Submitted to 2022 American Control Conference
and IEEE Control Systems Letters. 10 September 2021

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. We develop sufficient conditions for informativity of input-output data for stability, $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control, based on this S-procedure, which are 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.**

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] from data, and represent, for example, magnitude, energy and variance bounds on the noise.

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 Unions Horizon 2020 research and innovation programme (grant agreement No. 694504).

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 [9] as an alternative to magnitude bounds in parameter bounding identification, given its overly conservative noise characterization, cf. [10] for a comparison of instanteneous 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 in terms of cross-covariance bounds with practical relevance and informativity conditions for control based on input-output data. We provide sufficient conditions for informativity for stabilization, $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control, which are also necessary for input-state data.

II. Input-output data: cross-covariance bounds

Consider a class of linear systems described by

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

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 $B(\xi) = B_0 + B_1\xi + B_2\xi^2 + \cdots + B_n\xi^n$. By defining $\zeta(t) := \mathrm{col}(y(t-1), \ldots, y(t-n), u(t-1), \ldots, u(t-n))$, we can write (1) equivalently as

$$\begin{aligned} y(t) &= B_0u(t) + e(t) \\ &\quad + \begin{bmatrix} -A_1 & -A_2 & \cdots & -A_n & B_1 & B_2 & \cdots & B_n \end{bmatrix} \zeta(t), \end{aligned} \tag{2}$$

Hence, with $\zeta(t) \in \mathbb{R}^{n_z}$ as a state, a representation

$$\zeta(t+1) = \underbrace{\left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\bar{A}} & \multicolumn{2}{c}{\bar{B}} \\ \hline I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 \end{array}\right]}_{=:A_z} \zeta(t) + \underbrace{\begin{bmatrix} B_0 \\ \hline 0 \\ I \\ 0 \end{bmatrix}}_{=:B_z} u(t) + \underbrace{\begin{bmatrix} I \\ \hline 0 \\ 0 \\ 0 \end{bmatrix}}_{=:H_z} e(t),$$

$$y(t) = \begin{bmatrix} \bar{A} & \bar{B} \end{bmatrix} \zeta(t) + B_0u(t) + e(t), \tag{3}$$

is obtained with $\bar{A} := \mathrm{row}(-A_1, \ldots, -A_n)$ and $\bar{B} := \mathrm{row}(B_1, \ldots, B_n)$. Notice that (3) is a non-minimal representation of order n_z, where $n_z = 2n$ for the single-input-single output case $m = p = 1$. Defining the data matrices $Z_- := \begin{bmatrix} \zeta(0) & \zeta(1) & \cdots & \zeta(N-1) \end{bmatrix}$, $Y_- := \begin{bmatrix} y(0) & y(1) & \cdots & y(N-1) \end{bmatrix}$ and U_-, E_- accordingly, we obtain the data equation

$$Y_- = \begin{bmatrix} \bar{A} & \bar{B} \end{bmatrix} Z_- + B_0U_- + E_-, \tag{4}$$

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 and instrumental variable r, given by $\frac{1}{\sqrt{N}}\sum_{t=0}^{N-1} e(t)r(t)^\top = \frac{1}{\sqrt{N}}E_-R_-^\top$. We assume prior knowledge on the noise of the form

$$\frac{1}{N}E_-R_-^\top R_-E_-^\top \preceq H_u, \tag{5}$$

where H_u is an upper-bound on the squared sample cross-covariance matrix $\frac{1}{\sqrt{N}}E_-R_-^\top$. In a generalized form, we write

$$\begin{bmatrix} I \\ R_-E_-^\top \end{bmatrix}^\top \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} I \\ R_-E_-^\top \end{bmatrix} \succeq 0, \tag{6}$$

with $P_{22} \prec 0$. For $P_{11} = NH_u$, $P_{12} = P_{21}^\top = 0$ and $P_{22} = -I$, the bound (5) is recovered. In the state-space representation (3), the state-space matrices A_z and B_z contain unknown parameters $\bar{A}$, $\bar{B}$ and B_0. We write A_z and B_z as

$$A_z = \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array}\right] + \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \hline I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 \end{array}\right] =: \Lambda_e + J_1 \tag{7}$$

$$\text{and } B_z = \left[\begin{array}{c} B_0 \\ \hline 0 \\ 0 \\ 0 \end{array}\right] + \left[\begin{array}{c} 0 \\ \hline 0 \\ I \\ 0 \end{array}\right] =: B_e + J_2, \tag{8}$$

so that Λ_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 (Λ_e, B_e) that are compatible with the data is

$$\Sigma_{(U,Y)}^{RQ} := \{(\Lambda_e, B_e) \mid \exists E_- \text{ satisfying (6) so that (2) holds}\}.$$

