On Data-Driven Control: Informativity of Noisy Input-Output Data With Cross-Covariance Bounds

Citation for published version (APA):

DOI:
10.1109/LCSYS.2021.3139526

Document status and date:
Published: 01/01/2022

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 26. Oct. 2023
On data-driven control: informativity of noisy input-output data with cross-covariance bounds

Tom R.V. Steentjes, Mircea Lazar, Paul M.J. Van den Hof

Abstract—In this paper we develop new data informativity based controller synthesis methods that extend existing frameworks in two relevant directions: a more general noise characterization in terms of cross-covariance bounds and informativity conditions for control based on input-output data. Previous works have derived necessary and sufficient informativity conditions for noisy input-state data with quadratic noise bounds via an S-procedure. Although these bound do not capture cross-covariance bounds in general, we show that the S-procedure is still applicable for obtaining non-conservative conditions on the data. Informativity-conditions for stability, $H_\infty$ and $H_2$ control are developed, which are sufficient for input-output data and also necessary for input-state data. Simulation experiments illustrate that cross-covariance bounds can be less conservative for informativity, compared to norm bounds typically employed in the literature.

Index Terms—Data-driven control, LMIs, linear systems

I. INTRODUCTION

When mathematical models of dynamical systems are not available, data plays an essential role in the process of learning system characteristics. Indeed, data can contain information about the system from which a model of the system can be derived or a controller can be learned, either from a data-based model or directly from the data. A key problem for data-driven control is to determine whether a set of data collected from a system contains enough information to design a controller, independent of the methodology.

An indirect approach for controller design from data consists of two steps: obtaining a model from data through system identification [1] and subsequently designing a controller via a model-based method. In the field of identification for control, the problem of determining a suitable model for controller design is considered [2], aiming at minimizing performance degradation due to model mismatching. If the data used for obtaining a model are sufficiently rich for identification, is determined by a property called informativity.

Even if data are not informative for identification, data can still be informative for controller design. Necessary and sufficient conditions for informativity of data for control were developed in [3] for noiseless input-state data. These results were extended in [4] for noisy input-state data with prior knowledge on the noise in the form of quadratic bounds, via a matrix variant of the S-procedure. Quadratic noise bounds play a key role in data-driven controller design [4], [5], [6], distributed controller design [7] and dissipativity analysis [8], [9] from data, and represent, for example, magnitude, energy and variance bounds on the noise.

In this paper, we consider the problem of determining informativity of input-output and input-state data for control with prior knowledge of the noise in the form of a sample cross-covariance type bound with respect to a user-chosen instrumental signal. Bounds on the sample cross-covariance were introduced in [10] as an alternative to magnitude bounds in parameter bounding identification, given its overly conservative noise characterization, cf. [11] for a comparison of instantaneous and (quadratic) energy type bounds for data-driven control. Our approach to data-driven control extends existing frameworks in two relevant directions: a more general noise characterization, cf. [11] for a comparison of instantaneous and (quadratic) energy type bounds for data-driven control.

II. INPUT-OUTPUT DATA: CROSS-COVARIANCE BOUNDS

Consider a class of linear systems described by

$$A(q^{-1})y(t) = B(q^{-1})u(t) + e(t),$$

with $A(\xi) \in \mathbb{R}^{p \times p}[\xi]$ and $B(\xi) \in \mathbb{R}^{p \times m}[\xi]$ polynomial matrices, given by $A(\xi) = I + A_1\xi + A_2\xi^2 + \cdots + A_n\xi^n$ and $B(\xi) = B_0 + B_1\xi + B_2\xi^2 + \cdots + B_l\xi^l$ and $q^{-1}$ is the delay operator, i.e., $q^{-1}x(t) = x(t-1)$. By defining $\zeta(t) := \text{col}(y(t-1),\ldots,y(t-l),u(t-1),\ldots,u(t-l))$, $\bar{A} := \text{row}(-A_1,\ldots,-A_l)$ and $\bar{B} := \text{row}(B_1,\ldots,B_l)$, we can write (1) equivalently as

$$y(t) = B_0u(t) + e(t) + [\bar{A} \quad \bar{B}] \zeta(t),$$

Hence, with $\zeta \in \mathbb{R}^n$ as a state, a state-space representation is

$$\zeta(t+1) = \begin{bmatrix} \bar{A} & \bar{B} \\ I & 0 \end{bmatrix} \zeta(t) + \begin{bmatrix} B_0 \\ \text{0} \end{bmatrix} u(t) + \begin{bmatrix} I \\ \text{0} \end{bmatrix} e(t),$$

$$y(t) = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix} \zeta(t) + B_0u(t) + e(t),$$

Notice that (3) is a non-minimal representation of order $n = (p + m)l$. Defining the data matrices $Z := \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix}$ and $\zeta := \begin{bmatrix} (y(t-1),\ldots,(y(t-l),u(t-1),\ldots,u(t-l))}$, we can write (2) equivalently as

$$y(t) = Z\zeta(t),$$

with $Z \in \mathbb{R}^{m \times (n+1)}$. The main result of this paper is the following: if there exists an $\epsilon > 0$ such that

$$\text{trace}(Z^T Z) < \epsilon,$$

then the system is informatively controllable.

