Anisotropy-based robust performance analysis of finite horizon linear discrete time varying systems*

Igor Vladimirov †, Phil Diamond ‡, Peter Kloeden §

Abstract. We consider a problem of robust performance analysis of a finite-dimensional linear discrete time varying system on a bounded time interval. The system is assumed to be driven by a random disturbance whose probability distribution is not known precisely. The worst-case performance of the system is quantified by its a-anisotropic norm defined as the maximum ratio of the total root mean square value of the output signal to that of the disturbance over disturbances whose anisotropy is bounded from above by a given parameter $a \geq 0$. The anisotropy is an entropy theoretical distance of the probability distribution of a random vector from the Gaussian distributions with zero mean and scalar covariance matrices. Computing the a-anisotropic norm of the system is reduced to solving a set of two difference Riccati and Lyapunov equations and an equation involving determinants of matrices. In two limiting cases, where $a = 0$ and $a \to +\infty$, the a-anisotropic norm coincides with the scaled Frobenius norm and with the ℓ_2-induced operator norm of the system, respectively.

Contents

*The work is supported by Australian Research Council Grant A 1002 7063.

†Corresponding author, Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia, e-mail igv@maths.uq.edu.au.

‡Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia, e-mail pmd@maths.uq.edu.au.

§Department of Mathematics, J.-W. Goethe University, D–60054, Frankfurt am Main, Germany, e-mail kloeden@math.uni-frankfurt.de.

1 Introduction

We consider a problem of robust performance analysis of a finite-dimensional linear discrete time varying (LDTV) system on a bounded time interval. The system is assumed to be driven by a random disturbance. The influence of the disturbance on the output of the system is quantified by the ratio of the total root mean square values of the signals. This quantity coincides with the appropriately scaled Frobenius norm of the system if the disturbance is a sequence of uncorrelated random vectors with zero mean and identity covariance matrices. This last norm is a counterpart of the H_2-norm of time invariant systems. In general, the probability distribution of the disturbance is not known precisely and the worst case performance of the system can be described by the maximum root mean square gain over a set of probability distributions of the disturbance. In the absence of any prior information about this last, the setting leads to the ℓ_2-induced operator norm of the system.

Consider an intermediate situation where *anisotropy* of the disturbance is bounded from above by a given nonnegative parameter a. The anisotropy of a random vector is an entropy theoretical distance of its distribution from the set of Gaussian distributions with zero mean and scalar covariance matrices. In these conditions, the worst-case performance is described by the *a-anisotropic* norm of the system defined as its maximum root mean square gain over such disturbances. In the limiting cases where $a \to 0$ and $a \to \infty$, the a-anisotropic norm gives the scaled Frobenius and ℓ_2-induced operator norms of the system, respectively.

In the present paper, we revise the original definition of anisotropy of random vectors [15], to make the functional and the associated norm practically computable for finite horizon problems, retaining, at the same time, the basic properties of the original definition. The anisotropy as we use it in the paper is a multivariate version of a similar functional considered in [2] for scalar Gaussian random variables. The revised definition of anisotropy yields the same limit per unit time, when applied to long segments of stationary Gaussian sequences, and a similar consistency holds for the a-anisotropic norms. A finite-horizon LDTV system, as an input-output operator, can be identified with a finite matrix, and so all the norms of the former are those of the latter. Using the state space representation of the matrix, we reduce computing the a-anisotropic norm of the system to solving a set of a backward Riccati and a forward Lyapunov difference equations and an equation involving determinants of matrices.

The anisotropy-based approach to robust performance analysis of finite horizon LDTV systems which we develop in this paper is an extension of the one for infinite horizon time invariant systems in [5, 12], [15]–[19]. It pursues the aim of constructing controllers/estimators which would be more robust than the H_2 and less conservative than the H_∞-optimal ones. In this respect, the anisotropy-based approach is not dissimilar to the mixed H_2/H_∞-control [6, 9, 11, 14, 21] or to the minimum entropy H_∞-control [10] approaches where internally stabilizing controllers are sought to minimize the H_2-norm or, respectively, the entropy functional [1], of the closed-loop system under a prescribed H_∞-norm bound on the system. However, a feature of our approach is that it deals with stochastic minimax settings based on a single norm which incorporates the standard H_2 and H_∞ performance criteria as limiting cases.

The paper is organized as follows. In Section 2, we study the a-anisotropic norm for arbitrary finite matrices. In Section 2.1, the definition and basic properties of the anisotropy of random vectors are given. This definition is compared with the original one in Section 2.2. Section 2.3 defines the a-anisotropic norm of matrices. In Section 2.4, formulas for computing the norm are derived. Section 2.5 investigates the rates of convergence of the a-anisotropic norm to its limits as $a \to +0$ and $a \to +\infty$. In Section 3, the results of Section 2 are applied to anisotropy-based performance analysis of systems. Section 3.1 specifies the class of finite horizon LDTV systems considered. State-space formulas for computing the a-anisotropic norm are derived in Section 3.3 and based on the criterion of innerness established in Section 3.2. Section 3.4 outlines an algorithm for computing the anisotropic norm by Newton's iterations which employs the state-space formulas for the ℓ_4-norm of a system provided in Section 3.5.

2 Anisotropy and anisotropic norm

2.1 Anisotropy of random vectors

Let W be an $\mathbb{R}^m$-valued random vector whose probability distribution is absolutely continuous with pdf $f : \mathbb{R}^m \to \mathbb{R}_+$ with respect to the m-dimensional Lebesgue measure mes_m. The differential entropy [3, p. 229] of W is defined by

$$\mathbf{h}(W) = -\mathbf{E} \ln f(W) = -\int_{\mathbb{R}^m} f(w) \ln f(w) dw, \tag{1}$$

where $\mathbf{E}(\cdot)$ is expectation. Denote by $\mathbb{L}_2^m$ the class of absolutely continuously distributed random vectors W with values in $\mathbb{R}^m$ satisfying $\mathbf{E}|W|^2 < +\infty$. Define *anisotropy* of $W \in \mathbb{L}_2^m$ as

$$\mathbf{A}(W) = \frac{m}{2} \ln\left(\frac{2\pi \mathrm{e}}{m} \mathbf{E}|W|^2\right) - \mathbf{h}(W). \tag{2}$$

This definition of anisotropy differs from that in [15] and will be compared with the original in Section 2.2. As we shall see below, (2) yields the same limit per unit time, the mean anisotropy, when applied to long segments of stationary Gaussian sequences. Note that (2) is a multivariate version of a similar functional considered in [2] for scalar Gaussian random variables .

For what follows, denote by $\mathbb{G}^m(\mu, \Sigma)$ the class of $\mathbb{R}^m$-valued Gaussian distributed random vectors with mean $\mathbf{E}\, W = \mu$ and nonsingular covariance matrix $\mathbf{cov}(W) = \mathbf{E}\left((W-\mu)(W-\mu)^T\right) = \Sigma$, so that the corresponding pdf is

$$f(w) = (2\pi)^{-m/2} (\det \Sigma)^{-1/2} \exp(-\frac{1}{2}\|w-\mu\|^2_{\Sigma^{-1}}),$$

where $\|x\|_Q = \sqrt{x^T Q x}$ denotes the (semi-) norm of a vector x induced by a positive (semi-) definite symmetric matrix Q.

Lemma 1

(a) *The anisotropy $\mathbf{A}(W)$ defined by (2) is invariant under rotation and central dilatation of W, i.e. $\mathbf{A}(\lambda U W) = \mathbf{A}(W)$ for any $\lambda \in \mathbb{R} \setminus \{0\}$ and any orthogonal matrix $U \in \mathbb{R}^{m \times m}$;*

(b) *For any positive definite matrix* $\Sigma \in \mathbb{R}^{m\times m}$,

$$\min\left\{\mathbf{A}(W):\ W \in \mathbb{L}_2^m,\ \mathbf{E}\,(WW^T) = \Sigma\right\} = -\frac{1}{2}\ln\det\frac{m\Sigma}{\mathrm{Tr}\,\Sigma}, \tag{3}$$

and the minimum is attained only for $W \in \mathbb{G}^m(0,\Sigma)$;

(c) *For any* $W \in \mathbb{L}_2^m$, $\mathbf{A}(W) \geq 0$. *Moreover,* $\mathbf{A}(W) = 0$ *iff* $W \in \mathbb{G}^m(0,\lambda I_m)$ *for some* $\lambda > 0$, *where* I_m *denotes the identity matrix of order* m.