Lemma II.1 *Under the assumption that the noise E_- satisfies (6), the set of feasible pairs (Λ_e, B_e) is*

$$\Sigma_{(U,Y)}^{RQ} = \{(\Lambda_e, B_e) \mid (\Lambda_e, B_e) \text{ satisfies (9)}\},$$

where

$$\begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \begin{bmatrix} I & H_zY_-R_-^\top \\ 0 & -Z_-R_-^\top \\ 0 & -U_-R_-^\top \end{bmatrix} P_e \begin{bmatrix} I & H_zY_-R_-^\top \\ 0 & -Z_-R_-^\top \\ 0 & -U_-R_-^\top \end{bmatrix}^\top \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix} \succeq 0, \tag{9}$$

with $P_e := \begin{bmatrix} H_zP_{11}H_z^\top & H_zP_{12} \\ P_{12}^\top H_z^\top & P_{22} \end{bmatrix}$.

We denote all (A_z, B_z) that are compatible with the data by

$$\bar{\Sigma}_{(U,Y)}^{RQ} := \{(\Lambda_e + J_1, B_e + J_2) \mid (\Lambda_e, B_e) \in \Sigma_{(U,Y)}^{RQ}\}.$$

We have provided a parametrization of $\Sigma_{(U,Y)}^{RQ}$ based on the data equation (4). One can equivalently parametrize $\bar{\Sigma}_{(U,Y)}^{RQ}$ on the basis of the state data equation $Z_+ = A_zZ_- + B_zU_- + H_zE_-$. This leads to an equal set $\bar{\Sigma}_{(U,Y)}^{RQ}$, but the 'repeated' data in the parametrization contained in $Z_+ := \begin{bmatrix} \zeta(1) & \cdots & \zeta(N) \end{bmatrix}$, would render the evaluation more sensitive in terms of numerical precision.

B. Output-feedback control

Consider a (dynamic) output feedback controller described by the difference equation of the form [5]

$$C(q^{-1})y_c(t) = D(q^{-1})u_c(t), \tag{10}$$

with $C(\xi) \in \mathbb{R}^{m\times m}[\xi]$ and $D(\xi) \in \mathbb{R}^{m\times p}$ polynomial matrices given by $C(\xi) = I + C_1\xi + C_2\xi^2 + \cdots + C_n\xi^n$ and $D(\xi) = D_1\xi + D_2\xi^2 + \cdots + D_n\xi^n$ and q^{-1} the delay operator so that $q^{-1}x(t) = x(t-1)$. The interconnection of the controller (10) with the system (1) is described by the interconnection equations

$$u_c = y \quad \text{and} \quad y_c = u. \tag{11}$$

Similar to the definition of the state ξ for the system (1), we define the state ζ_c for the controller (10) as $\zeta_c := \mathrm{col}(y_c(t-1), \ldots, y(t-n), u_c(t-1), \ldots, u(t-n))$. A state-space representation for the controller is thus given by

$$\zeta_c(t+1) = \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\bar{C}} & \multicolumn{2}{c}{\bar{D}} \\ \hline I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 \end{array}\right]\zeta_c(t) + \left[\begin{array}{c} 0 \\ \hline 0 \\ I \\ 0 \end{array}\right]u_c(t), \tag{12}$$
$$y_c(t) = \begin{bmatrix} \bar{C} & \bar{D} \end{bmatrix}\zeta_c(t),$$

with $\bar{C} := \mathrm{row}(-C_1, \ldots, C_n)$ and $\bar{D} := \mathrm{row}(D_1, \ldots, D_n)$. By (11), it follows that $\zeta_c = \left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]\zeta$, which implies that $u(t) = y_c(t) = \begin{bmatrix} \bar{C} & \bar{D} \end{bmatrix}\zeta_c(t) = \begin{bmatrix} \bar{D} & \bar{C} \end{bmatrix}\zeta(t)$. Hence, the closed-loop system described by (3), (12) and (11) has a state-space representation

$$\zeta(t+1) = \underbrace{\left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\bar{A}+B_0\bar{D}} & \multicolumn{2}{c}{\bar{B}+B_0\bar{C}} \\ \hline I & 0 & 0 & 0 \\ \hline \multicolumn{2}{c|}{\bar{D}} & \multicolumn{2}{c}{\bar{C}} \\ \hline 0 & 0 & I & 0 \end{array}\right]}_{=:A_{\mathrm{cl}}}\zeta(t) + \begin{bmatrix} I \\ 0 \\ 0 \\ 0 \end{bmatrix}e(t). \tag{13}$$