T.R.V. Steentjes, M. Lazar and P.M.J. Van den Hof are with the Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands. E-mails: {t.r.v.steentjes, m.lazar, p.m.j.vandenhof}@tue.nl

This work is supported by the European Research Council (ERC), Advanced Research Grant SYSDYNET, under the European Union’s Horizon 2020 research and innovation programme (grant No. 694504).
\[
\begin{align*}
\zeta(0) & \cdots \zeta(N-1), \quad Y_- := \begin{bmatrix} y(0) & \cdots & y(N-1) \end{bmatrix} \\
U_- & \quad E_- \quad \text{accordingly, we obtain the data equation}
\end{align*}
\]
\[Y_- = \begin{bmatrix} \bar{A} & \bar{B} \end{bmatrix} Z_- + B_0 U_- + E_-, \quad (4)\]
where \(\bar{A}, \bar{B}, B_0\) are unknown system matrices. We consider the noise not to be measured, i.e., \(E_-\) is unknown, while prior knowledge on the cross-covariance of the noise with respect to an instrumental variable is available.

### A. Cross-covariance noise bounds

Consider the sample cross-covariance with respect to the noise \(e \in \mathbb{R}^p\) and a variable \(r \in \mathbb{R}^M\), given by
\[
\frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} e(t) r(t)^T = \frac{1}{\sqrt{N}} E_- R_+.\]
The variable \(r\) is instrumental and can be specified by the user (as discussed at the end of this subsection), i.e., it is a given variable in the upcoming analysis. We assume prior knowledge on the noise of the form
\[
\frac{1}{N} E_- R_+^T R_- E_-^T \preceq H_\mu, \quad (5)
\]
where \(H_\mu\) is an upper-bound on the squared cross-covariance matrix \(\frac{1}{N} E_- R_+^T R_- E_-^T\). In a generalized form, we write
\[
\begin{bmatrix} I & R_- E_-^T \end{bmatrix}^T \begin{bmatrix} Q_{11} & Q_{12} \\
Q_{12} & Q_{22} \end{bmatrix} \begin{bmatrix} I & R_- E_-^T \end{bmatrix} \succeq 0, \quad (6)
\]
with \(Q_{22} < 0\). For \(Q_{11} = NH\) and \(Q_{12} = 0\) and \(Q_{22} = -I\), the bound (5) is recovered. Note that (6) can be rewritten as the bound in [4, Assumption 1] with \(\Phi_{22} = R_-^T Q_{22} R_-,\) but in general only \(\Phi_{22} \preceq 0\) holds, while \(\Phi_{22} < 0\) is assumed in [4]. This means that the data informativity results of [4] cannot be used to establish data informativity for general cross-covariance noise bounds. However, the matrix S-lemma in [4, Theorem 13] can still be exploited to obtain necessary conditions for informativity of input-state data with cross-covariance bounds, as shown in Proposition 1 of this paper.

In the state-space representation (5), the state-space matrices \(A_2\) and \(B_2\) contain unknown parameters \(\bar{A}, \bar{B}, \) and \(B_0\). Write
\[
A_2 = \begin{bmatrix} \bar{A} & \bar{B} \\
0 & 0 \\
0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\
0 & 0 & 0 \\
I_{p \times (l-1)} & 0 & 0 \end{bmatrix} =: A_e, \quad (7)
\]
and
\[
B_2 = \begin{bmatrix} B_0 \\
0 \\
0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\
0 & I_m \\
0 & 0 \end{bmatrix} =: B_e, \quad (8)
\]
so that \(A_e\) and \(B_e\) are unknown parameter matrices, concatenated with zero rows, and \(J_1\) and \(J_2\) are binary matrices. The set of all pairs \((A_e, B_e)\) that are compatible with the data is
\[
\Sigma_{\text{RE}}^R(U,Y) := \left\{ \begin{bmatrix} \bar{A} & \bar{B} \\
0 & 0 \end{bmatrix}, \begin{bmatrix} B_0 \\
0 \\
0 \end{bmatrix} \right\} \quad \text{such that (4) and (4) hold}. \]

