**PHYSCON 2009, Catania, Italy, September, 1–September, 4 2009**

# ANISOTROPY-BASED APPROXIMATION OF LINEAR DISCRETE TIME-INVARIANT STOCHASTIC SYSTEM

**Michael M. Tchaikovsky**
Laboratory 1 ”Dynamics of Control Systems”
Institute of Control Sciences RAS
Moscow, Russia
mmtchaikovsky@hotmail.com

**Alexander P. Kurdyukov**
Laboratory 1 ”Dynamics of Control Systems”
Institute of Control Sciences RAS
Moscow, Russia
akurd@ipu.ru

**Abstract**

This paper represents an approach to approximative model reduction with anisotropic norm of approximation error as performance criterion. The anisotropic norm of a linear discrete time-invariant system is defined as its worst-case sensitivity to a stochastic Gaussian external disturbance with mean anisotropy not exceeding some known value. The mean anisotropy of a vector Gaussian sequence quantifies its temporal colouredness and spatial non-roundness. To solve the main problem, an auxiliary problem of weighted $\mathcal{H}_2$ $\epsilon$-optimal model approximation is stated and solved. Optimality conditions defining a solution to the anisotropy-based optimal approximation problem are expressed in form of a nonlinear matrix algebraic equation system. The presented approach guarantees stability of the obtained reduced-order model without any technical assumption. The reduced-order model approximates steady-state behaviour of the full-order system.

**Key words**

Stochastic disturbance, anisotropy, norm, optimal approximation.

## 1 Introduction

The stochastic approach to $\mathcal{H}_\infty$ optimization introduced in [Vladimirov, Kurdyukov and Semyonov, 1996-1; Vladimirov, Kurdyukov and Semyonov, 1996-2] is based on using the anisotropic norm of a system as performance criterion. The anisotropic norm being a special case of stochastic norm is a quantitative index of system sensitivity to random input disturbances with mean anisotropy bounded by a known nonnegative parameter. In turn, the mean anisotropy of a vector random sequence produced by a stable shaping filter from vector zero-mean Gaussian white noise with identity covariance matrix is a measure of colouredness of this sequence, that is a measure of correlation of vector components of the sequence (spatial part of the mean anisotropy), as well as a measure of correlation of different elements of this sequence (temporal part of the mean anisotropy). The latter coincides with the mutual information about an element of the sequence contained in the past history of this sequence. It has been shown earlier that $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms of a linear discrete time-invariant system are two limiting cases of the anisotropic norm as the mean anisotropy level of input random disturbance tends to zero or infinity, respectively. Therefore, this approach combines the attractive features of robust control and information theories holding an intermediate position between $\mathcal{H}_2$/LQG and $\mathcal{H}_\infty$ problems. The solution to the stochastic $\mathcal{H}_\infty$ problem presented in [Vladimirov, Kurdyukov and Semyonov, 1996-2] yields to the full-order controller, whereas it would be desirable to obtain a reduced-order one.

This paper represents an approach to approximative model reduction using anisotropic norm of approximation error as performance criterion. This result is assumed to be applied for designing the reduced-order anisotropic controller.

## 2 Problem Statement

Let us consider a linear discrete time-invariant system $F_n \in \mathcal{H}_\infty^{p\times m}$ with $n$-dimensional internal state $X$, $m$-dimensional input $W$, and $p$-dimensional output $Y = F_n W$. We assume that all these signals are infinite double-sided discrete-time sequences related to each other by the equations

$$F_n(z):\ \begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \begin{bmatrix} x_k \\ w_k \end{bmatrix}, \qquad (1)$$

where the constant matrices $A, B, C, D$ have appropriate dimensions, and the matrix $A$ is stable in discrete-time sense (i.e. its spectral radius $\rho(A) < 1$). The only prior information on the probability distribution of the random input sequence $W = (w_k)_{-\infty\leqslant k\leqslant+\infty}$ is assumed to be that $W$ is a stationary Gaussian sequence of random vectors $w_k$ with zero mean $\mathbf{E}\,(w_k) = 0$, unknown covariance matrix $\mathbf{E}\,(w_k w_k^{\mathrm{T}}) = \Sigma_W$, and Gaussian distribution density

$$p(w_k) = \frac{1}{\sqrt{(2\pi)^m \det \Sigma_W}} \exp\left(-\frac{1}{2}\|w_k\|^2_{\Sigma_W^{-1}}\right),$$

where $\|w_k\|_{\Sigma_W^{-1}} = \sqrt{w_k^{\mathrm{T}}\Sigma_W^{-1}w_k}$. At that it is supposed that the mean anisotropy of the sequence $W$ is upper-bounded by a known nonnegative parameter $\alpha$. This means that $W$ is produced from $m$-dimensional Gaussian white noise $V = (v_k)_{-\infty\leqslant k\leqslant+\infty}$ with zero mean $\mathbf{E}\,(v_k) = 0$ and scalar covariance matrix $\mathbf{E}\,(v_k v_k^{\mathrm{T}}) = \lambda I_m$, $\lambda \in \mathbb{R}^+$, by an unknown shaping filter $G$ belonging to the family

$$\mathcal{G}_\alpha \triangleq \left\{G \in \mathcal{H}_2^{m\times m}\colon\ \overline{\mathbf{A}}(G) \leqslant \alpha\right\}, \qquad (2)$$

where

$$\overline{\mathbf{A}}(G) = -\frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\det\left\{\frac{m}{\|G\|_2^2}\,\widehat{G}(\omega)\widehat{G}^*(\omega)\right\} d\omega \qquad (3)$$

is the mean anisotropy functional introduced in [Vladimirov, Kurdyukov and Semyonov, 1996-1], the angular boundary value $\widehat{G}(\omega) \triangleq \lim_{r\to 1-0} G(re^{i\omega})$.