With $K := \begin{bmatrix} \bar{D} & \bar{C} \end{bmatrix}$, the closed-loop system matrix A_{cl} satisfies $A_{\mathrm{cl}} = A_z + B_zK$. For some $(A_z, B_z) \in \bar{\Sigma}_{(U,Y)}^{RQ}$, we say that the controller (10) stabilizes (1) if the closed-loop system (13) is stable, i.e., if all eigenvalues of $A_z + B_zK$ are in the open unit circle, since this implies stability of the closed-loop system (1), (10) and (11). The notion of stabilization with respect to the state-space representation (3) was introduced in [5] for data-driven stabilization. We note that in the single-input-single-output case, (A_z, B_z) is controllable if and only if $A(\xi)$ and $B(\xi)$ are coprime [5].

III. Informativity for Stabilization

A. Informativity of input-output data

Definition III.1 The data (U, Y) are said to be informative for quadratic stabilization by by output-feedback controller (10) if there exist a K and $X \succ 0$ so that

$$\bar{\Sigma}_{(U,Y)}^{RQ} \subseteq \{(A, B) \mid (A + BK)X(A + BK)^\top - X \prec 0\}.$$

By (7), we find that the existence of K and $X \succ 0$ so that

$$(A_z + B_z K)^\top X (A_z + B_z K) - X \prec 0,$$

is equivalent with the existence of K and $X \succ 0$ such that (14) holds true. Now, for the data (U, Y) to be informative for quadratic stabilization, we require the existence of K and $X \succ 0$ so that (14) holds for all $(\Lambda_e, B_e) \in \Sigma^{RQ}_{(U,Y)}$. This is precisely a problem that can be solved by the S-procedure; more specifically, by the matrix-valued S-lemma [4].

Theorem III.1 *The data (U, Y) are informative for quadratic stabilization by output feedback controller (10) if there exist $L \in \mathbb{R}^{m \times n_z}$, $X \succ 0$, $\alpha \geq 0$ and $\beta > 0$ so that (15) holds true. Moreover, for L and X such that (15) is satisfied, $A_{cl} = A_z + B_z K$ is stable for all $(A_z, B_z) \in \bar{\Sigma}^{RQ}_{(U,Y)}$ with $K := LX^{-1}$.*

Proof: Let there exist L, $Y \succ 0$, $\alpha \geq 0$ and $\beta > 0$ so that (15) holds true and consider the matrix Π defined in (14). By the Schur complement, (15) is equivalent with

$$\Pi - \alpha \Lambda \succeq \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{where } K := LX^{-1} \quad \text{and}$$

$$\Lambda := \begin{bmatrix} I & H_z Y_- R_-^\top \\ 0 & -Z_- R_-^\top \\ 0 & -U_- R_-^\top \end{bmatrix} P_e \begin{bmatrix} I & H_z Y_- R_-^\top \\ 0 & -Z_- R_-^\top \\ 0 & -U_- R_-^\top \end{bmatrix}^\top .$$

Hence, (14) holds true for all $(\Lambda_e, B_e) \in \Sigma^{RQ}_{(U,Y)}$, cf. [4, Theorem 13]. This concludes the proof. ■

We remark that if there is a Z so that $\bar{Z}^\top \Lambda \bar{Z} \succ 0$ with $\bar{Z} := \mathrm{col}(I, Z)$, called the generalized Slater condition [4], then (15) is also a necessary condition for informativity of input-output data for quadratic stabilization. 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 $n \geq 1$, since the noise affects a subspace of the extended state space. The combination of noisy and noiseless states in ζ suggests that necessity could potentially proven in general by a 'fusion' of the matrix S-lemma and matrix Finsler's lemma [11].