Proof. The assertion (a) easily follows from the property [3, p. 234] that $\mathbf{h}(JW) = \mathbf{h}(W) + \ln|\det J|$ for any nonsingular matrix J. The assertion (b) immediately follows from what is generically known as the maximum entropy principle [3, p. 270],

$$\max\left\{\mathbf{h}(W):\ W \in \mathbb{L}_2^m,\ \mathbf{E}\,(WW^T) = \Sigma\right\} = \frac{1}{2}\ln\det(2\pi \mathrm{e}\,\Sigma) \tag{4}$$

where the maximum is attained only for $W \in \mathbb{G}^m(0,\Sigma)$. Combining (4) with (2) yields (3). To prove (c), note that the positive definite $(m\times m)$-matrix $M = m\Sigma/\mathrm{Tr}\,\Sigma$ on the right of (3) satisfies $\mathrm{Tr}\,M = m$. Hence, by the arithmetic-geometric mean inequality, $\det M \leq (\mathrm{Tr}\,M/m)^m = 1$ which implies that $\mathbf{A}(W) \geq 0$ for any $W \in \mathbb{L}_2^m$. The assertion (b) also shows that $\mathbf{A}(W) = 0$ iff W is Gaussian distributed with zero mean and $m\mathbf{E}\,(WW^T)/\mathbf{E}\,|W|^2 = I_m$, where the last equality is equivalent to $\mathbf{E}\,(WW^T)$ being scalar. This completes the proof of the lemma. ■

By Lemma 1(c), the anisotropy $\mathbf{A}(W)$ can be regarded as a distance of the probability distribution of W from the set of Gaussian distributions with zero mean and scalar covariance matrices. At the same time, $\mathbf{A}(W)$ quantifies noninvariance of the distribution with respect to the group of rotations.

2.2 Comparison with original definition of anisotropy

The anisotropy of a random vector $W \in \mathbb{L}_2^m$ was defined in [15] as

$$\mathbf{A}_o(W) = -\mathbf{h}(\sigma) = \mathbf{E}\,\ln g(\sigma) \tag{5}$$

where $g : \mathbb{S}_m \to \mathbb{R}_+$ is the pdf of $\sigma = W/|W|$ with respect to the uniform distribution U_m on the unit sphere $\mathbb{S}_m = \{s \in \mathbb{R}^m :\ |s| = 1\}$. Correspondingly, $\mathbf{h}(\sigma)$ is the differential entropy of the random vector σ with respect to U_m. To formulate the lemma below, define the absolutely continuous measure R_m on $\mathbb{R}_+$ by

$$R_m(dr) = S_m r^{m-1}dr, \qquad S_m = \mathrm{mes}_{\,m-1}\mathbb{S}_m. \tag{6}$$

Lemma 2 *For any* $W \in \mathbb{L}_2^m$, *the quantities (2) and (5) are related as*

$$\mathbf{A}(W) = \mathbf{A}_o(W) + \mathbf{I}(\rho;\sigma) + \mathbf{a}_m(\rho) \tag{7}$$

where

$$\mathbf{a}_m(\rho) = \frac{m}{2}\ln\left(\frac{2\pi\mathrm{e}}{m}\mathbf{E}\,\rho^2\right) - \mathbf{h}(\rho). \tag{8}$$

Here, $\mathbf{h}(\rho)$ *is the differential entropy of the random variable* $\rho = |W|$ *with respect to the measure* R_m *in (6), and* $\mathbf{I}(\rho;\sigma)$ *is the mutual information [3, p. 231] between* ρ *and* σ.

Proof. The joint pdf $\varphi : \mathbb{R}_+ \times \mathbb{S}_m \to \mathbb{R}_+$ of ρ and σ with respect to the measure $R_m \times U_m$ is expressed in terms of the pdf f of W as

$$\varphi(r, s) = f(rs). \tag{9}$$

From (5) and (9), it follows that the differential entropy (1) can be rewritten in the form

$$\begin{aligned} \mathbf{h}(W) &= -\mathbf{E}\,\ln\varphi(\rho,\sigma) = \mathbf{h}(\rho,\sigma) \\ &= -\mathbf{A}_o(W) + \mathbf{h}(\rho) - \mathbf{I}(\rho;\sigma). \end{aligned} \tag{10}$$

Plugging (10) in (2) and recalling the notation (8), obtain (7), and the the proof of the lemma is complete. ■

The quantity (8) is always nonnegative, with $\mathbf{a}_m(\rho) = 0$ iff $\frac{m\rho^2}{\mathbf{E}\rho^2}$ is χ^2-distributed with m degrees of freedom as is the random variable $|W|^2$ for $W \in \mathbb{G}^m(0, I_m)$ [20, pp. 183–184]. Hence, the representation (7) implies that $\mathbf{A}(W) \geq \mathbf{A}_o(W)$ for any $W \in \mathbb{L}_2^m$.

Let $W \in \mathbb{L}_2^m$ be partitioned into subvectors $w_1, \ldots, w_r$ of dimensions $m_1, \ldots, m_r$, respectively, so that $m_1 + \ldots + m_r = m$ and

$$W = \begin{bmatrix} w_1 \\ \vdots \\ w_r \end{bmatrix}. \tag{11}$$

For any $1 \leq s \leq t \leq r$, denote by $W_{s:t} = (w_k)_{s \leq k \leq t}$ the $(m_s + \ldots + m_t)$-dimensional subvector of (11) obtained by "stacking" $w_s, \ldots, w_t$.

Lemma 3 *The anisotropy (2) of the random vector (11) is representable as*

$$\mathbf{A}(W) = \mathbf{A}(U) + \sum_{k=1}^{r} \mathbf{A}(w_k) + \sum_{k=2}^{r} \mathbf{I}(w_k;\, W_{1:k-1}) \tag{12}$$

where $U \in \mathbb{G}^m(0, \Sigma)$ has block diagonal covariance matrix

$$\Sigma = \operatorname*{blockdiag}_{1 \leq k \leq r} \left(\frac{1}{m_k} \mathbf{E}\,|w_k|^2\, I_{m_k} \right). \tag{13}$$

Proof. From the chain rule for differential entropy (see for example, [3, p. 232]), it follows that

$$-\mathbf{h}(W) + \sum_{k=1}^{r} \mathbf{h}(w_k) = \sum_{k=2}^{r} \mathbf{I}(W_{1:k-1};\, w_k) = \sum_{k=1}^{r-1} \mathbf{I}(w_k;\, W_{k+1:r}). \tag{14}$$

On the other hand, by (13),

$$\frac{m}{2} \ln\left(\frac{2\pi \mathrm{e}}{m} \mathbf{E}\,|W|^2 \right) - \sum_{k=1}^{r} \frac{m_k}{2} \ln\left(\frac{2\pi \mathrm{e}}{m_k} \mathbf{E}\,|w_k|^2 \right) = -\frac{1}{2} \ln\det\frac{m\Sigma}{\operatorname{Tr}\Sigma} = \mathbf{A}(U). \tag{15}$$

The relations (14) and (15) immediately yield (12), and the proof of the lemma is complete. ■

Now let $W = (w_k)_{k\in\mathbb{Z}}$ be an ergodic strictly stationary sequence of m-dimensional random vectors with $\mathbf{E}\,|w_0|^2 < +\infty$. Assume that the finite dimensional probability distributions of W are all absolutely continuous. Application of Lemma 3 gives that $\frac{1}{N}\mathbf{A}(W_{0:N-1}) = \mathbf{A}(w_0) + \frac{1}{N}\sum_{k=1}^{N-1}\mathbf{I}(w_0;\, W_{-k:-1})$ for all $N \in \mathbb{N}$, and consequently,

$$\lim_{N\to+\infty}\frac{\mathbf{A}(W_{0:N-1})}{N} = \mathbf{A}(w_0) + \mathbf{I}(w_0; (w_k)_{k<0}). \tag{16}$$

Here, $\mathbf{I}(w_0; (w_k)_{k<0}) = \lim_{k\to+\infty}\mathbf{I}(w_0;\, W_{-k:-1})$ is the information about w_0 contained in the past history $(w_k)_{k<0}$ of the sequence W (see [7, Lemma 5.6.1 on p. 132]). One verifies that (16) yields the same formula for the mean anisotropy of a stationary Gaussian sequence W per unit time as that following from the original definition (5) of anisotropy of random vectors [5, Section 3].