#### Lemma 1

Let
\[
Q_e := \begin{bmatrix} H_2 Q_{11} H_2^T & H_2 Q_{12} Q_{22} \end{bmatrix}
\]
and consider
\[
\begin{bmatrix} I \\
A_e^T \\
B_e^T \end{bmatrix}^T \begin{bmatrix} I & H_2 Y_+ R_+^T \\
0 & -Z_+ R_+^T \end{bmatrix} Q_e \begin{bmatrix} I & H_2 Y_+ R_+^T \\
0 & -Z_+ R_+^T \end{bmatrix} \begin{bmatrix} I \\
A_e^T \\
B_e^T \end{bmatrix} \succeq 0. \quad (9)
\]
It holds that \(\Sigma_{\text{RE}}^R(U,Y) \subseteq \{ (A_e, B_e) \mid (9) \text{ holds} \} \). Moreover, if \(R_- [Z_+^T U_+^T] \) has full column rank, then \(\Sigma_{\text{RE}}^R(U,Y) = \{ (A_e, B_e) \mid (9) \text{ holds} \} \).

#### Proof

The first statement follows from (4), and the definition of \(\Sigma_{\text{RE}}^R(U,Y)\). Full-column rank of \(R_- [Z_+^T U_+]\) and (4) imply that the last \(ml + p(l-1)\) rows of \((A_e, B_e) = \text{col}(M_1, 0)\) are zero and \(E_- := Y_- - M_1 [Z_+^T U_+]^T\) satisfies (6). Hence, \((A_e, B_e) \in \Sigma_{\text{RE}}^R(U,Y)\) so that \(\Sigma_{\text{RE}}^R(U,Y) = \{ (A_e, B_e) \mid (9) \text{ holds} \} \).

We denote all \((A_2, B_2)\) that are compatible with the data by \(\Sigma_{\text{RE}}^R(U,Y) := \{ (A_e + J_1, B_e + J_2) \mid (A_e, B_e) \in \Sigma_{\text{RE}}^R(U,Y) \}\).

We have provided a parametrization of (a superset of) \(\Sigma_{\text{RE}}^R(U,Y)\) based on the data equation (4). One can equivalently parametrize \(\Sigma_{\text{RE}}^R(U,Y)\) on the basis of the state data equation
\[
Z_+ = A_2 Z_- + B_2 U_- + H_2 E_-.
\]
This leads to an equal set \(\Sigma_{\text{RE}}^R(U,Y)\), but the ‘repeated’ data in the parametrization contained in \(Z_+ := [\zeta(1) \cdots \zeta(N)]\), would render the evaluation numerically sensitive. Design methods in [4], [6] can yield a parametrization \(\Sigma_{\text{RE}}^R(U,Y)\) based on this state data equation, but with limited applicability to cross-covariance bounds (9), i.e., only if the dimension of \(r\) satisfies \(M \geq N\).

Existing guidelines [10] recommend choosing an instrumental variable \(r\) that is correlated with the input \(u\), but uncorrelated with the noise \(e\). Hence, this suggests the choice of (filtered/delayed versions) of the input for \(r\) in an open-loop case for data collection, and an external reference/dithering signal for \(r\) in a closed-loop case. Moreover, Lemma 1 provides an additional guideline for the choice of \(r\) to reduce conservatism in the case of input-output data, i.e., \(R_- [Z_+^T U_+]\) has full column rank only if the number of instrumental variables \(M\) satisfies \(M \geq pl + m(l+1)\).

### B. Output-feedback control

Consider a (dynamic) output feedback controller described by the difference equation of the form [5]
\[
C(q^{-1}) u(t) = D(q^{-1}) y(t), \quad (10)
\]
with \(C(\xi) \in \mathbb{R}^{m \times m}[\xi]\) and \(D(\xi) \in \mathbb{R}^{m \times p}[\xi]\) polynomial matrices given by \(C(\xi) = I + C_1 \xi + C_2 \xi^2 + \cdots + C_p \xi^p\) and \(D(\xi) = D_1 \xi + D_2 \xi^2 + \cdots + D_p \xi^p\). We define a state \(z_c\) for (10) as \(z_c := \text{col}(u(t-1), \ldots, u(t-l), y(t-l), \ldots, y(t-l))\), yielding a state-space representation for the controller:
\[
\zeta_c(t+1) = \begin{bmatrix} \tilde{C} & \tilde{D} \end{bmatrix} \zeta_c(t) + \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} y(t) \end{bmatrix}, \quad (11)
\]
with \(\tilde{C} := \text{row}(-C_1, \ldots, -C_p)\) and \(\tilde{D} := \text{row}(D_1, \ldots, D_p)\).
It follows that \(\zeta_c = [I \ 0] z_c\), which implies that \(u(t) = \begin{bmatrix} \tilde{C} & \tilde{D} \end{bmatrix} \zeta_c(t) = [D \ \tilde{C}] z_c(t)\). Hence, the closed-loop system described by (3) and (11) has a representation
\[
\zeta(t+1) = \begin{bmatrix} A + B_0 D & \tilde{B} + B_0 \tilde{C} \\
I & 0 \end{bmatrix} \zeta(t) + \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} y(t) \end{bmatrix} + \begin{bmatrix} 0 \end{bmatrix} e(t), \quad (11)
\]
where
With \( K := [\bar{D} \quad \bar{C}] \), the closed-loop system matrix \( A_{cl} \) satisfies \( A_{cl} = A_z + B_z K \). For some \((A_z, B_z) \in \Sigma_{(U, Y)}^{RQ}\), we say that the controller \((\bar{u}, \bar{y})\) stabilizes \((\bar{x}, \bar{z})\) if there exist a \( K \) and \( P > 0 \) so that

