Feedback, correlations, and propagation of mean anisotropy of signals in filter connections

Alexander P. Kurdyukov and Igor G. Vladimirov

Abstract—The anisotropy-based approach to robust control in stochastic systems occupies a unifying intermediate position between the \( H_2 \) and \( H_\infty \)-optimal control theories. Initiated at the interface of Information Theory and Robust Control about thirteen years ago, the approach employs the \( \alpha \)-anisotropic norm of a linear system as its worst-case sensitivity to input random disturbances whose mean anisotropy is bounded by a nonnegative parameter \( \alpha \). The latter quantifies the temporal “colouredness” and spatial “non-roundness” of the signal by its minimal relative entropy production rate with respect to Gaussian white noises with scalar covariance matrices. Revisiting the underlying definitions, the paper emphasizes the role of feedback in the construct of mean anisotropy of signals and discusses propagation of the latter through various filter connections. The results can be used to support physical and engineering intuition for a “rational” choice of the mean anisotropy level \( \alpha \) in the design of anisotropy-based robust controllers.

Index Terms—stochastic robust control, mean anisotropy, anisotropic norm, filter connections, Kolmogorov-Szegö formula.

I. INTRODUCTION

Performance analysis and optimal design of control systems and signal processing devices substantially rely on the statistical characteristics of the underlying random disturbances. The inevitable uncertainty in the knowledge of these last is a standard motivation for stochastic mini-max settings in Robust Control.

The anisotropy-based approach to stochastic robust control occupies a unifying intermediate position between the \( H_2 \) and \( H_\infty \)-optimal control theories. Initiated about thirteen years ago at the interface of Information Theory and Robust Control, the approach employs the concepts of mean anisotropy of signals and anisotropic norm of systems [1], [2], [3].

The mean anisotropy of a multidimensional random signal is defined via its relative entropy, or historically, the Kullback-Leibler informational divergence, with respect to Gaussian white noises with scalar covariance matrices [4]. It therefore quantifies both temporal correlations, that is, “colouredness” or predictability, of the signal and its spatial “non-roundness”.

The \( \alpha \)-anisotropic norm of a system is the worst-case sensitivity of its output measured by the largest root mean square gain of the system with respect to input random disturbances whose mean anisotropy is bounded by a nonnegative parameter \( \alpha \).

The anisotropy-based approach to controller design employs the \( \alpha \)-anisotropic norm of the closed-loop system as a performance index which is to be minimized, with the magnitude of the parameter \( \alpha \) governing the robustness, and hence, conservativeness of the controller.

Since the exogenous perturbations are often caused by superposition of various effects from other interacting systems, the present paper provides a collection of results on the changes in the mean anisotropy of signals propagating through filter connections.

In this context, a leading part is played by feedback since the latter constitutes a universal vehicle for creating the temporal correlations in a signal via “recycling” its past history. Furthermore, the role of feedback for the mean anisotropy of signals is closely related to the Kolmogorov-Szegö formula [5] for the Shannon entropy production rate in a stationary Gaussian sequence.

Revisiting the underlying definitions in Sec. II, the paper dedicates Sec. III to the role of feedback in the construct of mean anisotropy of signals and discusses the propagation of the latter through various types of filter connections in Sec. IV.

The results of the paper can be used to support physical and engineering intuition for a “rational” choice of the anisotropy level \( \alpha \) in the design of anisotropy-based robust controllers.

II. MEAN ANISOTROPY AND ANISOTROPIC NORM

Let \( V = (v_k)_{k \in \mathbb{Z}} \) be an \( m \)-dimensional discrete-time Gaussian white noise, that is, a sequence of independent Gaussian random vectors with zero mean \( \mathbf{E}v_k = 0 \) and identity covariance matrix \( \mathbf{E}(v_kv_k^T) = I_m \). Here, \( \mathbb{Z} \) is the set of integers which label equidistant moments of discrete time.