B. Informativity of input-state data

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

$$x(t+1) = Ax(t) + Bu(t) + e(t), \tag{16}$$

with the corresponding data equation

$$X_+ = AX_- + BU_- + E_-. \tag{17}$$

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

$$\Sigma^{RQ}_{(U_-,X)} := \{(A, B) \mid \exists E_- \text{ satisfying (6) so that (17) holds}\}$$

By (17), the set of feasible systems is $\Sigma^{RQ}_{(U_-,X)} = \{(A, B) \mid (A, B) \text{ satisfies (18)}\}$, where

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} I & X_+ R_-^\top \\ 0 & -X_- R_-^\top \\ 0 & -U_- R_-^\top \end{bmatrix} P \begin{bmatrix} I & X_+ R_-^\top \\ 0 & -X_- R_-^\top \\ 0 & -U_- R_-^\top \end{bmatrix}^\top}_{=:\Lambda_X} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \succeq 0. \tag{18}$$

Remark III.1 Consider a specific selection of $M = N$ instrumental variables defined by $r_i(t) := \delta(t - i)$, $i = 1, \ldots, N$, where the index i runs from 1 to 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 III.2 The data (U_-, X) are said to be informative for quadratic stabilization by state feedback if there exist a feedback gain K and $P \succ 0$ so that

$$\Sigma^{RQ}_{(U_-,X)} \subseteq \{(A, B) \mid (A + BK) P (A + BK)^\top - P \prec 0\}.$$

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

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top \Lambda_X \begin{bmatrix} I \\ Z \end{bmatrix} \succ 0. \tag{19}$$

Proposition III.1 *Suppose that there exists a Z so that (19) holds true. Then the data (U_-, X) are informative for*

$$\begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} X - (J_1 + J_2 K) X (J_1 + J_2 K)^\top & -(J_1 + J_2 K) X & -(J_1 + J_2 K) X K^\top \\ -X (J_1 + J_2 K)^\top & -X & -X K^\top \\ -K X (J_1 + J_2 K)^\top & -K X & -K X K^\top \end{bmatrix}}_{=:\Pi} \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix} \succ 0 \tag{14}$$

$$\begin{bmatrix} X - \beta I & -J_1 X - J_2 L & 0 & J_1 X + J_2 L \\ -X J_1^\top - L^\top J_2^\top & -X & -L^\top & 0 \\ 0 & -L & 0 & L \\ X J_1^\top + L^\top J_2^\top & 0 & L^\top & X \end{bmatrix} - \alpha \begin{bmatrix} I & H_z Y_- R_-^\top \\ 0 & -Z_- R_-^\top \\ 0 & -U_- R_-^\top \\ 0 & 0 \end{bmatrix} P_e \begin{bmatrix} I & H_z Y_- R_-^\top \\ 0 & -Z_- R_-^\top \\ 0 & -U_- R_-^\top \\ 0 & 0 \end{bmatrix}^\top \succeq 0 \tag{15}$$

quadratic stabilization if and only if there exist $L \in \mathbb{R}^{m\times n}$, $y \succ 0$, $\alpha \geq 0$ *and* $\beta > 0$ *so that*

$$\begin{bmatrix} X-\beta I & 0 & 0 & 0 \\ 0 & -X & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & X \end{bmatrix} - \alpha \begin{bmatrix} \Lambda_X & 0 \\ 0 & 0 \end{bmatrix} \succeq 0. \quad (20)$$

Moreover, K *is such that* $A+BK$ *is stable for all* $(A,B) \in \Sigma^{RQ}_{(U_-,X)}$ *if* $K := LY^{-1}$ *with* L *and* $Y \succ 0$ *satisfying* (20).

Proof: ($\Leftarrow$) This is proven by the same argument as in the proof of Theorem III.1. ($\Rightarrow$) Let the data be informative for quadratic stabilization, i.e., there exist K and $X \succ 0$ so that, with Π defined in (14) with $J_1 = 0$, $J_2 = 0$:

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \Pi\,(\star) \succ 0 \text{ for all } (A,B) \text{ with } \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \Lambda_X\,(\star) \succeq 0,$$

with

$$\Lambda_X = \left[\begin{array}{c|c} \Lambda^X_{11} & \Lambda^X_{12} \\ \hline \Lambda^X_{21} & \Lambda^X_{22} \end{array}\right] := \left[\begin{array}{cc} I & X_+R_-^\top \\ \hline 0 & -X_-R_-^\top \\ 0 & -U_-R_-^\top \end{array}\right] P\,(\star)^\top,$$