\[
\Sigma_{(U, Y)}^{RQ} = \{(A, B) \mid (A + BK)P(A + BK)^T - P < 0\}.
\]

By (7)-(8), we find that the existence of \( K \) and \( P > 0 \) so that \((A_z + B_z K)(A_z + B_z K)^T - P < 0\) is equivalent to the existence of \( K \) and \( P > 0 \) such that (13) holds true. For the data \((U, Y)\) to be informative for quadratic stabilization, we require the existence of \( K \) and \( P > 0 \) so that (13) holds true for all \((A_z, B_z) \in \Sigma_{(U, Y)}^{RQ}\). This is precisely a problem that can be solved by the S-procedure; more specifically, by the matrix-valued S-lemma [4].

**Theorem 1** The data \((U, Y)\) are informative for quadratic stabilization by output-feedback controller \((\bar{u}, \bar{y})\) if there exist \( L \in \mathbb{R}^{m \times n} \), \( P > 0 \), \( \alpha > 0 \), and \( \beta > 0 \) so that (14) holds true. Moreover, for \( L \) and \( P \) such that (14) is satisfied, \( A_{cl} = A_z + B_z K \) is stable for all \((A_z, B_z) \in \Sigma_{(U, Y)}^{RQ}\) with \( K := LP^{-1} \).

**Proof:** Let \( L \), \( P > 0 \), \( \alpha > 0 \), and \( \beta > 0 \) exist so that (14) holds true and consider the matrix \( \Pi \) defined in (13). By the Schur complement, (14) is equivalent to

\[
\Pi - \alpha \Lambda \succeq \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{where } K := LP^{-1} \text{ and } \\
\Lambda := \begin{bmatrix} I & H_y R_z^T \\ 0 & -R_z R_z^T \end{bmatrix} Q_e \begin{bmatrix} I & H_y R_z^T \\ 0 & -R_z R_z^T \end{bmatrix}^T.
\]

Hence, (13) holds true for all \((A_z, B_z) \in \Sigma_{(U, Y)}^{RQ}\), cf. [4, Theorem 13]. This concludes the proof.

We remark that if there is a \( Z \) so that \( Z^\top \Lambda Z > 0 \) with \( Z := \text{col}(I, Z) \), called the generalized Slater condition [4], then (14) is also a necessary condition for informativity of input-output data for quadratic stabilization, if \( R_- [Z^\top U_-^\top] \) has full column rank (Lemma 1). Unlike in the case of input-state data, which will be discussed next, we note that the generalized Slater condition can in general not hold true in the input-output case if \( l \geq 1 \), since the noise affects a subspace of the extended state space, yielding a degenerate matrix \( \Lambda \). The combination of noisy and noiseless states in \( \zeta \) suggests that necessity could potentially be proven in general by a ‘fusion’ of the matrix S-lemma and matrix Finsler’s lemma [12].

### III. INFORMATIVITY FOR STABILIZATION

#### A. Informativity of input-output data

**Definition 1** The data \((U, Y)\) are said to be informative for quadratic stabilization by output-feedback controller \((\bar{u}, \bar{y})\) if there exist a \( K \) and \( P > 0 \) so that

\[
\Sigma_{(U, Y)}^{RQ} = \{(A, B) \mid (A + BK)P(A + BK)^T - P < 0\}.
\]

We will now consider a special case, where input-state data is available instead of input output data. That is, we measure a state \( y(t) = x(t) \) and the class of systems considered is

\[
x(t + 1) = Ax(t) + Bu(t) + e(t),
\]

with the corresponding data equation

\[
X_+ = AX_- + BU_- + E_-.
\]

All systems that explain the data \((U_-, X)\) for some \( E_- \) satisfying the cross-covariance bound (6) are in the set

\[
\Sigma_{(U_, X)}^{RQ} = \{(A, B) \mid (A, B) \text{ satisfies (17)} \}.
\]