2.3 Anisotropic norm of matrices

Let $F \in \mathbb{R}^{p\times m}$ be a given matrix. Interpreting it as a linear operator with random input $W \in \mathbb{L}_2^m$, imagine a player who controls the pdf of W and tries to maximize the *root mean square gain* of F with respect to W defined by

$$\mathbf{Q}(F, W) = \sqrt{\frac{\mathbf{E}\,|FW|^2}{\mathbf{E}\,|W|^2}}. \tag{17}$$

Clearly, for any fixed $W \in \mathbb{L}_2^m$, the map $\mathbf{Q}(\,\cdot\,, W)$ determines a norm on the set $\mathbb{R}^{p\times m}$. If there are no other restrictions on the distribution of W, the quantity $\mathbf{Q}(F, W)$ can be made arbitrarily close to $\|F\|_\infty$, the largest singular value of F.

Suppose, however, that the player is constrained by the condition $\mathbf{A}(W) \le a$ where $a \ge 0$ is a given parameter. In particular, if $a = 0$, then, by Lemma 1(c), the player can only generate Gaussian distributed random vectors $W \in \bigcup_{\lambda>0}\mathbb{G}^m(0, \lambda I_m)$ with zero mean and scalar covariance matrices. For any such W, $\mathbf{Q}(F, W) = \frac{1}{\sqrt{m}}\|F\|_2$ where $\|F\|_2 = \sqrt{\mathrm{Tr}\,(F^T F)}$ is the Frobenius norm of F. In general, for every $a \ge 0$, define the *a-anisotropic norm* of F by

$$|\!|\!| F |\!|\!|_a = \sup\left\{\mathbf{Q}(F, W):\ W \in \mathbb{L}_2^m,\ \mathbf{A}(W) \le a\right\}. \tag{18}$$

Associate with (18) the set of *worst-case* inputs

$$\mathbb{W}_a(F) = \left\{W \in \mathbb{L}_2^m:\ \mathbf{A}(W) \le a,\ \mathbf{Q}(F, W) = |\!|\!| F |\!|\!|_a\right\}. \tag{19}$$

For any given $F \in \mathbb{R}^{p\times m}$, its a-anisotropic norm is a nondecreasing function of $a \ge 0$ satisfying

$$\frac{1}{\sqrt{m}}\|F\|_2 = |\!|\!| F |\!|\!|_0 \le \lim_{a\to+\infty}|\!|\!| F |\!|\!|_a = \|F\|_\infty. \tag{20}$$

These relations show that computing the norm $|\!|\!| F |\!|\!|_a$ for $a > 0$ is of any significance only when the matrix $F \in \mathbb{R}^{p\times m}$ satisfies

$$\|F\|_2 < \sqrt{m}\,\|F\|_\infty. \tag{21}$$

Clearly, the last inequality does not hold iff the matrix $F^T F$ is scalar, i.e. iff $F^T F = \lambda I_m$ for some $\lambda \geq 0$. In particular, (21) holds automatically if $F \neq 0$ and $p < m$. In the sequel, the matrices $F \in \mathbb{R}^{p\times m}$ satisfying (21) will be called *nonround*.

2.4 Computing anisotropic norm

Let $F \in \mathbb{R}^{p\times m}$ be a given nonround matrix. With F, associate the positive semi-definite matrix

$$\Lambda = F^T F \tag{22}$$

whose largest eigen-value is $\|F\|_\infty^2$. For every $q \in [0, \|F\|_\infty^{-2})$, define a positive definite matrix

$$\Sigma(q) = (I_m - q\Lambda)^{-1}. \tag{23}$$

Introduce the functions $\mathcal{A} : [0, \|F\|_\infty^{-2}) \to \mathbb{R}_+$ and $\mathcal{N} : [0, \|F\|_\infty^{-2}) \to [m^{-1/2}\|F\|_2,\ \|F\|_\infty)$ as

$$\mathcal{A}(q) = -\frac{1}{2}\ln\det\frac{m\Sigma(q)}{\mathrm{Tr}\,\Sigma(q)}, \tag{24}$$

$$\mathcal{N}(q) = \sqrt{\frac{\mathrm{Tr}\,(\Lambda\Sigma(q))}{\mathrm{Tr}\,\Sigma(q)}}. \tag{25}$$

It is straightforward to verify that these are analytic on $[0, \|F\|_\infty^{-2})$ and representable in the form

$$\mathcal{A}(q) = \frac{m}{2}(\ln\Phi(q) - \Psi(q)), \tag{26}$$

$$\mathcal{N}(q) = \sqrt{\frac{1}{q}\left(1 - \frac{1}{\Phi(q)}\right)}, \tag{27}$$

with $\mathcal{N}(0) = \lim_{q\to+0}\mathcal{N}(q) = \|F\|_2/\sqrt{m}$, where the functions $\Phi : [0, \|F\|_\infty^{-2}) \to [1, +\infty)$ and $\Psi : [0, \|F\|_\infty^{-2}) \to \mathbb{R}_+$ are given by

$$\Phi(q) = \frac{1}{m}\mathrm{Tr}\,\Sigma(q), \quad \Psi(q) = \frac{1}{m}\ln\det\Sigma(q). \tag{28}$$

The quantities (24) and (25) are related with the anisotropy (2) and root mean square gain (17) as

$$\mathcal{A}(q) = \mathbf{A}(W), \quad \mathcal{N}(q) = \mathbf{Q}(F, W), \quad \Phi(q) = \frac{1}{m}\mathbf{E}\,|W|^2$$

where $W \in \mathbb{G}^m(0, \Sigma(q))$. Such W can be produced from $V \in \mathbb{G}^m(0, I_m)$ as $W = GV$, where $G \in \mathbb{R}^{m\times m}$ is any matrix which factorizes (23) as

$$GG^T = \Sigma(q). \tag{29}$$

To formulate the lemma below, for any $k \in \mathbb{N}$, define the ℓ_{2k}-norm of a matrix $M \in \mathbb{R}^{s\times t}$ by

$$\|M\|_{2k} = (\mathrm{Tr}\,(MM^T)^k)^{1/(2k)}. \tag{30}$$

The ℓ_2-norm is merely the Frobenius norm. A probabilistic interpretation of the so-defined ℓ_4-norm is that if $V \in \mathbb{G}^t(0, I_t)$ then $\|M\|_4^4 = \frac{1}{2}\mathbf{Var}(|MV|^2)$ where $\mathbf{Var}\xi = \mathbf{E}\,(\xi - \mathbf{E}\,\xi)^2$ is the variance of a random variable ξ. Clearly, $\lim_{k\to+\infty}\|M\|_{2k} = \|M\|_\infty$.

Lemma 4 *Let $F \in \mathbb{R}^{p\times m}$ be nonround. Then the functions (24) and (25) are strictly increasing on $[0, \|F\|_\infty^{-2})$, with $\mathcal{A}$ strictly convex, and satisfy*

$$\mathcal{A}(q) \to +\infty, \quad \mathcal{N}(q) \to \|F\|_\infty \quad \textit{as } q \to \|F\|_\infty^{-2} - 0.$$

The derivatives of the functions are computed as

$$\mathcal{A}'(q) = \frac{1}{2q}\left(m\frac{\|G\|_4^4}{\|G\|_2^2} - \|G\|_2^2\right), \tag{31}$$

$$\mathcal{N}'(q) = \frac{1}{2}\mathcal{N}(q)\left(\frac{\Phi'(q)}{\Phi(q)(\Phi(q)-1)} - \frac{1}{q}\right), \tag{32}$$

$$\Phi'(q) = \frac{1}{mq}(\|G\|_4^4 - \|G\|_2^2) \tag{33}$$

where $G \in \mathbb{R}^{m\times m}$ is a matrix satisfying (29).