We are interested in finding an asymptotically stable reduced-order system $F_r \in \mathcal{H}_\infty^{p\times m}$ with $r$-dimensional internal state $\Xi$ $(r < n)$, $m$-dimensional input $W$, and $p$-dimensional output $\Psi = F_r W$, related by the equations

$$F_r(z):\ \left[\begin{array}{c}\xi_{k+1}\\ \hline \psi_k\end{array}\right] = \left[\begin{array}{c|c}A_r & B_r\\ \hline C_r & D_r\end{array}\right]\left[\begin{array}{c}\xi_k\\ \hline w_k\end{array}\right] \qquad (4)$$

with $(A_r, B_r, C_r, D_r) \in \mathcal{S} \triangleq \{(A_r, B_r, C_r, D_r)\colon \rho(A_r) < 1\}$. Such the reduced-order model is called admissible one.

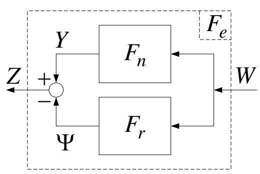


Figure 1. Block diagram of approximation error model $F_e$

Let us choose a model approximation performance criterion as the $\alpha$-anisotropic norm [Vladimirov, Kurdyukov and Semyonov, 1996-1] of difference between the transfer functions of systems (1) and (4):

$$J_\alpha(A_r, B_r, C_r, D_r) = \|\!|F_n - F_r|\!\|_\alpha \doteq \sup_{G\in\mathcal{G}_\alpha} \frac{\|(F_n - F_r)G\|_2}{\|G\|_2}, \qquad (5)$$

where $F_n(z) = C(zI - A)^{-1}B + D$, $F_r(z) = C_r(zI - A_r)^{-1}B_r + D_r$. Denote by $F_e(z) \triangleq F_n(z) - F_r(z)$ the transfer function of approximation error model with output $Z \triangleq Y - \Psi$. This transfer function is defined by the realization

$$F_e(z) \sim \left[\begin{array}{c|c}A_e & B_e\\ \hline C_e & D_e\end{array}\right] \triangleq \left[\begin{array}{cc|c}A & 0 & B\\ 0 & A_r & B_r\\ \hline C & -C_r & D - D_r\end{array}\right]. \qquad (6)$$

Block diagram of the approximation error model $F_e$ is represented at Fig. 1. Now the problem of anisotropy-based approximation of linear discrete time-invariant system can be stated as follows.

**Problem 1.** *Find the matrices* $(A_r, B_r, C_r, D_r)$ *from the admissible set* $\mathcal{S} \triangleq \{(A_r, B_r, C_r, D_r)\colon \rho(A_r) < 1\}$ *such that*

$$J_\alpha = \|\!|F_e|\!\|_\alpha \to \inf_{(A_r,B_r,C_r,D_r)\in\mathcal{S}}. \qquad (7)$$

## 3 Problem Solution

### 3.1 Saddle-Point Type Condition of Optimality

Taking into account definition (5) of the anisotropic norm, Problem 1 is a minimax problem. It gives the opportunity of applying the results of differential game theory to formulate a saddle-point type condition of optimality. In the considered case, the saddle point is a pair $(F_r^\star, G^\star)$ such that the inequalities

$$J_1(F_r^\star, G) \leqslant J_1(F_r^\star, G^\star) \leqslant J_1(F_r, G^\star) \qquad (8)$$

hold true for any admissible $F_r$ and $G$, where

$$J_1(F_r, G) \triangleq \|(F_n - F_r)G\|_2. \qquad (9)$$

Define the sets

$$\mathcal{S}^\diamond(G) \triangleq \operatorname*{Arg\,min}_{(A_r,B_r,C_r,D_r)\in\mathcal{S}} \|(F_n - F_r)G\|_2,\ G \in \mathcal{G}_\alpha, \qquad (10)$$

$$\mathcal{G}_\alpha^\diamond(F_r) \triangleq \operatorname*{Arg\,max}_{G\in\mathcal{G}_\alpha} \|(F_n - F_r)G\|_2/\|G\|_2,\ F_r \in \mathcal{S}. \qquad (11)$$

These sets are assumed to be nonempty. Set (10) consists of the admissible systems of order $r$ being the solutions to the following problem of weighted $\mathcal{H}_2$ optimal approximation

$$J_1(F_r, G) \to \inf_{(A_r,B_r,C_r,D_r)\in\mathcal{S}}, \quad G \in \mathcal{G}_\alpha,$$

under the assumption that the input sequence $W$ of error model (6) is produced by a known shaping filter