By (16), the set of feasible systems is

\[
\Sigma_{(U_, X)}^{RQ} = \{(A, B) \mid (A, B) \text{ satisfies (17)} \}.
\]

**Remark 1** Consider a specific selection of \( M = N \) instrumental variables defined by \( r_i(t) := \delta(t - i + 1), i = 1, \ldots, N, \) and \( \delta : \mathbb{Z} \to \{0, 1\} \) is the unit impulse defined as \( \delta(0) = 1 \) and \( \delta(x) = 0 \) for \( x \in \mathbb{Z} \setminus \{0\} \). It follows that \( R_- = I \) for this choice of instrumental signals. Then, with the generalized quadratic cross-covariance bound (6), we observe that for this special choice \( R_- = I \), we recover the set of feasible systems in [4], and, hence, the informativity conditions in [4].

**Definition 2** The data \((U_-, X)\) are said to be informative for quadratic stabilization by state feedback if there exist a feedback gain \( K \) and \( P > 0 \) so that

\[
\Sigma_{(U_-, X)}^{RQ} = \{(A, B) \mid (A + BK)P(A + BK)^T - P < 0\}.
\]

We will now provide a necessary and sufficient condition for informativity of input-state data for quadratic stabilization, given prior knowledge on the cross-covariance [6]. Consider the generalized Slater condition

\[
\begin{bmatrix} I \\ Z \end{bmatrix} A X \begin{bmatrix} I \\ Z \end{bmatrix} > 0.
\]

**Proposition 1** Suppose that there exists a \( Z \) so that (18) holds true. Then the data \((U_-, X)\) are informative for quadratic stabilization if

\[
\begin{bmatrix} I \\ A_{cl}^\top \\ B_{cl}^\top \end{bmatrix} P - (J_1 + J_2 K)P(J_1 + J_2 K)^T - (J_1 + J_2 K)P - (J_1 + J_2 K)PK^T - PK - -PK^T \geq 0
\]

\[
= 0
\]


stabilization if and only if there exist $L \in \mathbb{R}^{m \times n}$, $P > 0$, $\alpha \geq 0$ and $\beta > 0$ so that
\[
\begin{bmatrix}
P - \beta I & 0 & 0 & 0 \\
0 & -P & -L^T & 0 \\
0 & -L & 0 & L \\
0 & 0 & L^T & P
\end{bmatrix} - \alpha \begin{bmatrix}\Lambda_X & 0 \\
0 & 0 & 0
\end{bmatrix} \geq 0.
\tag{19}
\]
Moreover, $K$ is such that $A + BK$ is stable for all $(A, B) \in \Sigma^{RQ}_{(U, X)}$ if $K := LP^{-1}$ with $L$ and $P > 0$ satisfying (19).

Proof: $(\Rightarrow)$ Let the data be informative for quadratic stabilization, i.e., there exist $K$ and $P > 0$ so that, with $\Pi$ defined in (13) with $J_1 = 0$, $J_2 = 0$:
\[
\begin{bmatrix} I \\ A^T \\ B^T \end{bmatrix}^T \Pi (\ast) > 0 \text{ for all } (A, B)
\]
with $\Lambda_X = \begin{bmatrix} \Lambda_{X1} \\ \Lambda_{X2} \end{bmatrix}$ := $\begin{bmatrix} I \\ X^T \end{bmatrix}^T \begin{bmatrix} X & 1 \end{bmatrix} = 0$.

We will now show that $\ker \Lambda_{X2} \subseteq \Lambda_{X1}$, such that necessity follows by the matrix S-lemma [4]. First, notice that $\ker \Lambda_{X2} = \ker R_z [X^T \\ U^T]$. Now, take any $x \in \ker \Lambda_{X2}$, then $R_z [X^T \\ U^T]x = 0$. Clearly, we have that $(X^T \gamma X + Q_{z22} + Q_{12})R_z [X^T \\ U^T]x = 0$, which implies that $x \in \ker \Lambda_{X2}$. Since $x \in \ker \Lambda_{X2}$ was arbitrary, this shows that $\ker \Lambda_{X2} \subseteq \Lambda_{X1}$. By $\ker \Lambda_{X2} \subseteq \Lambda_{X1}$ and (13), there exist $\alpha \geq 0$ and $\beta > 0$ so that, by [4, Theorem 13]:
\[
\Pi - \alpha \Lambda_X \succeq \begin{bmatrix} \beta I & 0 \\ 0 & 0 \\
\end{bmatrix},
\]
which is equivalent to (19) for $L := KP$ by the Schur complement. This completes the proof.