The proof of the lemma is completely similar to that of [5, Lemma 4] and therefore omitted.

Theorem 1 *Let $F \in \mathbb{R}^{p\times m}$ be nonround. Then for any $a \geq 0$, the a-anisotropic norm (18) is computed as*

$$|\!|\!| F |\!|\!|_a = \mathcal{N}(q), \tag{34}$$

where $q = \mathcal{A}^{-1}(a)$ is the unique solution of the equation $\mathcal{A}(q) = a$. The set of worst-case inputs (19) is expressed in terms of the corresponding matrix (23) as

$$\mathbb{W}_a(F) = \bigcup_{\lambda>0} \mathbb{G}^m(0, \lambda\Sigma(q)). \tag{35}$$

Proof. By the invariance of the root mean square gain (17) with respect to transformations $W \mapsto \mu W$ for any $\mu \in \mathbb{R}\setminus\{0\}$, and by the assertions (a) and (b) of Lemma 1, in order to maximize the functional $\mathbf{Q}(F, W)$, it is advantageous for the hypothetical player of Section 2.3 to generate zero mean Gaussian distributed vectors W. More precisely,

$$\mathbb{W}_a(F) = \bigcup_{\lambda>0,\ \Sigma\in\mathbf{\Sigma}_a^\diamond} \mathbb{G}^m(0, \lambda\Sigma). \tag{36}$$

Here, $\mathbf{\Sigma}_a^\diamond$ is the set of solutions to the constrained optimization problem

$$\text{maximize } \operatorname{Tr}(\Lambda\Sigma) \text{ over } \Sigma \in \mathbf{\Sigma}_a \tag{37}$$

where

$$\mathbf{\Sigma}_a = \left\{\Sigma \in \Pi_m : \ \operatorname{Tr}\Sigma = m,\ -\frac{1}{2}\ln\det\Sigma \leq a\right\}, \tag{38}$$

with Π_m denoting the cone of positive definite symmetric matrices in $\mathbb{R}^{m\times m}$. Correspondingly, the a-anisotropic norm (18) is representable as

$$|\!|\!| F |\!|\!|_a^2 = \max\{\operatorname{Tr}(\Lambda\Sigma) : \ \Sigma \in \mathbf{\Sigma}_a\}. \tag{39}$$

By the concavity of the function $\ln\det(\cdot)$ on Π_m, the set $\mathbf{\Sigma}_a$ is convex. Moreover, it is strictly convex for any $a > 0$, while $\mathbf{\Sigma}_0 = \{I_m\}$. By the Frechet differentiability of $\ln\det(\,\cdot\,)$ with $\frac{\partial \ln\det\Sigma}{\partial\Sigma}[\Delta] = \mathrm{Tr}\,(\Sigma^{-1}\Delta)$, $\Delta = \Delta^T \in \mathbb{R}^{m\times m}$, the convex optimization problem (37) is easily solved by the Lagrange multipliers method. Omitting the technical details which are similar to those in the proof of [5, Theorem 3], note that if F is nonround, the solutions to the problem (37) all have the form $m\Sigma(q)/\mathrm{Tr}\,\Sigma(q)$. Here, the parameter $q \in [0, \|F\|_\infty^{-2})$ is found from the equation $-\frac{1}{2}\ln\det\frac{m\Sigma(q)}{\mathrm{Tr}\,\Sigma(q)} = a$ which, by (24), coincides with $\mathcal{A}(q) = a$ and, by Lemma 4, has a unique solution $q = \mathcal{A}^{-1}(a)$ for any $a \geq 0$. Hence, $\mathbf{\Sigma}_a^\diamond = \{m\Sigma(q)/\mathrm{Tr}\,\Sigma(q)\}$ which, in combination with (36), gives (35). The corresponding maximum value in (39) is $|\!|\!|F|\!|\!|_a^2 = \frac{\mathrm{Tr}\,(\Lambda\Sigma(q))}{\mathrm{Tr}\,\Sigma(q)}$ which, by (25), yields (34), and the proof of the lemma is complete. ■

The representation (39) implies that $|\!|\!|F|\!|\!|_a$ is concave in $a \geq 0$. Indeed, from the concavity of $\ln\det(\,\cdot\,)$, it follows that the sets (38) satisfy $\alpha\mathbf{\Sigma}_a + \beta\mathbf{\Sigma}_b \subset \mathbf{\Sigma}_{\alpha a+\beta b}$ for any $a, b, \alpha, \beta \geq 0$, with $\alpha+\beta = 1$. Hence, by the linearity of the function maximized in (39), $|\!|\!|F|\!|\!|_{\alpha a+\beta b}^2 \geq \alpha|\!|\!|F|\!|\!|_a^2 + \beta|\!|\!|F|\!|\!|_b^2$ which means that $|\!|\!|F|\!|\!|_a^2$ is concave in a and consequently, so is $|\!|\!|F|\!|\!|_a$ by the monotonicity and concavity of the square root function. A slight refinement of these considerations yields that for any nonround matrix F, $|\!|\!|F|\!|\!|_a$ is strictly concave in a.

2.5 Asymptotic behaviour of a-anisotropic norm as $a \to 0$ and $a \to \infty$

The two lemmas below investigate the rates of convergence of the a-anisotropic norm (18) to its limiting values in (20).

Lemma 5 *Let $F \in \mathbb{R}^{p\times m}$ be nonround. Then*

$$|\!|\!|F|\!|\!|_a - \frac{\|F\|_2}{\sqrt{m}} \sim \frac{\sqrt{\mathbf{V}(F)\,a}}{\|F\|_2} \quad \textit{as } a \to +0, \tag{40}$$

where

$$\mathbf{V}(F) = \frac{\|F\|_4^4}{m} - \left(\frac{\|F\|_2^2}{m}\right)^2. \tag{41}$$

Proof. Note that the quantity (41) is nonnegative for any $F \in \mathbb{R}^{p\times m}$, and $\mathbf{V}(F) > 0$ iff F is nonround. By (22) and (23), the functions (28) are analytic on $[0, \|F\|_\infty^{-2})$, with

$$\Phi(q) = 1 + \frac{1}{m}\sum_{k=1}^{+\infty}\|F\|_{2k}^{2k}\,q^k, \quad \Psi(q) = \frac{1}{m}\sum_{k=1}^{+\infty}\frac{\|F\|_{2k}^{2k}}{k}\,q^k \tag{42}$$

where the definition (30) is used. Differentiating (26) and (42), obtain that $\mathcal{A}'(0) = \frac{m}{2}(\Phi'(0)/\Phi(0) - \Psi'(0)) = 0$, and consequently, since $\mathcal{A}(0) = 0$,

$$\mathcal{A}(q) = \frac{\mathcal{A}''(0)}{2}q^2 + O(q^3) \quad \text{as } q \to +0, \tag{43}$$

where

$$\mathcal{A}''(0) = \frac{m}{2}\left(\frac{\Phi''(0)}{\Phi(0)} - \left(\frac{\Phi'(0)}{\Phi(0)}\right)^2 - \Psi''(0)\right) = \frac{m\mathbf{V}(F)}{2} \tag{44}$$

is strictly positive by the nonroundness of F as mentioned above. Therefore, (43) implies that

$$\mathcal{A}^{-1}(a) = \sqrt{\frac{2a}{\mathcal{A}''(0)}} + O(a) \quad \text{as } a \to +0. \tag{45}$$

From (42), it follows that $\Psi'(q) = \frac{\Phi(q)-1}{q}$ and consequently, by (27), $\mathcal{N}(q) = \sqrt{\frac{\Psi'(q)}{\Phi(q)}}$. The last representation and the equality $\mathcal{N}(0) = \frac{\|F\|_2}{\sqrt{m}}$ imply that

$$\mathcal{N}(q) = \frac{\|F\|_2}{\sqrt{m}} + \mathcal{N}'(0)q + O(q^2) \quad \text{as } q \to +0, \tag{46}$$

where

$$\mathcal{N}'(0) = \frac{\mathcal{N}(0)}{2}\left(\frac{\Psi''(0)}{\Psi'(0)} - \frac{\Phi'(0)}{\Phi(0)}\right) = \frac{\sqrt{m}\,\mathbf{V}(F)}{2\|F\|_2}. \tag{47}$$