We will now show that $\ker \Lambda^X_{22} \subseteq \Lambda^X_{12}$, such that necessity follows by the matrix S-lemma [4]. First, notice that $\ker \Lambda^X_{22} = \ker R_- \begin{bmatrix} X_-^\top & U_-^\top \end{bmatrix}$. Now, take any $x \in \ker \Lambda^X_{22}$. Then $R_- \begin{bmatrix} X_-^\top & U_-^\top \end{bmatrix} x = 0$. Clearly, we have that $X_+R_-^\top R_- \begin{bmatrix} X_-^\top & U_-^\top \end{bmatrix} x = 0$, which implies that $x \in \ker \Lambda^X_{12}$. Since $x \in \ker \Lambda^X_{22}$ was chosen arbitrary, this shows that $\ker \Lambda^X_{22} \subseteq \Lambda^X_{12}$. By $\ker \Lambda^X_{22} \subseteq \Lambda^X_{12}$ and (19), 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 with (20) for $L := KX$ 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 $z(t) = C\zeta(t) + Du(t)$. For any pair (A_z, B_z), the controller (10) yields the closed loop system

$$\zeta(t+1) = (A_z + B_zK)\zeta(t) + H_ze(t),$$
$$z(t) = (C+DK)\zeta(t).$$

Hence, the transfer matrix from e to z is given by

$$T(z) := (C+DK)(zI - A_z - B_zK)^{-1}H_z,$$

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 γ, $\|T\|_{\mathcal{H}_\infty} < \gamma$, if and only if there exists $X \succ 0$ such that [12]

$$\begin{bmatrix} X & 0 & A_K^\top X & C_K^\top \\ 0 & \gamma I & H_z^\top X & 0 \\ XA_K & XH_z & X & 0 \\ C_K & 0 & 0 & \gamma I \end{bmatrix} \succ 0, \quad (21)$$

where $A_K := A_z + B_zK$ and $C_K := C + DK$.

Definition IV.1 The data (U,Y) are said to be informative for common $\mathcal{H}_\infty$ control by output-feedback controller (10) with performance γ if there exist a K and $X \succ 0$ so that

$$\bar{\Sigma}^{RQ}_{(U,Y)} \subseteq \{(A_z, B_z) \,|\, (21) \text{ holds true}\}.$$

Theorem IV.1 *The data* (U,Y) *are informative for common* $\mathcal{H}_\infty$ *control with performance* γ *if there exist* $L \in \mathbb{R}^{m\times n_z}$, $Y \succ 0$, $\alpha \geq 0$ *and* $\beta > 0$ *so that* (22) *holds true.*

Proof: By a congruence transformation of (21) with $\text{diag}(Y, I, Y, I)$ with $Y := X^{-1}$ and the application of the Schur complement (twice), the existence of K and $X \succ 0$ so that (21) holds, is equivalent with the existence of Y and L so that $Y \succ 0$ and

$$Y - V_z \underbrace{(Y - \gamma^{-1}F^\top F)^{-1}}_{=:S} V_z^\top - \gamma^{-1}H_zH_z^\top \succ 0 \quad (23)$$

and $Y - \gamma^{-1}F^\top F \succ 0$, where $V_z := A_zY + B_zL$ and $F := CY + DL$. We can now rewrite (23) as

$$\begin{bmatrix} I \\ A_z^\top \\ B_z^\top \end{bmatrix}^\top \begin{bmatrix} Y - \gamma^{-1}H_zH_z^\top & 0 \\ 0 & -\begin{bmatrix} Y \\ L \end{bmatrix} S \begin{bmatrix} Y \\ L \end{bmatrix}^\top \end{bmatrix} \begin{bmatrix} I \\ A_z^\top \\ B_z^\top \end{bmatrix} \succ 0,$$

which is equivalent with

$$\begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \Pi_{\mathcal{H}_\infty} \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix} := \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \begin{bmatrix} Y - \gamma^{-1}H_zH_z^\top & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}$$
$$- \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \begin{bmatrix} J_1Y + J_2L \\ Y \\ L \end{bmatrix} S\,(\star)^\top \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix} \succ 0. \quad (24)$$