IV. INCLUDING PERFORMANCE SPECIFICATIONS

We will now consider the problem of finding a controller (10) for which the closed-loop system achieves an $\mathcal{H}_\infty$ or $\mathcal{H}_2$ performance bound from the input-output data $(U, Y)$. Consider the performance output $z$, given by $z(t) = C_z \zeta(t) + D_z u(t)$. For any pair $(A_z, B_z)$, the controller (10) yields the closed loop system
\[
\zeta(t + 1) = (A_z + B_z K) \zeta(t) + H_z e(t),
\]
\[
z(t) = (C + D K) \zeta(t).
\]
Hence, the transfer matrix from $e$ to $z$ is given by
\[
T(z) := (C_z + D_z K) (z I - A_z - B_z K)^{-1} H_z,
\]
for which the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norm are denoted $\|T\|_{\mathcal{H}_\infty}$ and $\|T\|_{\mathcal{H}_2}$, respectively.

For given $K$, the $\mathcal{H}_\infty$ norm of $T$ is less than $\gamma$, $\|T\|_{\mathcal{H}_\infty} < \gamma$, if and only if there exists $X > 0$ such that [13, p. 125]
\[
\begin{bmatrix} X & 0 & A_z^T X & C_z^T \\
0 & \gamma I & H_z X & 0 \\
X A_K & X H_z & X & 0 \\
C_K & 0 & 0 & \gamma I
\end{bmatrix} \succ 0,
\tag{20}
\]
where $A_K := A_z + B_z K$ and $C_K := C_z + D_z K$.

Definition 3 The data $(U, Y)$ are said to be informative for common $\mathcal{H}_\infty$ control by output-feedback controller (10) with performance $\gamma$ if there exist $K$ and $X > 0$ so that
\[
\Sigma^{RQ}_{(U, Y)} \subseteq \{(A_z, B_z) \mid (20) \text{ holds true}\}.
\]

Theorem 2 The data $(U, Y)$ are informative for common $\mathcal{H}_\infty$ control with performance $\gamma$ if there exist $L \in \mathbb{R}^{m \times n}$, $P > 0$, $\alpha \geq 0$ and $\beta > 0$ so that (21) holds true.

Proof: By a congruence transformation of (20) with $\text{diag}(P, I, P, I)$ with $P := X^{-1}$ and the application of the Schur complement (twice), the existence of $K$ and $X > 0$ so that (20) holds, is equivalent to the existence of $P$ and $L$ so that $P > 0$ and
\[
P - V_z (P - \gamma^{-1} F^T F)^{-1} V_z^T - \gamma^{-1} H_z H_z^T > 0
\tag{22}
\]
and $P - \gamma^{-1} F^T F > 0$, where $V_z := A_z P + B_z L$ and $F := C_z P + D_z L$. We can now rewrite (22) as
\[
\begin{bmatrix} I \\ A_z^T \\ B_z^T \end{bmatrix}^T \Pi \mathcal{H}_\infty \begin{bmatrix} I \\ A_z^T \\ B_z^T \end{bmatrix} := \begin{bmatrix} I \\ A_z^T \\ B_z^T \end{bmatrix}^T \Pi \mathcal{H}_\infty \begin{bmatrix} I \\ A_z^T \\ B_z^T \end{bmatrix} \succeq 0.
\tag{23}
\]
Hence, the data $(U, Y)$ are informative for common $\mathcal{H}_\infty$ control with performance $\gamma$ if and only if there exist $P > 0$ and $L$ such that $P - \gamma^{-1} F^T F > 0$ and (23) holds for all $(A_z, B_z) \in \Sigma^{RQ}_{(U, Y)}$. By assumption, there exist $P > 0$, $L$, $\alpha \geq 0$ and $\beta > 0$ such that (21) holds true. By the Schur complement, (21) is equivalent to $\Pi \mathcal{H}_\infty - \alpha \Lambda \succeq \begin{bmatrix} \beta I \\ 0 \\
\end{bmatrix}$, which implies that (23) holds for all $(A_z, B_z) \in \Sigma^{RQ}_{(U, Y)}$. The conditions (21) are linear with respect to $P$, $L$, $\alpha$ and $\beta$. By a straightforward additional application of the Schur complement, (21) can also be made linear with respect to $\gamma$. 