Consider the \( m \)-dimensional stationary Gaussian signal \( W = (w_k)_{k \in \mathbb{Z}} = GV \) generated from \( V \) by a linear discrete time invariant causal system \( G \) with impulse response function \( 0 \leq k \mapsto g_k \in \mathbb{R}^{m \times m} \) as

\[
w_j = \sum_{k=0}^{+\infty} g_k v_{j-k}\]

The generating filter \( G \) is identified with its matrix transfer function

\[
G(z) = \sum_{k=0}^{+\infty} g_k z^{-k}
\]
which is assumed to be in the Hardy space $H^2_{m \times m}$, endowed with the $H_2$-norm
\[ \|G\|_2 = \sqrt{\sum_{k \geq 0} \|g_k\|_2^2} = \sqrt{\mathbf{E}(|w_j|^2)}, \]

extending the Frobenius norm $\|M\|_2 = \sqrt{\text{Tr}(MM^T)}$ of a finite-dimensional real matrix $M$. The mean anisotropy [2] of the signal $W = GV$ is defined by
\[ \mathcal{A}(G) = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \det\left( \frac{m \mathcal{G}(\omega)(\mathcal{G}(\omega))^*}{\|G\|_2^2} \right) \, d\omega \tag{2} \]

where $\mathcal{G}(\omega) = \lim_{\rho \to 1^+} G(\rho e^{i\omega})$ is the angular boundary value of the transfer function (1), with the limit understood in the sense of the $L_2$-norm, so that $\mathcal{G}$ is the Fourier transform of the impulse response function.

Note that (2) extends to systems $G \in H^m_{2 \times r}$ of full row rank $m \leq r$. More precisely, $0 \leq \mathcal{A}(G) < +\infty$ iff rank $G(\omega) = m$ for almost all $\omega \in [-\pi, \pi]$, in which case the mean anisotropy $\mathcal{A}(G)$ is completely specified by the singular values of $G$; otherwise, $\mathcal{A}(G) = +\infty$.

Furthermore, $\mathcal{A}(G) = 0$ iff $G$ is an all-pass system up to a nonzero constant scalar multiplier, that is, $G(\omega)(\mathcal{G}(\omega))^* = \lambda I_m$ for almost all $\omega \in [-\pi, \pi]$ for some $\lambda > 0$, or equivalently, iff $W$ is a Gaussian white noise with scalar covariance matrix $E(w_k w'_k) = \lambda I_m$.

In general, the mean anisotropy $\mathcal{A}(G)$ can be equivalently defined as the minimal relative entropy production rate of $W = GV$ with respect to the family of Gaussian white noises with scalar covariance matrices [4, pp. 1266–1269]. More precisely,
\[ \mathcal{A}(G) = \lim_{N \to +\infty} \min_{\lambda > 0} \frac{D(P_N||Q_{mN,\lambda})}{N}, \]

Here, $P_N$ is the probability distribution of the $mN$-dimensional random vector $(w_k)_{0 \leq k < N}$, the fragment of the signal $W$ on the time interval $\{0, \ldots, N-1\}$, and $Q_{mN,\lambda}$ is the $r$-variate Gaussian probability measure with zero mean and covariance matrix $\lambda I_r$, whose probability density function is given by
\[ q_{mN,\lambda}(w) = (2\pi \lambda)^{-r/2} \exp\left(-\frac{|w|^2}{2\lambda}\right), \quad w \in \mathbb{R}^r. \]

Accordingly, writing $p_N$ for the density of $P_N$ with respect to the $mN$-variate Lebesgue measure,
\[ D(P_N||Q_{mN,\lambda}) = \int_{\mathbb{R}^{mN}} p_N(w) \ln \left( \frac{p_N(w)}{q_{mN,\lambda}(w)} \right) \, dw, \]

is the relative entropy, or Kullback-Leibler informational divergence, of $P_N$ with respect to $Q_{mN,\lambda}$.

For a given anisotropy level $a \geq 0$, the $a$-anisotropic norm of a linear discrete time invariant system $F$ in $H^m_{\infty \times m}$ is defined by
\[ \|F\|_a = \sup_{G \in H^m_{2 \times m}: \mathcal{A}(G) \leq a} \frac{\|FG\|_2}{\|G\|_2}, \tag{3} \]

The quantity $\|F\|_a$ is nondecreasing and concave in $a \geq 0$, with
\[ \frac{\|F\|_2}{\sqrt{m}} = \|F\|_0 \leq \lim_{a \to +\infty} \|F\|_a = \|F\|_{\infty}, \]

so that the standard $H_2$ and $H_{\infty}$-norms are the limiting cases of the anisotropic norm.

