis $X_N(s)$ of $G_{o,1}(s)$ is

$$X_N(s) = \frac{\sqrt{2}}{2} \left[\begin{array}{c} \dfrac{s+1}{s+1.5} \\ -\dfrac{s+2}{s+1.5} \end{array} \right]$$

The particular solution $X_0(s)$ has McMillan degree 5 and we can determine $Y(s)$ such that $\widehat{X}(\lambda) = X_0(\lambda) + X_N(\lambda)Y(\lambda)$ has the least possible McMillan degree 4. An appropriate proper $Y(s)$ is

$$Y(s) = -\frac{\sqrt{2}}{4}\,\frac{s(s+2.137)(s-1.637)}{(s-\sqrt{5/2})(s+\sqrt{5/2})(s+3)}$$

for which the corresponding solution of order 4 is

$$\widehat{X}(\lambda) = \left[\begin{array}{c} \dfrac{0.5s(s+1)(s-1.637)(s+2.137)}{(s-\sqrt{5/2})(s+\sqrt{5/2})(s+3)} \\ -\dfrac{0.5s(s+2)(s-1.637)(s+2.137)}{(s-\sqrt{5/2})(s+\sqrt{5/2})(s+3)} \end{array} \right]$$

By choosing

$$M(s) = \frac{s-\sqrt{5/2}}{(s+1)(s+2)}$$

we get a proper and stable solution of order 3

$$X(s) = \left[\begin{array}{c} \dfrac{0.5s(s+2.137)(s-1.637)}{(s+\sqrt{5/2})(s+2)(s+3)} \\ \dfrac{0.5s(s+2.137)(s-1.637)}{(s+\sqrt{5/2})(s+1)(s+3)} \end{array} \right]$$

Interestingly, with the above choice of $M(s)$ we can directly solve $G_{o,1}(s)X(s) = \widehat{F}_1(s)M(s)$ and get an order 2 proper and stable solution

$$X(s) = \left[\begin{array}{c} \dfrac{1.05s^2 - 1.628s + 0.3216}{(s+\sqrt{5/2})(s+3)} \\ \dfrac{-0.05025s^2 + 0.077865s - 0.6433}{(s+\sqrt{5/2})(s+3)} \end{array} \right]$$

The explanation of this fact is that the above choice of $M(s)$ leads to poles-zeros cancellations when forming $\widehat{F}_1(s)M(s)$, thus the resulting particular solution $X_0(s)$ has a lower order too. The $\mathcal{H}_2$-norm of the residual $\mathcal{R}(s) = \widehat{F}_2(s)M(s)$ is in both cases 0.0645.

The $\mathcal{H}_\infty$-solution: With $M(s) = I$, the lower and upper bounds for γ are $\gamma_l = \|\widehat{F}_2(s)\|_\infty = 0.1544$ and $\gamma_u = \|F(s)\|_\infty = 1.4142$. After 10 iterations, we obtain $\gamma = 0.1556$ for which the spectral factor $W(s)$ in (12) is

$$W(s) = \frac{0.15558(s^2 + 0.575s + 4.743)}{(s + 3)(s + \sqrt{5/2})}$$

The stable-unstable additive decomposition (13) gives

$$L_s(s) = \frac{-9.09s^3 - 33.74s^2 - 29.43s + 1.357}{(s + 1)(s^2 + 0.575s + 4.743)}$$

$$L_u(s) = \frac{0.4523}{s - \sqrt{5/2}}$$

The solution of the Nehari problem (15) for $\gamma_1 = 1.01\|L_u(-s)\|_H = 0.1445$ is

$$Y_s(s) = \frac{-22.5}{s + 158.9}$$

Since $Y(s) = (Y_s(s) + L_s(s))W(s)$ is proper, while $G_{o,1}(s)$ is strictly proper (having an infinite zero), the resulting solution of (16) with $\widetilde{M}(s) = I$ will be improper. To obtain a proper solution, we chose $\widetilde{M}(s) = 1/(0.1s + 1)$ for which the least order solution of (16) is

$$X(s) = \begin{bmatrix} \dfrac{5s^4 + 822.6s^3 + 2510s^2 + 1654s - 37.88}{(s + \sqrt{5/2})(s + 10)(s + 3)(s + 158.9)} \\ \dfrac{5s^4 + 827.6s^3 + 3328s^2 + 3347s - 75.77}{(s + \sqrt{5/2})(s + 10)(s + 3)(s + 158.9)} \end{bmatrix}$$

The resulting $\mathcal{H}_\infty$-norm of the residual $R(s)$ is 0.1510. A second order approximation of $X(s)$ obtained by using the singular perturbation approximation approach [23] leads to

$$X_r(s) = \begin{bmatrix} \dfrac{5.088s^2 + 5.944s - 0.1311}{(s + 2.564)(s + 10.17)} \\ \dfrac{5.064s^2 + 11.36s - 0.2621}{(s + 2.564)(s + 10.17)} \end{bmatrix}$$

with a residual norm of 0.1511.