$G \in \mathcal{G}_\alpha$, i.e. $W = GV$. Set (11) is formed by the stable filters $G$ generating multidimensional Gaussian random sequences with spectral densities which are the worst (i.e. the most adverse) for the approximation error model $F_e$. Although the set $\mathcal{G}^\diamond_\alpha(F_r)$ is invariant under right multiplication by an all-pass system, all of them generate sequences with the unique spectral density (up to a scalar multiplier) [Vladimirov, Kurdyukov and Semyonov, 1996-1]. Such the filters are called the worst-case shaping filters. Thus, the relation

$$(\mathcal{S}^\diamond_\alpha \circ \mathcal{G}^\diamond_\alpha)(F_r) \triangleq \bigcup\nolimits_{G\in\mathcal{G}^\diamond_\alpha(F_r)} \mathcal{S}^\diamond_\alpha(G), \quad F_r \in \mathcal{S}$$

defines generally set-valued composition $\mathcal{S}^\diamond_\alpha \circ \mathcal{G}^\diamond_\alpha \colon \mathcal{S} \to 2^{\mathcal{S}}$ of the mappings $\mathcal{S}^\diamond_\alpha \colon \mathcal{G}_\alpha \to 2^{\mathcal{S}}$ and $\mathcal{G}^\diamond_\alpha \colon \mathcal{S} \to 2^{\mathcal{G}_\alpha}$.

**Lemma 1.** *If a realization $(A_r, B_r, C_r, D_r)$ of a reduced-order model $F_r$ is a stationary point of the mapping $\mathcal{S}^\diamond_\alpha \circ \mathcal{G}^\diamond_\alpha$, i.e. if there exists a shaping filter $G$ such that*

$$F_r = (A_r, B_r, C_r, D_r) \in \mathcal{S}^\diamond_\alpha(G),\ G \in \mathcal{G}^\diamond_\alpha(F_r), \quad (12)$$

*then this system is a solution to Problem 1.*

Proof of this result is omitted due to lack of space and will be introduced elsewhere.

### 3.2 Worst-Case Shaping Filter for Error Model

Let us further suppose that the approximation error model $F_e(z)$ satisfies the strict inequality

$$m^{-1/2}\|F_e\|_2 < \|F_e\|_\infty. \quad (13)$$

Otherwise, the anisotropic norm of the system $F_e(z)$ trivially coincides with its scaled $\mathcal{H}_2$ norm. Note that inequality (13) does not hold iff the system $F_e(z)$ is inner [Zhou and Doyle, 1998] up to a constant nonzero multiplier. In the last case there exists a number $\lambda > 0$ such that $\widehat{F}^*_e(\omega)\widehat{F}_e(\omega) = \lambda I_m$ for almost all $\omega \in [-\pi, \pi]$. For nonzero system $F_e \in \mathcal{H}^{p\times m}_\infty$ inequality (13) holds true if the inequality $p < m$ does [Vladimirov, Kurdyukov and Semyonov, 1996-1].

**Lemma 2.** *Let the system $F_n(z)$ given by (1) be asymptotically stable, and let the system $F_e(z) = F_n(z) - F_r(z)$ given by (6) not be inner. Then for any admissible reduced-order model $F_r(z) = (A_r, B_r, C_r, D_r) \in \mathcal{S}$ and mean anisotropy level $\alpha \geqslant 0$ of the input sequence there exists a unique pair $(q, R)$ of the scalar parameter $q \in [0, \|F_e\|^{-2}_\infty)$ and stabilizing solution $R$ of the algebraic Riccati equation*

$$\left.\begin{array}{l} R = A_e^{\mathrm{T}} R A_e + qC_e^{\mathrm{T}} C_e + L^{\mathrm{T}}\Sigma^{-1}L \\ L = [\,L_1\ L_2\,] \triangleq \Sigma(B_e^{\mathrm{T}} R A_e + qD_e^{\mathrm{T}} C_e) \\ \Sigma \triangleq (I_m - B_e^{\mathrm{T}} R B_e - qD_e^{\mathrm{T}} D_e)^{-1} \end{array}\right\}, \quad (14)$$

*such that*

$$\alpha = -\frac{1}{2}\ln\det\frac{m\Sigma}{\mathrm{tr}(LP_cL^{\mathrm{T}} + \Sigma)}, \quad (15)$$

*where $P_c = P_c^{\mathrm{T}} > 0$ is the controllability gramian of the shaping filter*

$$G(z) \sim \left[\begin{array}{cc|c} A + BL_1 & BL_2 & B\Sigma^{1/2} \\ B_rL_1 & A_r + B_rL_2 & B_r\Sigma^{1/2} \\ \hline L_1 & L_2 & \Sigma^{1/2} \end{array}\right] \quad (16)$$

*satisfying the Lyapunov equation*

$$P_c = (A_e + B_eL)P_c(A_e + B_eL)^{\mathrm{T}} + B_e\Sigma B_e^{\mathrm{T}}. \quad (17)$$

*At that, filter* (16) *is a representative of family* (2) *of the worst-case shaping filters.*

Proof of this lemma follows immediately from Theorem 2 in [Vladimirov, Kurdyukov and Semyonov, 1996-1] applied to error model (6).