A more detailed account on the properties of the mean anisotropy (2) of stationary Gaussian sequences and the anisotropic norm (3) of linear time invariant systems can be found in [6], [7].

### III. FEEDBACK IN MEAN ANISOTROPY

Although it is not obvious from (2), the construct of mean anisotropy incorporates feedback as a universal mechanism for creating temporal correlations in a signal via “recycling” its past history. The property originates from the Kolmogorov-Szegö formula for the Shannon entropy production rate in a stationary Gaussian sequence [5].

To this end, a signal $W = GV$ generated by a system $G \in H^m_{2 \times m}$ as described in Sec. II is split into a $W$-predictable component $\hat{W} = (\hat{w}_k)_{k \in \mathbb{Z}}$ and an innovation $\tilde{W} = (\tilde{w}_k)_{k \in \mathbb{Z}}$ as
\[ W = \hat{W} + \tilde{W}. \tag{4} \]

Here,
\[ \tilde{w}_k = E(w_k | (w_j)_{j < k}) \]
is the one-step predictor of $W$ based on its past history by time $k$, with $E(\cdot | \cdot)$ standing for the conditional expectation. Therefore,
\[ \hat{W} = SW, \quad \tilde{W} = MV, \tag{5} \]

where $S$ is a strictly causal system, and $M \in \mathbb{R}^{m \times m}$ is a constant matrix interpreted as a memoryless system. From (4) and (5), the generating filter is representable as
\[ G = (I_m - S)^{-1}M, \tag{6} \]

so that $S$ specifies the feedback law whereby the past history of $W$ is processed prior to getting “mixed” with the innovation signal $SV$; see Fig. 1.

![Fig. 1. The prediction-innovation representation of the generating filter.](image)
Proof: At any time \( k \), the predictor \( \hat{\bar{w}}_k \) and the innovation \( \bar{w}_k \) are independent zero mean Gaussian random vectors whose covariance matrices, in view of (4), add up to

\[
\mathbf{E}(\bar{w}_k \bar{w}_k^T) = \mathbf{E}(\hat{\bar{w}}_k \hat{\bar{w}}_k^T) + \mathbf{E}(\bar{w}_k \bar{w}_k^T).
\]  

(8)

The variances of the signals \( W = G\nu \) and \( \hat{W} \) and the covariance matrix of \( \hat{W} \) are given by

\[
\begin{align*}
\mathbf{E}(|\bar{w}_k|^2) &= ||G||_2^2, \\
\mathbf{E}(|\hat{\bar{w}}_k|^2) &= ||SG||_2^2, \\
\mathbf{E}(|\bar{w}_k|^2) &= MMM^T,
\end{align*}
\]

(9) \hspace{1cm} (10) \hspace{1cm} (11)

where we have also used (5). Substituting (9) and (11) into the Kolmogorov-Szegö type representation of the mean anisotropy [3, 6] yields

\[
\mathcal{A}(G) = -\frac{1}{2} \ln \det \left( mMMM^T \right). 
\]

(12)

Now, taking the trace of the covariance matrices on both sides of (8) and using (9)–(11) gives

\[
||G||_2^2 = ||SG||_2^2 + ||M||_2^2 \
\leq \left( ||S||_{\mathcal{H}\infty norm} ||G||_2 \right)^2 + ||M||_2^2.
\]

Here, the inequality is obtained by applying the definition (3) to \( S \) and implies the upper bound

\[
||G||_2^2 \leq \frac{||M||_2^2}{1 - ||S||_{\mathcal{H}\infty norm}^2}, 
\]

(13)

where the positiveness of the denominator is ensured by the assumption \( ||S||_{\infty} < 1 \) and by the property that the anisotropic norm is always majorized by the \( \mathcal{H}\infty \)-norm. Substitution of (13) into the right hand side of (12) gives (7).

It is worth noting the recursiveness of the inequality (7) in that the upper bound on the mean anisotropy \( \mathcal{A}(G) \) involves the \( \mathcal{A}(G) \)-anisotropic norm of the prediction operator \( S \) which is a subsystem of the generating filter \( G \); see Fig. 1. The inequality can be weakened by replacing the anisotropic norm with \( ||S||_{\infty} \).