Recalling Theorem 1 and using (45) and (46), obtain that

$$|||F|||_a = \frac{\|F\|_2}{\sqrt{m}} + \mathcal{N}'(0)\sqrt{\frac{2a}{\mathcal{A}''(0)}} + O(a) \quad \text{as } a \to +0,$$

which, in combination with (44) and (47), yields (40), and the proof of the lemma is complete. ∎

Lemma 6 *Let $F \in \mathbb{R}^{p\times m}$ be nonround. Then*

$$\begin{aligned} \|F\|_\infty - |||F|||_a \ \sim\ & \frac{1}{2}\|F\|_\infty (1-\theta)\,\theta^{\frac{\theta}{1-\theta}} \\ & \times \exp\left(-\frac{2}{1-\theta}\left(\frac{a}{m} + \mathbf{J}(F)\right)\right) \quad \textit{as } a \to +\infty, \end{aligned} \tag{48}$$

where

$$\mathbf{J}(F) = -\frac{1}{2m}\sum_{1\le k\le m:\ \lambda_k < \|F\|_\infty^2} \ln(1 - \lambda_k \|F\|_\infty^{-2}), \qquad \theta = \frac{\mu}{m}. \tag{49}$$

Here, $\lambda_1, \ldots, \lambda_m$ denote the eigenvalues of the matrix (22), and μ is the multiplicity of the largest eigenvalue $\|F\|_\infty^2$.

Proof. Using (23) and (49), it is straightforward to verify that the functions (28) satisfy

$$\Phi(K(\varepsilon)) = \frac{\theta}{\varepsilon} + O(1), \quad \Psi(K(\varepsilon)) = -\theta\ln\varepsilon + 2\mathbf{J}(F) + O(\varepsilon) \quad \text{as } \varepsilon \to +0, \tag{50}$$

where $K(\varepsilon) = (1-\varepsilon)\,\|F\|_\infty^{-2}$. Plugging (50) in (26) and (27), obtain that

$$\mathcal{A}(K(\varepsilon)) \ = \ \frac{m}{2}\left(\ln\theta - (1-\theta)\ln\varepsilon - 2\mathbf{J}(F) + O(\varepsilon)\right), \tag{51}$$

$$\mathcal{N}(K(\varepsilon)) \ = \ \|F\|_\infty\left(1 - \frac{1-\theta}{2\,\theta}\varepsilon + O(\varepsilon^2)\right) \quad \text{as } \varepsilon \to +0. \tag{52}$$

Note that $\theta < 1$ by the nonroundness of F. From (51), it now follows that

$$(\mathcal{A}\circ K)^{-1}(a) \sim \theta^{\frac{1}{1-\theta}} \exp\left(-\frac{2}{1-\theta}\left(\frac{a}{m} + \mathbf{J}(F)\right)\right) \quad \text{as } a \to +\infty. \tag{53}$$

Rephrasing Theorem 1 as $|||F|||_a = (\mathcal{N}\circ K)\left((\mathcal{A}\circ K)^{-1}(a)\right)$ and combining (53) with (52), obtain (48), and the proof of the lemma is complete. ∎

3 Anisotropy-based performance analysis

3.1 Problem statement

Consider a LDTV system F on a bounded time interval $[0, N]$ with m-dimensional input disturbance $W = (w_k)_{0\le k\le N}$, p-dimensional output $Z = (z_k)_{0\le k\le N} = F * W$ and n-dimensional internal state $X = (x_k)_{0\le k\le N}$ governed by the equations

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k w_k, &(54)\\ z_k &= C_k x_k + D_k w_k, &(55)\end{aligned}$$

with initial condition $x_0 = 0$. Here, $A_\bullet$, $B_\bullet$, $C_\bullet$, $D_\bullet$ are appropriately dimensioned matrix-valued functions of time. That the system F has the state space realization (54)–(55), will be written as

$$F = \left[\begin{array}{c|c} A_\bullet & B_\bullet \\ \hline C_\bullet & D_\bullet \end{array}\right]. \tag{56}$$

For any $0 \le s \le t \le N$, the segments of the signals W and Z on the time interval $[s, t]$ will be organized as the vectors

$$W_{s:t} = \begin{bmatrix} w_s \\ \vdots \\ w_t \end{bmatrix}, \qquad Z_{s:t} = \begin{bmatrix} z_s \\ \vdots \\ z_t \end{bmatrix}. \tag{57}$$

Clearly, the assumption $x_0 = 0$ implies that $Z_{0:t} = F_{0:t} W_{0:t}$ where

$$F_{s:t} = \operatorname*{block}_{s\le j,k\le t}(f_{jk}) \tag{58}$$

is a block lower triangular matrix whose $(p \times m)$-blocks are given by

$$f_{jk} = \begin{cases} C_j T_{j,k+1} B_k & \text{for } j > k \\ D_k & \text{for } j = k \\ 0 & \text{otherwise} \end{cases}, \tag{59}$$

with

$$T_{jk} = A_{j-1} \times \ldots \times A_k \tag{60}$$

denoting the state transition matrix. In what follows, the system F on $[0, N]$ is identified, as an input-output operator, with the matrix $F_{0:N}$. Correspondingly, all the norms of F are those of $F_{0:N}$. In Sections 3.3, state space formulas will be derived for computing the a-anisotropic norm $|\!|\!|F|\!|\!|_a$ under the assumption that F is nonround in the sense that so is $F_{0:N}$. This will be based on the lemma of the next section.

3.2 Criterion of innerness

The system F with state space realization (56) is said to be *inner* on the time interval $[0, N]$ if the corresponding matrix $F_{0:N}$ defined by (58) satisfies $F_{0:N}^T F_{0:N} = I_{(N+1)m}$.

The innerness means that $|Z_{0:N}| = |W_{0:N}|$ for any input W. To formulate the lemma below, denote by

$$
\begin{aligned}
P_j &= \sum_{k=0}^{j-1} T_{j,k+1} B_k B_k^T T_{j,k+1}^T, && (61)\\
Q_k &= \sum_{j=k}^{N} T_{jk}^T C_j^T C_j T_{jk} && (62)
\end{aligned}
$$

respectively the controllability and observability gramians of F where T_{jk} is the state transition matrix (60). These satisfy the difference Lyapunov equations

$$
\begin{aligned}
P_{k+1} &= A_k P_k A_k^T + B_k B_k^T, && (63)\\
Q_k &= A_k^T Q_{k+1} A_k + C_k^T C_k && (64)
\end{aligned}
$$

propagating over $0 \le k \le N$ in the opposite directions with boundary conditions $P_0 = Q_{N+1} = 0$.