**Remark 1.** *Recall that a solution $R = R^{\mathrm{T}} \in \mathbb{R}^{(n+r)\times(n+r)}$ of algebraic Riccati equation* (14) *is called stabilizing if the matrix $A_e + B_eL$ is stable ($\rho(A_e + B_eL) < 1$) and the matrix $\Sigma = \Sigma^{\mathrm{T}} > 0$. For any $F_r = (A_r, B_r, C_r, D_r) \in \mathcal{S}$ and $q \in [0, \|F_e\|^2_\infty)$ equation* (14) *has a unique positive-definite stabilizing solution.*

**Remark 2.** *The internal state of the worst-case filter $G$ actually is a copy of that of the error model $F_e$. Thus, realization* (6) *combined with $w_k = L_1x_k + L_2\xi_k + \Sigma^{1/2}v_k$ relates the input $V$, output $W = GV$, and internal state $(X, \Xi)$ of worst-case filter* (16).

### 3.3 Weighted $\mathcal{H}_2$ Optimal Model Approximation

For the known worst-case shaping filter $G \in \mathcal{G}^\diamond_\alpha(F_r)$ defined by realization (16) and equations (14)–(15), the anisotropy-based model approximation problem (7) is equivalent to the problem of weighted $\mathcal{H}_2$ optimal approximation

$$J_1(F_r, G) = \|(F_n - F_r)G\|_2 \to \inf_{(A_r,B_r,C_r,D_r)\in\mathcal{S}}, \quad G \in \mathcal{G}^\diamond_\alpha(F_r). \quad (18)$$

Let us consider the realization of the weighted approximation error model

$$\widetilde{F}_e(z) = F_n(z)G(z) - F_r(z)G(z). \quad (19)$$

The transfer function of the weighted error model

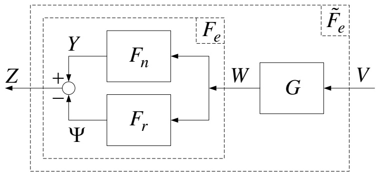


Figure 2. Block diagram of weighted approximation error model

$\widetilde{F}_e(z)$ in (19) is given by the state-space realization

$$\widetilde{F}_e(z) \sim \left[\begin{array}{c|c} \widetilde{A}_e & \widetilde{B}_e \\ \hline \widetilde{C}_e & \widetilde{D}_e \end{array}\right]$$

$$= \left[\begin{array}{cc|c} A + BL_1 & BL_2 & B\Sigma^{1/2} \\ B_r L_1 & A_r + B_r L_2 & B_r \Sigma^{1/2} \\ \hline (D - D_r)L_1 + C & (D - D_r)L_2 - C_r & (D - D_r)\Sigma^{1/2} \end{array}\right]. \tag{20}$$

Block diagram of the weighted error model is given at Fig. 2.

**Lemma 3.** *For any matrices* $(A_r, B_r, C_r) \in \mathbb{S}$, *the infimum of functional* (18) *is reached with* $D_r \equiv D$.

Proof is omitted due to lack of space.

From Lemma 3 it follows that we can assume $D_r = D = 0$ in systems (1) and (4) without prejudice to generality of problem statement (18).

Let us denote

$$\widetilde{F}_n(z) \triangleq F_n(z)G(z) \sim \left[\begin{array}{c|c} \widetilde{A} & \widetilde{B} \\ \hline \widetilde{C} & 0 \end{array}\right] \tag{21}$$

and define the reduced-order model of the system $\widetilde{F}_n(z)$ as

$$\widetilde{F}_r(z) \triangleq F_r(z)G(z) \sim \left[\begin{array}{c|c} \widetilde{A}_r & \widetilde{B}_r \\ \hline \widetilde{C}_r & 0 \end{array}\right]. \tag{22}$$

Now consider the auxiliary performance criterion for $\mathcal{H}_2$ model approximation

$$J_2(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) = \|\widetilde{F}_e\|_2^2. \tag{23}$$

From the representation

$$J_2(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) = \operatorname{tr}(\widetilde{C}_e P_c \widetilde{C}_e^{\mathrm{T}}),\ (\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}, \tag{24}$$

it can be easily seen that

$$\inf_{(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}} J_2 = \left\{ \inf_{(A_r, B_r, C_r) \in \mathbb{S}} J_1 \right\}^2.$$

Denote by $J_2^\star$ the infimum of $J_2$ over $\widetilde{\mathbb{S}}$ :

$$J_2^\star \triangleq \inf\{J_2(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \colon (\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}\},$$

$$\widetilde{\mathbb{S}} \triangleq \{(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) : \rho(\widetilde{A}_r) < 1\}. \tag{25}$$

It is obvious that for any $\epsilon > 0$ there exist some matrices $(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}$ such that

$$0 \leqslant J_2(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) - J_2^\star < \epsilon.$$

Further, instead of optimal weighted $\mathcal{H}_2$ approximation problem (18), we will consider the following $\epsilon$-optimal problem.