If the feedback is “weak”, \( ||S||_{\infty} \ll 1 \), so that the temporal correlations in the signal \( W = G\nu \) are insignificant, then its mean anisotropy is due basically to the relative gap between the extreme singular values of the matrix \( M \).

A. State-space representation

As an example of an explicit representation for the aforementioned prediction operator \( S \), let the generating filter \( G \), which produces \( W = G\nu \), have an \( n \)-dimensional state \( X = (x_k)_{k \in \mathbb{Z}} \), and let the signals \( X \) and \( W \) be governed by

\[
\begin{align*}
x_{k+1} &= Ax_k + Bu_k, \\
w_k &= Cx_k + Du_k.
\end{align*}
\]

(14) \hspace{1cm} (15)

Here, \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \), \( C \in \mathbb{R}^{m \times n} \) and \( D \in \mathbb{R}^{m \times m} \) are constant matrices, with \( A \) assumed asymptotically stable, so that its spectral radius satisfies \( \rho(A) < 1 \). The state-space representation (14)–(15) is written as

\[
\begin{bmatrix}
A \\
B \\
C \\
D
\end{bmatrix}
\]

(16)

Theorem 2: Let the matrix \( D \) in (15) be nonsingular, and let both matrices \( A \) and \( A - BD^{-1}C \) be asymptotically stable. Then the generating filter (16) is representable in the prediction-innovation form (6) where the memoryless part is given by

\[
M = D,
\]

(17)

and the prediction operator \( S \) is described by the state-space realization

\[
S = \begin{bmatrix}
A - BD^{-1}C & BD^{-1} \\
C & 0
\end{bmatrix}.
\]

(18)

The mean anisotropy of the generating filter can be computed as

\[
\mathcal{A}(G) = -\frac{1}{2} \ln \det \left( \frac{mDD^T}{\text{Tr}(CPC^T + DD^T)} \right),
\]

(19)

where \( P \) is the controllability gramian of (16) satisfying the matrix algebraic Lyapunov equation

\[
P = APA^T + BBT.
\]

(20)

Proof: By (14), the operator \( V \rightarrow X \) is strictly causal and, therefore, the state signal \( X \) is \( V \)-predictable. The nonsingularity of \( D \) and asymptotic stability of \( A - BD^{-1}C \) imply that the filter \( G \) is invertible, with its inverse filter described by

\[
G^{-1} = \begin{bmatrix}
A - BD^{-1}C & BD^{-1} \\
-DBD^{-1} & D^{-1}
\end{bmatrix}.
\]

The last state-space realization is obtained by expressing \( v_k \) in terms of \( x_k \) and \( w_k \) from (15),

\[
v_k = -D^{-1}C x_k + D^{-1} w_k,
\]

and then by substituting the equation for \( v_k \) into (14) and regrouping the terms,

\[
x_{k+1} = Ax_k + B(-D^{-1}C x_k + D^{-1} w_k) \\
= (A - BD^{-1}C) x_k + BD^{-1} w_k,
\]

(21)

which also shows that \( G^{-1} \) and \( G \) share the state \( X \). Hence, the latter is not only \( V \)-predictable but is also \( W \)-predictable. Combining (15) with (21), one verifies that the prediction-innovation decomposition of the signal \( W \) in (4) holds with

\[
\tilde{W} = CX = SW, \quad \hat{W} = DV,
\]

where \( S \) is a strictly causal system whose state-space realization is described by (18). To establish (19), it now remains to use the Kolmogorov-Szegö type representation (12) together with (17), and to recall the relation

\[
||G||_2^2 = \mathbf{E}(|\bar{w}_k|^2 + |\hat{\bar{w}}_k|^2) \\
= \text{Tr}(CPC^T + DD^T),
\]

for \( G \).
which employs the property that the controllability gramian $P$ of the generating filter (16) is the covariance matrix of the state signal, $P = \mathbb{E}(x_k x_k^T)$.

An alternative state-space representation of the mean anisotropy can be found in [6, Theorem 1 on p. 31] which does not require the asymptotic stability of $A - BD^{-1}C$ but, in addition to (20), employs the stabilizing solution of a matrix algebraic Riccati equation.

**IV. MEAN ANISOTROPY IN FILTER CONNECTIONS**

We will now discuss the results on the changes in the mean anisotropy of signals propagating through several types of filter connections.