Lemma 7 *The system F with state space realization (56) is inner on the time interval $[0, N]$ iff for every $0 \le k \le N$,*

$$
\begin{aligned}
B_k^T Q_{k+1} B_k + D_k^T D_k &= I_m, && (65)\\
P_k(A_k^T Q_{k+1} B_k + C_k^T D_k) &= 0. && (66)
\end{aligned}
$$

Proof. For every $0 \le k \le N$, denote

$$
H_k(W_{0:k}) = |Z_{0:k}|^2 + \|x_{k+1}\|_{Q_{k+1}}^2. \qquad (67)
$$

Clearly, (67) is a quadratic form in $W_{0:k}$ which, by (62), satisfies

$$
H_k(W_{0:k}) = |Z_{0:N}|^2 \quad \text{if} \quad W_{k+1:N} = 0. \qquad (68)
$$

From (54), (55) and (64), it follows that for any input W,

$$
\begin{aligned}
H_k(W_{0:k}) &= |Z_{0:k-1}|^2 + |C_k x_k + D_k w_k|^2 + \|A_k x_k + B_k w_k\|_{Q_{k+1}}^2\\
&= H_{k-1}(W_{0:k-1}) + \|w_k\|_{B_k^T Q_{k+1} B_k + D_k^T D_k}^2\\
&+ 2x_k^T (A_k^T Q_{k+1} B_k + C_k^T D_k) w_k. && (69)
\end{aligned}
$$

Now let (65)–(66) hold for all $0 \le k \le N$. Then the last term on the right of (69) vanishes since $x_k \in \mathrm{im}\,(P_k) = P_k \mathbb{R}^n$ and hence, by induction,

$$
H_k(W_{0:k}) = H_{k-1}(W_{0:k-1}) + |w_k|^2 = |W_{0:k}|^2.
$$

In particular, $H_N(W_{0:N}) = |W_{0:N}|^2$ for any W, which, by (68), implies the innerness of F on $[0, N]$, thereby proving the sufficiency of (65)–(66). Now let F be inner. Then for any $0 \le k \le N$ and any signal W such that $W_{0:k-1} = 0$ and $W_{k+1:N} = 0$, we have $x_k = 0$ and hence, by (69),

$$
H_k([0, w_k^T]^T) = \|w_k\|_{B_k^T Q_{k+1} B_k + D_k^T D_k}^2 = |w_k|^2
$$

which, by the arbitrariness of w_k, implies (65). Now assume that $W_{k+1:N} = 0$ while $W_{0:k}$ is arbitrary. Then, by the innerness and by (68),

$$H_k(W_{0:k}) = H_{k-1}(W_{0:k-1}) + |w_k|^2 + 2x_k^T(A_k^T Q_{k+1} B_k + C_k^T D_k)w_k = |W_{0:k}|^2,$$

and consequently,

$$H_{k-1}(W_{0:k-1}) - |W_{0:k-1}|^2 = -2x_k^T(A_k^T Q_{k+1} B_k + C_k^T D_k)w_k. \tag{70}$$

The left hand side of (70) is a quadratic form in $W_{0:k-1}$ while the right hand side is a bilinear function of $W_{0:k-1}$ and w_k. Hence, by the arbitrariness of $W_{0:k}$ in (70), $x_k^T(A_k^T Q_{k+1} B_k + C_k^T D_k) = 0$ for any $W_{0:k-1}$ which implies (66), and the proof of the lemma is complete. ■

Note that the criterion of innerness (65)–(66) is similar to the one for linear time invariant systems [8, Lemma 5 on p. 1693]. Omitting the controllability gramians P_k in (66), obtain sufficient conditions for innerness.

3.3 State space formulas for anisotropic norm

For any given $q \in [0, \|F\|_\infty^{-2})$, consider the matrices $S_k \in \mathbb{R}^{m\times m}$ and $L_k \in \mathbb{R}^{m\times n}$ associated with a sequence of matrices $Q_k \in \mathbb{R}^{n\times n}$, $0 \le k \le N$, satisfying the backward difference Riccati equation

$$\begin{aligned} Q_k &= A_k^T Q_{k+1} A_k + q C_k^T C_k + L_k^T S_k^{-1} L_k & (71)\\ S_k &= (I_m - B_k^T Q_{k+1} B_k - q D_k^T D_k)^{-1}, & (72)\\ L_k &= S_k(B_k^T Q_{k+1} A_k + q D_k^T C_k), & (73) \end{aligned}$$

with boundary condition $Q_{N+1} = 0$. If $q = 0$, a simple induction shows that $Q_k = 0$, $S_k = I_m$ and $L_k = 0$ for all $0 \le k \le N$. The statement below is a counterpart of the discrete bounded real lemma for time invariant systems [4] (also see [13, Theorem 4.6.6 on p. 71]).

Lemma 8 *For every $q \in [0, \|F\|_\infty^{-2})$, the matrices Q_k, governed by the Riccati equation (71)–(73), are all positive semi-definite and S_k positive definite.*

Proof. Let $q < \|F\|_\infty^{-2}$. Then the matrix $I_{m(N+1)} - qF_{0:N}^T F_{0:N}$ is positive definite and hence, the problem of minimizing the quadratic form $|W_{0:N}|^2 - q|Z_{0:N}|^2$ over all possible vectors $W_{0:N}$ has the unique solution 0. On the other hand, the minimum value is $J_0(0)$ where for every $0 \le k \le N$, $J_k(x)$ is the optimal cost in the following control problem for the system F over the time interval $[k, N]$ starting at a state $x \in \mathbb{R}^n$,

$$J_k(x) = \inf\{|W_{k:N}|^2 - q|Z_{k:N}|^2 : \ x_k = x, \ W_{k:N} \in \mathbb{R}^{m(N-k+1)}\}. \tag{74}$$

Clearly, $J_k : \mathbb{R}^n \to \mathbb{R}_-$ since the function minimized in (74) is nonpositive for $W_{k:N} = 0$. By (54)–(55), the functions J_k satisfy the Bellman equation

$$J_k(x) = \inf_{w\in\mathbb{R}^m}(J_{k+1}(A_k x + B_k w) + |w|^2 - q|C_k x + D_k w|^2), \tag{75}$$

with boundary condition $J_{N+1} \equiv 0$. Now prove by induction that the minimum in (74) is achievable and that

$$J_k(x) = -\|x\|^2_{Q_k}, \tag{76}$$

using (75) and establishing at each step the assertion of the lemma. For $k = N+1$, the representation (76) is obvious by the conditions $Q_{N+1} = 0$ and $J_{N+1} \equiv 0$. Assume that (76) is already proved for J_{k+1} for some $0 \le k \le N$, with Q_{k+1} positive semi-definite. Then the function to be minimized on the right of (75) is a quadratic form in x and w,

$$\begin{aligned} & -\|A_k x + B_k w\|^2_{Q_{k+1}} + |w|^2 - q|C_k x + D_k w|^2 \\ = & -\|x\|^2_{A_k^T Q_{k+1} A_k + q C_k^T C_k} - 2w^T(B_k^T Q_{k+1} A_{k+1} + q D_k^T C_k)x \\ & + w^T(I_m - B_k^T Q_{k+1} B_k - q D_k^T D_k)w. \end{aligned} \tag{77}$$

Now note that if the symmetric matrix $I_m - B_k^T Q_{k+1} B_k - q D_k^T D_k$ is not positive definite then there exists $w \in \mathbb{R}^m \setminus \{0\}$ such that the last quadratic form on the right of (77) takes a nonpositive value for $x = 0$. But this implies the existence of a vector $W_{k:N}$, with $w_k = w$, such that $|W_{k:N}|^2 - q|Z_{k:N}|^2 \le 0$ with $x_k = 0$ which contradicts the positive definiteness of the matrix $I_{m(N-k+1)} - qF_{k:N}^T F_{k:N}$ following from that of $I_{m(N+1)} - qF_{0:N}^T F_{0:N}$. The contradiction shows that the matrix S_k in (72) is well defined and positive definite. Now recalling the whole Riccati equation (71)–(73), and minimizing the right hand side of (77) over $w \in \mathbb{R}^m$, obtain (76). This completes the inductive step and the proof of the lemma. ■

Now let $V = (v_k)_{0 \le k \le N}$ be a sequence of independent random vectors $v_k \in \mathbb{G}^m(0, I_m)$. Consider the input disturbance W to the system (54)–(55) generated from V as

$$w_k = L_k x_k + \sqrt{S_k}\, v_k, \quad 0 \le k \le N, \tag{78}$$

where the matrices L_k and S_k are associated with the Riccati equation (71)–(73) for some $q \in [0, \|F\|_\infty^{-2})$. By Lemma 8, the matrix square roots $\sqrt{S_k}$ are all well defined and assumed symmetric and positive definite, for convenience. Therefore, $W = G * V$ where G is a LDTV shaping filter with state-space realization

$$G = \left[\begin{array}{c|c} A_\bullet + B_\bullet L_\bullet & B_\bullet \sqrt{S_\bullet} \\ \hline L_\bullet & \sqrt{S_\bullet} \end{array}\right]. \tag{79}$$

The scenario is depicted in Fig. 1. By the nonsingularity of the matrices $\sqrt{S_k}$, the filter is invertible and

$$G^{-1} = \left[\begin{array}{c|c} A_\bullet & B_\bullet \\ \hline -S_\bullet^{-1/2} L_\bullet & S_\bullet^{-1/2} \end{array}\right]. \tag{80}$$

Clearly, the invertibility is equivalent to nonsingularity of the matrix $G_{0:N}$ which is defined in the same way as the matrix $F_{0:N}$ in (58) for the system F. The probabilistic meaning of the matrices S_k is that $S_k = \mathbf{cov}(w_k \mid W_{0:k-1})$ under the scenario $W = G * V$.[1]

[1]Here, $\mathbf{cov}(\xi \mid \eta) = \mathbf{E}\left((\xi - \mathbf{E}(\xi \mid \eta))(\xi - \mathbf{E}(\xi \mid \eta))^T \mid \eta\right)$ is the conditional covariance matrix of a random vector ξ with respect to the σ-algebra induced by another vector η, with $\mathbf{E}(\cdot \mid \cdot)$ denoting the conditional expectation. If the joint distribution of the vectors is Gaussian, the matrix is nonrandom.