**Problem 2.** *For given* $\epsilon > 0$, *find the reduced-order realization* $\widetilde{F}_r = (\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}$ *such that*

$$|J_2(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) - J_2^\star| < \epsilon. \tag{26}$$

**Remark 3.** *The solution to Problem 2 for continuous-time case was introduced in* [Huang and Teo, 2001]. *The further results of this paper are mainly similar to ideas of* [Huang and Teo, 2001] *extended to the discrete-time case.*

### 3.4 $\mathcal{H}_2$ $\epsilon$-Optimal Approximation of Linear Discrete Time-Invariant System

Let us begin on solving $\mathcal{H}_2$ $\epsilon$-optimal approximation problem (26) with introducing an auxiliary functional

$$J_\beta(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) = J(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) + \operatorname{tr}(P)/\beta, \quad (\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}, \tag{27}$$

where $\beta \in \mathbb{R}^+$, $P$ is a unique positive definite solution to the Lyapunov equation

$$P = \widetilde{A}_r P \widetilde{A}_r^{\mathrm{T}} + \widetilde{B}_r \widetilde{B}_r^{\mathrm{T}} + \widetilde{A}_r^{\mathrm{T}} \widetilde{A}_r + \widetilde{C}_r^{\mathrm{T}} \widetilde{C}_r + I_r, \tag{28}$$

where $I_r \in \mathbb{R}^{r \times r}$ is the identity matrix. The following lemma guarantees existence of global minimum of auxiliary functional (27) over the set of the reduced-order admissible realizations.

**Lemma 4.** *For any* $\beta > 0$, *the auxiliary functional* $J_\beta$ *defined by* (27) *has a global minimum over the set* $\widetilde{\mathbb{S}}$.

Proof of this lemma is omitted due to lack of space and will be introduced elsewhere.

The following lemma shows that the minimum of the auxiliary functional $J_\beta$ over $\widetilde{\mathbb{S}}$ tends to the infimum of the auxiliary $\mathcal{H}_2$ performance criterion $J_2$ over $\widetilde{\mathbb{S}}$ as $\beta \to +\infty$. Moreover, the global minimum point of the functional $J_\beta$ is the solution to $\epsilon$-optimal Problem 2 for some sufficiently large $\beta$.

**Lemma 5.** *Consider the* $\mathcal{H}_2$ *performance criterion* $J_2 = \|\widetilde{F}_e\|_2^2$ *of model approximation and the auxiliary functional* $J_\beta$ *defined by expression* (27). *Let*

$(\widetilde{A}_r^\star, \widetilde{B}_r^\star, \widetilde{C}_r^\star)$ *be a global minimum point of the functional* $J_\beta$. *Then*

$$\lim_{\beta\to+\infty} J_\beta(\widetilde{A}_r^\star, \widetilde{B}_r^\star, \widetilde{C}_r^\star) = \lim_{\beta\to+\infty} J_2(\widetilde{A}_r^\star, \widetilde{B}_r^\star, \widetilde{C}_r^\star) = J_2^\star. \quad (29)$$

Proof of this result is omitted for brevity and will be introduced elsewhere.

**Remark 4.** *The second equality in expression* (29) *means that for any* $\epsilon > 0$ *there exists* $\beta^\star > 0$ *such that any global minimum point* $(\widetilde{A}_r^\star, \widetilde{B}_r^\star, \widetilde{C}_r^\star)$ *of the functional* $J_\beta$ *over* $\widetilde{\mathbb{S}}$ *is a solution to auxiliary* $\epsilon$*-optimal problem* (26) *on condition that* $\beta > \beta^\star$.

**Remark 5.** *From Lemma 5 it follows that solving auxiliary problem of* $\mathcal{H}_2$ $\epsilon$*-optimal model approximation* (26) *reduces to finding the global minimum of the functional* $J_\beta$ *for sufficiently large* $\beta$.

**Remark 6.** *The problem of finding the minimum of* $J_\beta$ *is the smooth constrained nonlinear optimization problem because of the stability constraint on the matrix* $\widetilde{A}_r$. *The solution set of this optimization problem is an open set. Any global minimum point of* $J_\beta$ *is a critical point, at which the gradient of the functional* $J_\beta$ *is equal to zero.*

Let us obtain the expression for the gradient of $J_\beta$. For a smooth transformation $F$ defined on $\mathbb{R}^{r\times r}\times\mathbb{R}^{r\times m}\times\mathbb{R}^{p\times r}$, let $dF(\theta)$ denotes its Frechet derivative in the direction $\theta = [\,\theta_1\ \theta_2\ \theta_3\,] \in \mathbb{R}^{r\times r}\times\mathbb{R}^{r\times m}\times\mathbb{R}^{p\times r}$ at a point $(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \mathbb{R}^{r\times r}\times\mathbb{R}^{r\times m}\times\mathbb{R}^{p\times r}$. Then we have

$$dJ_\beta(\theta) = dJ_2(\theta) + \operatorname{tr}(dP(\theta))/\beta. \quad (30)$$

Let us obtain the expression for $dJ_2(\theta)$ differentiating (24):

$$\begin{aligned} dJ_2(\theta) &= \operatorname{tr}(2d\widetilde{C}_e(\theta)P_c\widetilde{C}_e^{\mathrm{T}} + \widetilde{C}_e dP_c(\theta)\widetilde{C}_e^{\mathrm{T}}) = \\ &= \operatorname{tr}(-2P_c\widetilde{C}_e^{\mathrm{T}}[\,0\ \theta_3\,]) + \operatorname{tr}(\widetilde{C}_e^{\mathrm{T}}\widetilde{C}_e dP_c(\theta)). \end{aligned} \quad (31)$$