\[
\begin{bmatrix}
P - \beta I & -J_1 P - J_2 L & 0 & J_1 P + J_2 L \\
-P J_1^T - L^T J_2 & -P & -L^T & 0 \\
0 & -L & 0 & L \\
P J_1^T + L^T J_2 & 0 & L^T & P
\end{bmatrix} - \alpha \begin{bmatrix} I \\ H_z Y \gamma R_z^T \\
0 & -\gamma R_z^T \\
0 & -U R_z^T \\
0 & 0
\end{bmatrix}^T \succeq 0
\tag{24}
\]
For a given controller parameter matrix $K$, the $\mathcal{H}_2$ norm of $T$ is less than $\gamma$, $\|T\|_{\mathcal{H}_2} < \gamma$, if and only if there exists $X > 0$ such that [14, Proposition II.1], cf. [4]

$$\text{trace } H_z^T X H_z < \gamma^2 \quad \text{and} \quad X > A_K^T X A_K + C_K^T C_K. \tag{24}$$

**Definition 4** The data $(U, Y)$ are said to be informative for common $\mathcal{H}_2$ control by output-feedback controller (10) with performance $\gamma$ if there exist a $K$ and $X > 0$ so that $\Sigma_{RQ}^{(c, x)} \subseteq \{(A_z, B_z) | \text{[24] holds true}\}$.

**Theorem 3** The data $(U, Y)$ are informative for common $\mathcal{H}_2$ control with performance $\gamma$ if there exist $L \in \mathbb{R}^{m \times n}$, symmetric $Z$, $P > 0$, $0 \geq \alpha > 0$ and $\beta > 0$ so that trace $Z < \gamma^2$, \text{[25]} holds true,

$$[P \quad F^\top] \succ 0 \quad \text{and} \quad \begin{bmatrix} Z & H_z^T \\ H_z & P \end{bmatrix} \succeq 0. \tag{26}$$

Proof: By a congruence transformation of \text{[24]} with $P := X^{-1}$ and the Schur complement (in both directions), it follows that \text{[24]} is equivalent to $P - F^T F > 0$, \text{[27]}

$$P - V_z (P - F^T F)^{-1} V_z^T > 0 \tag{27}$$

and trace $H_z^T P^{-1} H_z < \gamma^2$. Now, we can rewrite \text{[27]} as

$$\begin{bmatrix} \begin{bmatrix} I & \Sigma_{RQ}^{(c; x)} \end{bmatrix}^\top \\ \begin{bmatrix} 0 \\ \alpha \end{bmatrix} \end{bmatrix} \begin{bmatrix} P \\ 0 \end{bmatrix} \begin{bmatrix} 0 - P L (P - F^T F)^{-1} P L^\top \\ L \end{bmatrix} \begin{bmatrix} I & \Sigma_{RQ}^{(c; x)} \end{bmatrix} \succeq 0,$$

which, by \text{[28]}, holds if and only if

$$\begin{bmatrix} I \quad A_e^T \quad B_e^T \\ B_e \quad P \quad L \end{bmatrix} \begin{bmatrix} J_t P + J_2 L \\ P \quad S(\ast) \end{bmatrix} \begin{bmatrix} I \quad A_e^T \quad B_e^T \\ 0 \quad 0 \quad 0 \end{bmatrix} \succeq 0. \tag{28}$$

There exist $P > 0$, $L$ so that \text{[28]}, $P - F^T F > 0$ and trace $H_z^T P^{-1} H_z < \gamma^2$ if and only if there exist $P > 0$, $L$, $Z$ so that \text{[28]}, $P - F^T F > 0$, $Z - H_z^T P^{-1} H_z \succeq 0$ and trace $Z < \gamma^2$. Indeed, for $Z := H_z^T P^{-1} H_z$, we infer trace $Z < \gamma^2$. Sufficiency follows from $H_z^T P^{-1} H_z \preceq Z \Rightarrow$ trace $H_z^T P^{-1} H_z \preceq$ trace $Z$. Hence, the data $(U, Y)$ are informative for common $\mathcal{H}_2$ control with performance $\gamma$ if and only if there exist $P > 0$, $L$, $Z$ so that $Z - H_z^T P^{-1} H_z \preceq Z \Rightarrow$ trace $Z < \gamma^2$, $P - F^T F > 0$ and $Z \in \Sigma_{RQ}^{(c; x)}$ for all $(A_e, B_e) \in \Sigma_{RQ}^{(U; Y)}$. By assumption, trace $Z < \gamma^2$ is satisfied, $P - F^T F > 0$, $Z - H_z^T P^{-1} H_z \preceq 0$ follow by \text{[26]} and via an analogue argument as in the proof of Theorem 2, \text{[28]} holds for all $(A_e, B_e) \in \Sigma_{RQ}^{(U; Y)}$ by \text{[25]}.