**A. Serial connection**

Let $W = G_N \times \ldots \times G_1 V$ be generated from the white noise $V$ by $N$ serially connected filters as shown in Fig. 2.

\[
\begin{array}{c}
W \\
G_N \quad \ldots \quad G_1 \\
\end{array} \quad V
\]

Fig. 2. The serial connection of filters.

**Lemma 1:** For the serial connection of generating filters $G_1 \in \mathcal{H}_2^{m \times m}$ and $G_2, \ldots, G_N \in \mathcal{H}_\infty^{m \times m}$, its mean anisotropy satisfies

\[
\overline{A}(G_N \times \ldots \times G_1) \leq a_N,
\]

where the nonnegative reals $a_1, \ldots, a_N$ are computed by the recurrence equation

\[
a_k = \overline{A}(G_k) + a_{k-1} + m \ln \frac{||G_k||_{a_{k-1}}}{||G_k||_0}
\]

with initial condition $a_0 = 0$.

Proof of Lemma 1 can be found in [6] and employs an “anisotropy leverage” identity

\[
\overline{A}(FG) = \overline{A}(F) + \overline{A}(G) + m \ln \frac{\sqrt{m}||FG||_2}{||F||_2||G||_2},
\]

which follows from (2) and holds for any systems $F, G \in \mathcal{H}_2^{m \times m}$ satisfying $FG \in \mathcal{H}_2^{m \times m}$. Hence, in application to a system $F \in \mathcal{H}_\infty^{m \times m}$, combining the identity (23) with (3) gives

\[
\sup_{G \in \mathcal{H}_2^{m \times m} : \overline{A}(G) \leq a} \overline{A}(FG) = \overline{A}(F) + a + m \ln \frac{||F||_a}{||F||_0},
\]

which holds for any $a \geq 0$ and underlies the recurrence (22); see also, [6, Theorem 2 on p. 32].

Following [7, Lemma 3 on p. 13], to formulate another corollary of (23), we define the condition number of an invertible filter $G \in \mathcal{H}_2^{m \times m}$ by

\[
\text{cond}(G) = \frac{||G||_2||G^{-1}||_2}{m}.
\]

The latter satisfies $\text{cond}(G) \geq 1$, with the inequality turning to equality iff $G$ is an all-pass system up to a nonzero constant scalar multiplier.

**Lemma 2:** For an invertible filter $G \in \mathcal{H}_2^{m \times m}$ and its inverse $G^{-1}$, their mean anisotropies satisfy

\[
\overline{A}(G) + \overline{A}(G^{-1}) = m \ln \text{cond}(G).
\]

Proof: Applying the identity (23) to $F = G^{-1}$, we have

\[
0 = \overline{A}(I_m) = \overline{A}(G^{-1}) + \overline{A}(G) + m \ln \frac{\sqrt{m}||I_m||_2}{||G^{-1}||_2||G||_2},
\]

whence (24) follows immediately in view of the relation $||I_m||_2 = \sqrt{m}$.

Since the mean anisotropies $\overline{A}(G)$ and $\overline{A}(G^{-1})$ are both nonnegative, $m \ln \text{cond}(G)$ is an upper bound for each of them.

**B. Summation of independent signals**

Let $W = \sum_{s=1}^N W_s$ be obtained by summation of $N$ signals $W_s = G_s V_s$ generated from mutually independent white noises $V_1, \ldots, V_N$ by filters $G_1, \ldots, G_N \in \mathcal{H}_2^{m \times m}$; see Fig. 3. Such $W$ can be equivalently modeled as $W = GV$, where the “effective” generating filter $G \in \mathcal{H}_2^{m \times m}$ satisfies the factorization

\[
\hat{G} \hat{G}^* = \sum_{s=1}^N \hat{G}_s \hat{G}_s^*.
\]

The latter implies that $||G||_2 = \sum_{s=1}^N ||G_s||_2^2$, and hence, the ratio

\[
R_s = \left( \frac{||G_s||_2}{||G||_2} \right)^2
\]

quantifies the relative power contribution of $W_s$ to the resulting signal $W$.