To find $dP_c(\theta)$, let us differentiate Lyapunov equation (17) and obtain

$$\begin{aligned} &\widetilde{A}_e dP_c(\theta)\widetilde{A}_e^{\mathrm{T}} - dP_c(\theta) + d\widetilde{A}_e(\theta)P_c\widetilde{A}_e^{\mathrm{T}} + \widetilde{A}_e P_c d\widetilde{A}_e^{\mathrm{T}}(\theta) \\ &+d\widetilde{B}_e(\theta)\widetilde{B}_e^{\mathrm{T}} + \widetilde{B}_e d\widetilde{B}_e^{\mathrm{T}}(\theta) = \widetilde{A}_e dP_c(\theta)\widetilde{A}_e^{\mathrm{T}} - dP_c(\theta) \\ &+ \begin{bmatrix} 0 & 0 \\ 0 & \theta_1 \end{bmatrix} P_c\widetilde{A}_e^{\mathrm{T}} + \widetilde{A}_e P_c \begin{bmatrix} 0 & 0 \\ 0 & \theta_1^{\mathrm{T}} \end{bmatrix} + \begin{bmatrix} 0 \\ \theta_2 \end{bmatrix} \widetilde{B}_e^{\mathrm{T}} \\ &+\widetilde{B}_e[\,0\ \theta_2^{\mathrm{T}}\,] = \widetilde{A}_e dP_c(\theta)\widetilde{A}_e^{\mathrm{T}} - dP_c(\theta) + XP_c\widetilde{A}_e^{\mathrm{T}} \\ &+\widetilde{A}_e P_c X^{\mathrm{T}} + Y\widetilde{B}_e^{\mathrm{T}} + \widetilde{B}_e Y^{\mathrm{T}} = 0, \end{aligned} \quad (32)$$

where

$$X \triangleq d\widetilde{A}_e(\theta) = \begin{bmatrix} 0 & 0 \\ 0 & \theta_1 \end{bmatrix},\ Y \triangleq d\widetilde{B}_e(\theta) = \begin{bmatrix} 0 \\ \theta_2 \end{bmatrix}. \quad (33)$$

Consider the observability gramian $P_o$ of weighted error model (20), satisfying the Lyapunov equation

$$P_o = \widetilde{A}_e^{\mathrm{T}} P_o \widetilde{A}_e + \widetilde{C}_e^{\mathrm{T}}\widetilde{C}_e. \quad (34)$$

From (32) and (34) it follows that

$$\begin{aligned} &\operatorname{tr}(\widetilde{C}_e^{\mathrm{T}}\widetilde{C}_e dP_c(\theta)) = 2\operatorname{tr}(P_c\widetilde{A}_e^{\mathrm{T}}\widetilde{A}_e^{\mathrm{T}}P_o\widetilde{A}_e X) \\ &-2\operatorname{tr}(P_c\widetilde{A}_e^{\mathrm{T}}P_o X) + 2\operatorname{tr}(\widetilde{B}_e^{\mathrm{T}}\widetilde{A}_e^{\mathrm{T}}P_o\widetilde{A}_e Y) \\ &-2\operatorname{tr}(\widetilde{B}_e^{\mathrm{T}}P_o Y). \end{aligned} \quad (35)$$

Taking into account (33), we have

$$\begin{aligned} &2\operatorname{tr}(P_c\widetilde{A}_e^{\mathrm{T}}\widetilde{A}_e^{\mathrm{T}}P_o\widetilde{A}_e X) \\ &\quad = 2\operatorname{tr}([\,0\ I_r\,]P_c\widetilde{A}_e^{\mathrm{T}}\widetilde{A}_e^{\mathrm{T}}P_o[\,0\ I_r\,]^{\mathrm{T}}\widetilde{A}_e\theta_1), \\ &2\operatorname{tr}(\widetilde{B}_e^{\mathrm{T}}\widetilde{A}_e^{\mathrm{T}}P_o\widetilde{A}_e Y) \\ &\quad = 2\operatorname{tr}(\widetilde{B}_e^{\mathrm{T}}\widetilde{A}_e^{\mathrm{T}}P_o[\,0\ I_m\,]^{\mathrm{T}}\widetilde{A}_e\theta_2), \\ &-2\operatorname{tr}(P_c\widetilde{A}_e^{\mathrm{T}}P_o X) \\ &\quad = -2\operatorname{tr}([\,0\ I_r\,]P_c\widetilde{A}_e^{\mathrm{T}}P_o[\,0\ I_r\,]^{\mathrm{T}}\theta_1), \\ &-2\operatorname{tr}(\widetilde{B}_e^{\mathrm{T}}P_o Y) = -2\operatorname{tr}(\widetilde{B}_e^{\mathrm{T}}P_o[\,0\ I_m\,]^{\mathrm{T}}\theta_2). \end{aligned} \quad (36)$$