**Remark 2** The conditions in Theorem 2/3 are also necessary for informativity of input-state data for $\mathcal{H}_2$ control, where $H_z = I$, $J_t = 0$, $J_2 = 0$ and $Y_-$ and $Z_-$ are replaced by $X_+$ and $X_-$, if \text{[18]} holds for some $Z$.

\begin{align*}
&P - \gamma^{-1} H_z H_z^\top - \beta I \\
&0 0 J_t P + J_2 L 0 \\
&0 0 0 P 0 \\
&P^\top J_z^T \quad + \quad L^\top J_2^T \\
&0 0 0 0 F^\top \gamma I \\
&\begin{bmatrix} I & H_z P^{-1} R_z^\top \\
&0 & -Z_2 R_z^\top \\
&0 & -U R_z^\top \\
&0 & 0 \end{bmatrix} \begin{bmatrix} I & H_z P^{-1} R_z^\top \\
&0 & -Z_2 R_z^\top \\
&0 & -U R_z^\top \\
&0 & 0 \end{bmatrix}^\top \succeq 0, \tag{21}
\end{align*}

We assume prior knowledge in the sense that $E_{-} \in \mathcal{E}_{RQ} = \{E_{-} | E_{-} R_{-}^\top R_{-} E_{-}^\top \preceq \chi_{HA} \}$, where $H_u$ is taken as $H_u = I$, independent of $N$. The cross-covariance bounds hold true for all generated data sets. We verify that there exists some $Z$ so that \text{[18]} holds true for all data sets. Hence, by Remark 2, the data are informative for common $\mathcal{H}_2$ control with performance $\gamma$ if and only if the conditions in Theorem 3 are feasible. The relative number of data sets for which this necessary and sufficient condition is feasible for some $\gamma > 0$ is visualized in Figure \text{[1a]} for each data length $N$, in blue. For $N \geq 20$, all data sets are informative for common $\mathcal{H}_2$ control. For these data sets, the smallest $\mathcal{H}_2$ norm upper bounds $\gamma^2$ are visualized in Figure \text{[1b]} where the median performance is indicated by a solid line and the shaded area is bounded by the 25th and 75th percentiles. In comparison, the $\mathcal{H}_2$ norm that can be achieved by a state feedback controller with knowledge of $(A_0, B_0)$ is \text{[18]}.
We consider three cases: with feasible $\gamma$ (red) and quadratic cross-covariance bounds (blue) and (b) for $H_2$ control versus data length $N$ for noise-norm bounds (red) and quadratic cross-covariance bounds (blue). For increasing data length, informativity becomes expected. For increasing data length, informativity becomes expected. 

Now, consider that noisy output measurements are available instead of state measurements. Consider system (1) with $A(q^{-1})$ and $B(q^{-1})$ such that $T_0(q^{-1}) = A^{-1}(q^{-1})B(q^{-1})$ with $T_0 := C_0(qI - A_0)^{-1}B_0$, where $C_0$ is the output matrix. We consider three cases: $C_0 = [1 0 1]$, $C_0 = [0 1 0]$, and $C_0 = [1 0 0]$. The noise is uniformly drawn from $[-0.35, 0.35]$. For each choice of output, we generate 50 data sets for data lengths ranging from $N = 2$ to $N = 250$. We choose an instrumental signal containing lagged input signals as before, which is therefore independent on the choice of output. The upper-bound is chosen $H_0 = 0.3$, which holds for all data sets. By Theorem 3, feasibility of the conditions for informativity for $H_2$ control for some $\gamma > 0$ is verified for each data set. The results are depicted in Figure 2. We observe that the data sets are not informative for low data lengths, which can be expected. For increasing data length, informativity becomes dependent on the choice of output. For $N = 30$, for example, 90% of the data sets yielded feasible informativity conditions for the choice of $C_0 = [1 0 0]$, compared to less than 50% of the data sets for the other two choices for $C_0$.

VI. CONCLUDING REMARKS

We have considered the problem of informativity of input-output data for control, with prior knowledge of the noise in the form of quadratic sample cross-covariance bounds. Sufficient informativity conditions for stabilization, $H_\infty$ and $H_2$ control via dynamic output feedback were derived, which are also necessary if the state is measured. We have provided a numerical case study where data-informativity can be concluded with cross-covariance bounds, while the data are concluded to be non-informative with magnitude bounds. Finally, we have illustrated how the choice of output affects the informativity of input-output data via a numerical example.

REFERENCES


\[
\begin{bmatrix}
P - \beta I & 0 & 0 & J_1P + J_2L & 0 \\
0 & 0 & 0 & P & 0 \\
0 & 0 & 0 & L & 0 \\
P^T J_1^T + L^T J_2^T & P L^T & P F^T & I
\end{bmatrix} - \alpha
\begin{bmatrix}
I & H_z P - R_z^T & 0 & -Z R_z^\top \\
0 & -Z R_z & 0 & -U R_z^\top \\
0 & 0 & Q_e & 0 \\
0 & 0 & 0 & 0
\end{bmatrix} \succeq 0
\]