Theorem 2 *Let the system F with state-space realization (56) be nonround. Then for any $a \geq 0$, its a-anisotropic norm is computed as $\|F\|_a = \mathcal{N}(q)$, where $q = \mathcal{A}^{-1}(a) \in [0, \|F\|_\infty^{-2})$ is the unique solution of the equation $\mathcal{A}(q) = a$. Here,*

$$\begin{aligned} \mathcal{A}(q) &= \frac{m(N+1)}{2}(\ln \Phi(q) - \Psi(q)), & (81)\\ \mathcal{N}(q) &= \sqrt{\frac{1}{q}\left(1 - \frac{1}{\Phi(q)}\right)}, & (82)\\ \Phi(q) &= \frac{1}{m(N+1)}\sum_{k=0}^{N} \mathrm{Tr}\,(L_k P_k L_k^T + S_k), & (83)\\ \Psi(q) &= \frac{1}{m(N+1)}\sum_{k=0}^{N} \ln\det S_k, & (84) \end{aligned}$$

where the matrices S_k and L_k are associated with the Riccati equation (71)–(73), and P_k are the controllability gramians of the filter G in (79) satisfying the difference Lyapunov equation

$$P_{k+1} = (A_k + B_k L_k)P_k(A_k + B_k L_k)^T + B_k S_k B_k^T \tag{85}$$

with initial condition $P_0 = 0$. Moreover, G is a worst-case shaping filter for the system F.

Proof. Using (80), consider a LDTV system Θ on the time interval $[0, N]$ with input W and state-space realization

$$\Theta = \begin{bmatrix} \sqrt{q}F \\ \\ G^{-1} \end{bmatrix} = \left[\begin{array}{c|c} A_\bullet & B_\bullet \\ \hline \sqrt{q}C_\bullet & \sqrt{q}D_\bullet \\ -S_\bullet^{-1/2}L_\bullet & S_\bullet^{-1/2} \end{array}\right]. \tag{86}$$

From (71) and from the boundary condition $Q_{N+1} = 0$, it follows by induction that the matrices Q_k are the controllability gramians of Θ. Recalling the rest of the Riccati equation (72)–(73) and applying Lemma 7 to (86), obtain that the system Θ is inner and consequently,

$$G_{0:N}G_{0:N}^T = (I_{m(N+1)} - qF_{0:N}^T F_{0:N})^{-1}.$$

Therefore, by Theorem 1, the signal $W = G * V$ will be made a worst-case disturbance for the system F, i.e. $W_{0:N} \in \mathbb{W}_a(F_{0:N})$, if the parameter $q \in [0, \|F\|_\infty^{-2})$ is adjusted so as to satisfy the equation $\mathbf{A}(W_{0:N}) = a$. This last reduces to $\mathcal{A}(q) = a$, where the function $\mathcal{A}$ is given by (81) and

$$\Phi(q) = \frac{\mathbf{E}\,|W_{0:N}|^2}{m(N+1)}, \quad \Psi(q) = \frac{\det\mathbf{cov}(W_{0:N})}{m(N+1)}. \tag{87}$$

By the structure of the shaping filter (79), the enumerators in (87) are computed as

$$\mathbf{E}\,|W_{0:N}|^2 = \|G\|_2^2 = \sum_{k=0}^{N} \mathrm{Tr}\,(L_k P_k L_k^T + S_k), \quad \det\mathbf{cov}(W_{0:N}) = \prod_{k=0}^{N}\det S_k, \tag{88}$$

where the probabilistic meaning of the controllability gramians in (85) is that P_k is the covariance matrix $\mathbf{cov}(x_k)$ of the internal state of the system F under the scenario $W = G * V$. Plugging (88) in (87) yields (83) and (84), and the proof of the theorem is complete. ■

3.4 Computing anisotropic norm by Newton's iterations

By Theorem 2, computation of the a-anisotropic norm of a nonround system (56) for a given $a > 0$ reduces to solution of the equation $\mathcal{A}(q) = a$ in $q \in (0, \|F\|_\infty^{-2})$. This can be accomplished using Newton's iterations. More precisely, define the sequence of numbers $(q_k)_{k\geq 0}$ recursively by

$$q_{k+1} = \begin{cases} (\|F\|_\infty^{-2} + q_k)/2 & \text{if } \mathcal{A}(q_k) < a \\ q_k + (a - \mathcal{A}(q_k))\,/\mathcal{A}'(q_k) & \text{otherwise} \end{cases} \tag{89}$$

with initial condition $q_0 = 0$. From smoothness, monotonicity and convexity of the function $\mathcal{A}$ in (81), the sequence q_k converges to the solution sought, $\lim_{k\to+\infty} q_k = \mathcal{A}^{-1}(a)$. Moreover, $0 = q_0 < \ldots < q_{k_*} \geq q_k \geq q_{k+1} \geq \mathcal{A}^{-1}(a)$ for all $k \geq k_*$. Here,

$$k_* = \min\left\{k \geq 1 : \ \mathcal{A}((1 - 2^{-k})\|F\|_\infty^{-2}) \geq a\right\}.$$

Therefore, the Newton's method is activated after the k_*-th iteration in (89). In practice, the iterations can be halted upon fulfillment of the condition

$$(\mathcal{A}(q_k) - a)\ \max\left(1,\ \frac{\mathcal{N}'(q_k)}{\mathcal{A}'(q_k)\mathcal{N}(q_k)}\right) < \varepsilon \tag{90}$$

thus providing, asymptotically, the relative accuracy $\varepsilon \ll 1$ for computing the a-anisotropic norm $\|F\|_a = \mathcal{N}(\mathcal{A}^{-1}(a))$. The derivatives of the functions $\mathcal{A}$ and $\mathcal{N}$ defined by (81)–(84) are computed by the following formulas similar to (31)–(33),

$$\begin{aligned}
\mathcal{A}'(q) &= \frac{1}{2q}\left(m(N+1)\frac{\|G\|_4^4}{\|G\|_2^2} - \|G\|_2^2\right), \\
\mathcal{N}'(q) &= \frac{1}{2}\mathcal{N}(q)\left(\frac{\Phi'(q)}{\Phi(q)(\Phi(q)-1)} - \frac{1}{q}\right), \\
\Phi'(q) &= \frac{\|G\|_4^4 - \|G\|_2^2}{m(N+1)q},
\end{aligned}$$

where G is the shaping filter described in Theorem 2, and these give explicit expressions for (89) and (90). The formulas involve the ℓ_4-norm of the filter which, by Lemma 9 of the next section, can be computed as

$$\|G\|_4^4 = \sum_{k=0}^{N} \mathrm{Tr}\left((L_k P_k L_k^T + S_k)^2 + 2M_k^T R_{k+1} M_k\right), \quad M_k = (A_k + B_k L_k)P_k L_k^T + B_k S_k.$$

Here, P_k and R_k are the controllability and observability gramians of the filter satisfying respectively (85) and the backward difference Lyapunov equation

$$R_k = (A_k + B_k L_k)^T R_{k+1}(A_k + B_k L_k) + L_k^T L_k$$

with boundary condition $R_{N+1} = 0$.