To find the derivative $dP(\theta)$, let us differentiate auxiliary Lyapunov equation (28) and obtain

$$\begin{aligned} dP(\theta) &= \widetilde{A}_r dP(\theta)\widetilde{A}_r^{\mathrm{T}} + \theta_1 P\widetilde{A}_r^{\mathrm{T}} + \widetilde{A}_r P\theta_1^{\mathrm{T}} + \widetilde{A}_r^{\mathrm{T}}\theta_1 \\ &+ \theta_1^{\mathrm{T}}\widetilde{A}_r + \widetilde{B}_r\theta_2^{\mathrm{T}} + \theta_2\widetilde{B}_r^{\mathrm{T}} + \widetilde{C}_r^{\mathrm{T}}\theta_3 + \theta_3^{\mathrm{T}}\widetilde{C}_r. \end{aligned} \quad (37)$$

Let the matrix $Q = Q^{\mathrm{T}} > 0$ be a unique positive definite solution to the Lyapunov equation

$$Q = \widetilde{A}_r^{\mathrm{T}} Q\widetilde{A}_r + I_r/\beta. \quad (38)$$

Taking into account (38), from (37) it follows that

$$\begin{aligned} &\operatorname{tr}((I_r - \widetilde{A}_r^{\mathrm{T}}\widetilde{A}_r)dP(\theta)) \\ &= 2\operatorname{tr}((P + I_r)\widetilde{A}_r^{\mathrm{T}}\theta_1) + 2\operatorname{tr}(\widetilde{B}_r^{\mathrm{T}}\theta_2) + 2\operatorname{tr}(\widetilde{C}_r^{\mathrm{T}}\theta_3), \end{aligned}$$

from which

$$\begin{aligned} \operatorname{tr} dP(\theta)/\beta &= 2\operatorname{tr}(Q(P + I_r)\widetilde{A}_r^{\mathrm{T}}\theta_1) \\ &+ 2\operatorname{tr}(Q\widetilde{B}_r^{\mathrm{T}}\theta_2) + 2\operatorname{tr}(Q\widetilde{C}_r^{\mathrm{T}}\theta_3). \end{aligned} \quad (39)$$

Substituting (31), (35), (36), and (39) into (30), with (34) in mind we obtain the following expression for the Frechet derivative of the functional $J_\beta$ in the direction $\theta = [\,\theta_1\ \theta_2\ \theta_3\,]$ at the point $(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r)$ :

$$\begin{aligned} &dJ_\beta(\theta) = \\ &= 2\operatorname{tr}\{(Q(P + I_r)\widetilde{A}_r^{\mathrm{T}} - [\,0\ I_r\,]P_c\widetilde{A}_e^{\mathrm{T}}\widetilde{C}_e^{\mathrm{T}}[\,0\ I_r\,]^{\mathrm{T}}\widetilde{C}_e)\theta_1\} \\ &+2\operatorname{tr}\{(Q\widetilde{B}_r^{\mathrm{T}} - \widetilde{B}_e^{\mathrm{T}}\widetilde{C}_e^{\mathrm{T}}[\,0\ I_m\,]^{\mathrm{T}}\widetilde{C}_e)\theta_2\} \\ &+2\operatorname{tr}\{(Q\widetilde{C}_r^{\mathrm{T}} - [\,0\ I_p\,]P_c\widetilde{C}_e^{\mathrm{T}})\theta_3\}. \end{aligned} \quad (40)$$

Thus, the gradient of the functional $J_\beta$ at the point $(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}$ is

$$\nabla J_\beta(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) = \left[ \frac{\partial J_\beta}{\partial \widetilde{A}_r} \; \frac{\partial J_\beta}{\partial \widetilde{B}_r} \; \frac{\partial J_\beta}{\partial \widetilde{C}_r} \right], \qquad (41)$$

where

$$\frac{\partial J_\beta}{\partial \widetilde{A}_r} = 2\widetilde{A}_r^{\mathrm{T}}(P + I_r)Q - 2\widetilde{C}_e^{\mathrm{T}}[\,0\ I_r\,]\widetilde{C}_e\widetilde{A}_e P_c[\,0\ I_r\,]^{\mathrm{T}}, \qquad (42)$$

$$\frac{\partial J_\beta}{\partial \widetilde{B}_r} = 2\widetilde{B}_r Q - 2\widetilde{C}_e^{\mathrm{T}}[\,0\ I_m\,]\widetilde{C}_e\widetilde{B}_e, \qquad (43)$$

$$\frac{\partial J_\beta}{\partial \widetilde{C}_r} = 2\widetilde{C}_r Q - 2\widetilde{C}_e P_c[\,0\ I_p\,]^{\mathrm{T}}. \qquad (44)$$

So the optimality conditions for $(\widetilde{A}_r, \widetilde{B}_r, \widetilde{C}_r) \in \widetilde{\mathbb{S}}$ to minimize $J_\beta$ are given by

$$\frac{\partial J_\beta}{\partial \widetilde{A}_r} = 0, \qquad \frac{\partial J_\beta}{\partial \widetilde{B}_r} = 0, \qquad \frac{\partial J_\beta}{\partial \widetilde{C}_r} = 0.$$