**Lemma 3:** The mean anisotropy of the signal $W = GV$ generated by a filter $G \in \mathcal{H}_2^{m \times m}$ satisfying (25) affords the upper bound

\[
\overline{A}(G) \leq \min \left( \sum_{s=1}^N R_s \overline{A}(G_s) \right),
\]

\[
\min_{1 \leq s \leq N} \left( \overline{A}(G_s) - \frac{m}{2} \ln R_s \right).
\]

Proof of Lemma 3 can also be found in [6, pp. 33–34] and employs the concavity and monotonicity of $\ln \text{det}(\cdot)$ on the cone of positive definite Hermitian matrices. The right hand side of (27) describes two upper bounds for $\overline{A}(G)$ which complement each other due to an interplay between
the relative power contributions \( R_s \) in (26) and the “partial” mean anisotropies \( \overline{A}(G_s) \).

C. Multiplex connection

Consider the multiplex connection of \( N \) filters \( G_s \in \mathcal{H}^{m_s \times m_s}_\infty \) in Fig. 4, where \( V_1, \ldots, V_N \) are mutually independent white noises which can be combined into one white noise \( V \) of dimension \( m = m_1 + \ldots + m_N \). The resultant generating filter \( G = G_1 \otimes \ldots \otimes G_N \in \mathcal{H}^{m \times m}_\infty \) which produces \( W = GV \) has a block diagonal matrix transfer function

\[
G = \begin{bmatrix}
G_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
0 & \cdots & G_N & 0
\end{bmatrix}.
\]

The latter implies that \( \det G = \prod_{s=1}^N \det G_s \) and, similarly to the summation of independent signals in Sec. IV-B, \( \|G\|_2^2 = \sum_{s=1}^N \|G_s\|_2^2 \). Substitution of these expressions into (2) gives

\[
\overline{A}(G) = \sum_{s=1}^N \left( \overline{A}(G_s) - m_s \ln \frac{\|G_s\|_0}{\|G_s\|_0} \right).
\]

D. Feedback connection

Let \( W = GV \) be produced from the white noise \( V \) by the generating filter

\[
G = (I_m - G_0 \Delta)^{-1} G_0
\]

which is obtained from a “nominal” filter \( G_0 \in \mathcal{H}^{m \times m}_\infty \) by a strictly causal feedback perturbation \( \Delta \in \mathcal{H}^{m \times m}_\infty \) interpreted as an uncertainty in the noise model; see Fig. 5. By the

\[
\begin{array}{c}
W \\
\downarrow \\
G_0 \\
\uparrow \\
\downarrow \\
\Delta \\
\uparrow \\
V
\end{array}
\]

Fig. 5. Feedback perturbation of the nominal filter.

Small Gain Theorem, the linear fractional transformation on the right hand side of (28) is well-defined if

\[
\|G_0\|_\infty \|\Delta\|_\infty < 1.
\]

Assuming that \( G_0 \) is given in the prediction-innovation form (6), that is,

\[
G_0 = (I_m - S)^{-1} M,
\]

the appropriate representation for the perturbed generating filter (28) is obtained by an additive modification of the prediction operator \( S \), so that

\[
G = (I_m - T)^{-1} M,
\]

where

\[
T = S + M \Delta.
\]

If the nominal filter \( G_0 \) is known and invertible, then \( W = G_0 U \), with all the noise model uncertainty due to the perturbation \( \Delta \) incorporated in the signal \( U = FV \). The latter is produced from the white noise \( V \) by the generating filter

\[
F = G_0^{-1} G = (I_m - \Delta G_0)^{-1}
\]

whose well-posedness too is ensured by (29). Since \( \Delta G_0 \) inherits the strict causality from \( \Delta \) and is contractive, application of Theorem 1 with \( S = \Delta G_0 \) and \( M = I_m \) to the system (30) yields

\[
\overline{A}(F) \leq -\frac{m}{2} \ln \left( 1 - \|G_0\|_\infty^2 \right)
\]

\[
\leq -\frac{m}{2} \ln \left( 1 - \|G_0\|_\infty^2 \|\Delta\|_\infty^2 \right)
\]

where use has also been made of the submultiplicative property of the \( \mathcal{H}_\infty \)-norm.

ACKNOWLEDGEMENT

The authors thank Dr E.A. Maksimov for helpful discussions.

REFERENCES