3.5 State space formulas for ℓ_4-norm

The lemma below provides state-space formulas for computing the ℓ_4-norm of a system, as defined by (30).

Lemma 9 *The ℓ_4-norm of the system F with state space realization (56) on the time interval $[0, N]$ is expressed in terms of its controllability and observability gramians (61)–(64) as*

$$\begin{aligned}\|F\|_4^4 &= \sum_{k=0}^{N} \mathrm{Tr}\,((C_k P_k C_k^T + D_k D_k^T)^2 \\ &+ 2\,(A_k P_k C_k^T + B_k D_k^T)^T Q_{k+1}(A_k P_k C_k^T + B_k D_k^T)).\end{aligned} \tag{91}$$

Proof. Let the vector $W_{0:N}$ in (57) be Gaussian distributed with zero mean and identity covariance matrix. Then

$$\|F\|_4^4 = \mathrm{Tr}\, H^2 = \sum_{k=0}^{N} \mathrm{Tr}\,(h_{kk}^2 + 2 \sum_{j=k+1}^{N} h_{jk}^T h_{jk}) \tag{92}$$

where $H = \mathbf{cov}(Z_{0:N}) = \mathrm{block}_{0 \le j,k \le N}(h_{jk})$, with $h_{jk} = \mathbf{cov}(z_j, z_k)$. Here, $\mathbf{cov}(\xi, \eta) = \mathbf{E}\,((\xi - \mathbf{E}\,\xi)(\eta - \mathbf{E}\,\eta)^T)$ denotes the cross-covariance matrix of random vectors ξ and η. By (54), for every $0 \le k \le N$, x_k is $W_{0:k-1}$-measurable and, since w_k is independent of $W_{0:k-1}$, the diagonal blocks of H are

$$h_{kk} = C_k \mathbf{cov}(x_k) C_k^T + D_k D_k^T. \tag{93}$$

By a similar reasoning, for any $N \ge j > k \ge 0$, $\mathbf{E}\,(z_j \mid W_{0:k}) = C_j T_{j,k+1} x_{k+1}$, where T_{jk} is the state transition matrix (60). Consequently, by the $W_{0:k}$-measurability of z_k in (55), $h_{jk} = \mathbf{E}\,(\mathbf{E}\,(z_j \mid W_{0:k})\, z_k^T) = C_j T_{j,k+1} \mathbf{cov}(x_{k+1}, z_k)$ which, in combination with (62), yields

$$\sum_{j=k+1}^{N} h_{jk}^T h_{jk} = (\mathbf{cov}(x_{k+1}, z_k))^T Q_{k+1} \mathbf{cov}(x_{k+1}, z_k). \tag{94}$$

It now remains to note that $\mathbf{cov}(x_k) = P_k$ and $\mathbf{cov}(x_{k+1}, z_k) = A_k P_k C_k^T + B_k D_k^T$. Plugging these last relations in (93) and (94), from (92) obtain (91), and the proof of the lemma is complete. ■

References

[1] Arov D.Z. and Krein M.G. [1983] On computing the entropy integrals and their minima in generalized extension problems, *Act. Sci. Mat.*, **45**, 33–50.

[2] Bernhard H.-P. [1998] A tight upper bound on the gain of linear and nonlinear predictors for stationary stochastic processes, *IEEE Transactions on Signal Processing*, **46**, 11, 2909–2917.

[3] Cover T.M., and Thomas J.A. [1991] *Elements of Information Theory* (John Wiley & Sons, New York).

[4] De Souza C.E., and Xie L. [1992] On the discrete-time bounded real lemma with application in the characterization of static state feedback H_∞-controllers, *Systems and Control Letters*, **18**, 61–71.

[5] Diamond P., Vladimirov I., Kurdjukov A., and Semyonov A. [2001] Anisotropy-based performance analysis of linear discrete time invariant control systems, *Int. J. Control*, **74**, 1, 28–42.

[6] Doyle J., Zhou K., Glover K., and Bodenheimer B. [1994] Mixed H_2 and H_∞ performance objectives II: Optimal control, *IEEE Transactions on Automatic Control*, **39**, 8, 1575–1587.

[7] Gray R.M. [1990] *Entropy and Information Theory* (Springer, New York).

[8] Gu D.-W., Tsai M.C., O'Young S.D. and Postlethwaite I. [1989] State-space formulae for discrete-time H^∞ optimization, *International Journal of Control*, **49**, 1683–1723.

[9] Haddad W.M., Bernstein D.S., and Mustafa D. [1991] Mixed H_2/H_∞ regulation and estimation: The discrete-time case, *Systems & Control Letters*, **16**, 235–247.

[10] Mustafa D. and Glover K. [1991] *Minimum Entropy H_∞-Control* (Springer-Verlag, New York).

[11] Rotea M.A. and Khargonekar P.P. [1991] H_2-optimal control with an H_∞-constraint: The state feedback case, *Automatica*, **27**(2), 307–316.

[12] Semyonov A.V., Vladimirov I.G., Kurdjukov A.P. [1994] Stochastic approach to H_∞-optimization, *Proceedings of the 33rd Conference on Decision and Control, Florida, USA, December 14–16*, **3**, 2249–2250.

[13] Skelton R, Iwasaki T., and Grigoriadis K. [1998] *A Unified Algebraic Approach to Linear Control Design* (Taylor & Francis, London).

[14] Stoorvogel A.A. [1993] The robust H_2 control problem: a worst case design, *IEEE Trans. Automat. Contr.*, **AC–38**(9), 1358–1370.

[15] Vladimirov I.G., Kurdjukov A.P. and Semyonov A.V. [1995] Anisotropy of signals and the entropy of linear stationary systems, *Doklady Akademii Nauk, Mathematics,* **342**, 5, 583–585 (in Russian) (English translation: 1995, **51**, 3, 388–390).

[16] Vladimirov I.G., Kurdjukov A.P. and Semyonov A.V. [1995] The stochastic problem of H_∞-optimization, *Doklady Akademii Nauk, Control Theory*, **343**, 5, 607–609 (in Russian) (English translation: 1995, **52**, 1, 155–157).

[17] Vladimirov I.G., Kurdjukov A.P., and Semyonov A.V. [1996] On computing the anisotropic norm of linear discrete-time-invariant systems, *Proceedings of the 13th IFAC World Congress, San-Francisco, California, USA, June 30–July 5*, **G**, 179–184, Paper IFAC-2d-01.6.

[18] Vladimirov I.G., Kurdjukov A.P., and Semyonov A.V. [1996] State-space solution to anisotropy-based stochastic H_∞-optimization problem, *Proceedings of the 13th IFAC World Congress, San-Francisco, California, USA, June 30–July 5*, **H**, 427–432, Paper IFAC-3d-01.6.

[19] Vladimirov I.G., Kurdjukov A.P., and Semyonov A.V. [1999] Asymptotic of anisotropic norm of linear time invariant systems, *Avtomatika i Telemekhanika*, **3**, 78–87 (in Russian).

[20] Wilks S.S. [1963] *Mathematical Statistics* (John Wiley & Sons, New York).

[21] Zhou K., Glover K., Bodenheimer B., and Doyle J. [1994] Mixed H_2 and H_∞ performance objectives I: Robust performance analysis, *IEEE Transactions on Automatic Control*, **39**, 8 1564–1574.

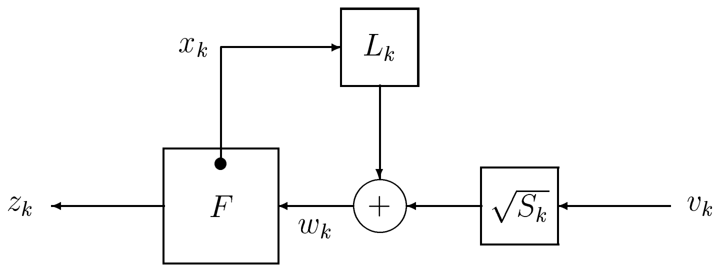

Figure 1: The structure of the shaping filter (79).