VII. Conclusions

We proposed general approaches to solve $\mathcal{H}_2$- or $\mathcal{H}_\infty$-norm optimal FDI filter design problems. The new approaches reformulate the filter design problems as equivalent model matching problems for which algorithms are proposed able to solve these problems in the most general setting. In this way, the technical difficulties often encountered by the existing methods when trying to reduce the approximation problems to standard $\mathcal{H}_2$- or $\mathcal{H}_\infty$-norm synthesis problems are completely avoided. For example, the presence of zeros or poles on the boundary of stability domains or problems with non-full rank and even improper transfer-function matrices can be easily handled. The underlying main computational algorithms are based on descriptor system representations and rely on orthogonal matrix pencil reductions. For all basic computations, reliable numerical software tools are available for MATLAB in the Descriptor Systems Toolbox [14] and HTOOLS Toolbox [21]. Prototype implementations of the proposed methods are available and will be part of a forthcoming Fault Detection Toolbox for MATLAB.

References

[1] A. Varga, "Computational issues in fault-detection filter design," in *Proc. of CDC'2002, Las Vegas, Nevada*, 2002.

[2] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control.* Springer-Verlag, Berlin, 2003.

[3] A. Varga, "Computation of least order solutions of linear rational equations," in *Proc. of MTNS'04, Leuven, Belgium*, 2004.

[4] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems.* Kluwer Academic Publishers, London, 1999.

[5] A. Edelmayer, J. Bokor, and L. Kevitzky, "An H_∞-filtering approach to robust detection of failures in dynamic systems," in *Proc. of 33rd CDC, Lake Buena Vista, USA*, 1994, pp. 3037–3039.

[6] A. A. Stoorvogel, H. Niemann, A. Saberi, and P. Sannuti, "Optimal fault signal estimation," *Int. J. on Robust and Nonlinear Control*, vol. 12, pp. 697–727, 2002.

[7] Y. S. Hung, "Model-matching approach to H_∞ filtering," *IEE Proc.-D: Contr. Theory & Appl.*, vol. 140, pp. 133–139, 1993.

[8] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control.* Prentice Hall, 1996.

[9] H. Niemann and J. Stoustrup, "Desifgn of fault detectors using H_∞ optimization," in *Proc. of CDC'00, Sydney, Australia*, 2000, pp. 4237–4238.

[10] P. M. Frank and X. Ding, "Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis," *Automatica*, vol. 30, pp. 789–804, 1994.

[11] A. Varga, "New computational approach for the design of fault detection and isolation filters," in *Advances in Automatic Control*, ser. The Kluwer International Series in Engineering and Computer Science, M. Voicu, Ed., vol. 754. Kluwer Academic Publishers, 2003, pp. 367–381.

[12] B. A. Francis, *A Course in H^∞ Theory.* Springer-Verlag, New York, 1987.

[13] C. Oară and A. Varga, "Computation of general inner-outer and spectral factorizations," *IEEE Trans. Automat. Control*, vol. 45, pp. 2307–2325, 2000.

[14] A. Varga, "A DESCRIPTOR SYSTEMS toolbox for MATLAB," in *Proc. CACSD'2000 Symposium, Anchorage, Alaska*, 2000.

[15] ——, "Computation of Kronecker-like forms of a system pencil: Applications, algorithms and software," in *Proc. CACSD'96 Symposium, Dearborn, MI*, 1996, pp. 77–82.

[16] T. Beelen and P. Van Dooren, "An improved algorithm for the computation of Kronecker's canonical form of a singular pencil," *Lin. Alg. & Appl.*, vol. 105, pp. 9–65, 1988.

[17] A. Varga, "On computing least order fault detectors using rational nullspace bases," in *Proc. of IFAC Symp. SAFEPROCESS'2003, Washington D.C.*, 2003.

[18] A. S. Morse, "Minimal solutions to transfer matrix equations," *IEEE Trans. Automat. Control*, vol. 21, pp. 131–133, 1976.

[19] A. Varga, "Reliable algorithms for computing minimal dynamic covers for descriptor systems," in *Proc. of MTNS'04, Leuven, Belgium*, 2004.

[20] ——, "Computation of coprime factorizations of rational matrices," *Lin. Alg. & Appl.*, vol. 271, pp. 83–115, 1998.

[21] A. Varga and V. Ionescu, "HTOOLS - A Toolbox for solving H_∞ and H_2 synthesis problems," in *Proc. of IFAC/IMACS Symp. on Computer Aided Design of Control Systems, Swansea, UK*, July 1991, pp. 508–511.

[22] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds," *Int. J. Control*, vol. 39, pp. 1115–1193, 1984.

[23] Y. Liu and B. D. O. Anderson, "Singular perturbation approximation of balanced systems," *Int. J. Control*, vol. 50, pp. 1379–1405, 1989.