Let us formulate the obtained results as

**Theorem 1.** *Let us consider the linear discrete time-invariant system $F_n(z) \in \mathcal{H}_\infty^{p\times m}$ given by state-space equations* (1)*. With given $\alpha > 0$ and $\epsilon > 0$, as well as sufficiently large $\beta > \beta^\star > 0$, the state-space realization $(A_r, B_r, C_r, D_r)$ of the stable reduced-order system $F_r(z)$ given by* (4) *such that the condition*

$$J_\alpha(A_r, B_r, C_r, D_r) = \|\!|F_n - F_r|\!\|_\alpha < \epsilon \qquad (45)$$

*holds true is defined by the solution of the following nonlinear algebraic equation system*

$$\left.\begin{array}{l} R = A_e^{\mathrm{T}} R A_e + qC_e^{\mathrm{T}} C_e + L^{\mathrm{T}}\Sigma^{-1}L \\ L \triangleq \Sigma B_e^{\mathrm{T}} R A_e \\ \Sigma \triangleq (I_m - B_e^{\mathrm{T}} R B_e)^{-1} \end{array}\right\}, \qquad (46)$$

$$P_c = \widetilde{A}_e P_c \widetilde{A}_e^{\mathrm{T}} + \widetilde{B}_e \widetilde{B}_e^{\mathrm{T}}, \qquad (47)$$

$$P = \widetilde{A}_r P \widetilde{A}_r^{\mathrm{T}} + \widetilde{B}_r \widetilde{B}_r^{\mathrm{T}} + \widetilde{A}_r^{\mathrm{T}} \widetilde{A}_r + \widetilde{C}_r^{\mathrm{T}} \widetilde{C}_r + I_r, \qquad (48)$$

$$Q = \widetilde{A}_r^{\mathrm{T}} Q \widetilde{A}_r + I_r/\beta, \qquad (49)$$

$$-\frac{1}{2}\ln\det\frac{m\Sigma}{\mathrm{tr}(LP_cL^{\mathrm{T}} + \Sigma)} = \alpha, \qquad (50)$$

$$\widetilde{A}_r^{\mathrm{T}}(P + I_r)Q - \widetilde{C}_e^{\mathrm{T}}[\,0\ I_r\,]\widetilde{C}_e\widetilde{A}_e P_c[\,0\ I_r\,]^{\mathrm{T}} = 0, \qquad (51)$$

$$\widetilde{B}_r Q - \widetilde{C}_e^{\mathrm{T}}[\,0\ I_m\,]\widetilde{C}_e\widetilde{B}_e = 0, \qquad (52)$$

$$\widetilde{C}_r Q - \widetilde{C}_e P_c[\,0\ I_p\,]^{\mathrm{T}} = 0, \qquad (53)$$

*where*

$$\left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & 0 \end{array}\right] = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & A_r & B_r \\ \hline C & -C_r & 0 \end{array}\right],$$

$$\left[\begin{array}{c|c} \widetilde{A}_e & \widetilde{B}_e \\ \hline \widetilde{C}_e & 0 \end{array}\right] = \left[\begin{array}{cc|c} A + BL_1 & BL_2 & B\Sigma^{1/2} \\ B_rL_1 & A_r + B_rL_2 & B_r\Sigma^{1/2} \\ \hline C & -C_r & 0 \end{array}\right],$$

$$\left[\begin{array}{c|c} \widetilde{A}_r & \widetilde{B}_r \\ \hline \widetilde{C}_r & 0 \end{array}\right] = \left[\begin{array}{c|c} A_r + B_rL_2 & B_r[\,L_1\ \Sigma^{1/2}\,] \\ \hline C_r & 0 \end{array}\right],$$

*with $D_r = D$.*

## 4 Conclusion

A solution to the problem of optimal anisotropy-based approximative model reduction has been proposed. To solve the main problem, an auxiliary problem of weighted $\mathcal{H}_2$ $\epsilon$-optimal discrete-time model approximation has been stated and solved. Optimality conditions defining a solution to the anisotropy-based approximative model reduction problem have been expressed in form of a nonlinear algebraic matrix equation system consisting of Riccati equation, three Lyapunov equations, and four special-type equations. The presented approach guarantees stability of the obtained reduced-order model, which approximates steady-state behaviour of the full-order system, but does not reflects the dynamics of the full-order system.


## Acknowledgements

This work was supported by Russian Foundation for Basic Research (grants 08-08-00567-a and 07-01-92166-NCNI-a).


## References


Vladimirov, I.G., Kurdjukov, A.P., and Semyonov, A.V. (1996-1). On Computing the Anisotropic Norm of Linear Discrete-Time-Invariant Systems. In *Proc. 13th IFAC World Congress.* San-Francisco, CA, USA. pp. 179–184.

Vladimirov, I.G., Kurdjukov, A.P., and Semyonov, A.V. (1996-2). State-Space Solution to Anisotropy-Based Stochastic $\mathcal{H}_\infty$-Optimization Problem. In *13th IFAC World Congress.* San-Francisco, CA, USA. pp. 427–432.

Zhou, K. and Doyle, J.C. (1998). *Essentials of Robust Control*. Prentice Hall Inc., NJ.

Huang, W.Y. and Teo, K.L. (2001). $\mathcal{H}_2$ Near-Optimal Model Reduction. *IEEE Trans. AC*, **46**, pp. 1279–1284.