Hence, the data (U,Y) are informative for common $\mathcal{H}_\infty$ control with performance γ if and only if there exist $Y \succ 0$ and L such that $Y - \gamma^{-1}F^\top F \succ 0$ and (24) holds for all $(\Lambda_e, B_e) \in \Sigma^{RQ}_{(U,Y)}$. By assumption, there exist $Y \succ 0$, L, $\alpha \geq 0$ and $\beta > 0$ such that (22) holds true. By the Schur complement, (22) is equivalent with $\Pi_{\mathcal{H}_\infty} - \alpha\Lambda \succeq \left[\begin{smallmatrix} \beta I & 0 \\ 0 & 0 \end{smallmatrix}\right]$, which implies that (24) holds for all $(\Lambda_e, B_e) \in \Sigma^{RQ}_{(U,Y)}$. ∎

The conditions (22) are linear with respect to Y, L, α and β. By a straightforward additional application of the Schur complement, (22) can also be made linear with respect to γ.

For a given controller parameter matrix K, the $\mathcal{H}_2$ norm of T is less than γ, $\|T\|_{\mathcal{H}_2} < \gamma$, if and only if there exists $X \succ 0$ such that

$$\text{trace}\, X < \gamma^2 \quad \text{and} \quad X \succ A_K^\top XA_K + C_K^\top C_K. \quad (25)$$

Definition IV.2 The data (U,Y) are said to be informative for common $\mathcal{H}_2$ control by output-feedback controller (10) with performance γ if there exist a K and $X \succ 0$ so that $\bar{\Sigma}^{RQ}_{(U,Y)} \subseteq \{(A_z, B_z) \,|\, (25) \text{ holds true}\}$.

Theorem IV.2 *The data* (U, Y) *are informative for common* $\mathcal{H}_2$ *control with performance* γ *if there exist* $L \in \mathbb{R}^{m \times n_z}$, *symmetric* $Z, Y \succ 0$, $\alpha \geq 0$ *and* $\beta > 0$ *so that* $\operatorname{trace} Z < \gamma^2$, (26) *holds true,*

$$\begin{bmatrix} Y & F^\top \\ F & I \end{bmatrix} \succ 0 \quad \textit{and} \quad \begin{bmatrix} Z & H_z^\top \\ H_z & Y \end{bmatrix} \succeq 0. \tag{27}$$

Proof: By a congruence transformation of (25) with $Y := X^{-1}$ and the Schur complement (in both directions), it follows that (25) is equivalent with $Y - F^\top F \succ 0$ and

$$Y - V_z (Y - F^\top F)^{-1} V_z^\top \succ 0. \tag{28}$$

Now, we can rewrite inequality (28) as

$$\begin{bmatrix} I \\ A_z^\top \\ B_z^\top \end{bmatrix}^\top \begin{bmatrix} Y & 0 \\ 0 & -\begin{bmatrix} Y \\ L \end{bmatrix} (Y - F^\top F)^{-1} \begin{bmatrix} Y \\ L \end{bmatrix}^\top \end{bmatrix} \begin{bmatrix} I \\ A_z^\top \\ B_z^\top \end{bmatrix} \succ 0,$$

which, by (7), holds if and and only if

$$\begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \begin{bmatrix} J_1 Y + J_2 L \\ Y \\ L \end{bmatrix} S(\star)^\top \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix} \prec \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}. \tag{29}$$

Hence, the data (U, Y) are informative for common $\mathcal{H}_2$ control with performance γ if and only if there exist $Y \succ 0$, L and Z such that $\operatorname{trace} Z < \gamma^2$, $Y - F^\top F \succ 0$ and (29) hold for all $(\Lambda_e, B_e) \in \Sigma_{(U,Y)}^{RQ}$. By assumption, $\operatorname{trace} Z < \gamma^2$ is satisfied, $Y - F^\top F \succ 0$ follows by (27) and via an analogue argument as in the proof of Theorem IV.1, (29) holds for all $(\Lambda_e, B_e) \in \Sigma_{(U,Y)}^{RQ}$ by (26). ■

Remark IV.1 The conditions in Theorem IV.1/IV.2 are also necessary for informativity of input-state data for $\mathcal{H}_\infty$/$\mathcal{H}_2$ control, where $H_z = I$, $J_1 = 0$, $J_2 = 0$ and Y_- and Z_- are replaced by X_+ and X_-, if (19) holds for some Z.

V. Numerical Example

Consider the system (16) with true system matrices

$$A_0 = \begin{bmatrix} -0.2414 & -0.8649 & 0.6277 \\ 0.3192 & -0.0301 & 1.0933 \\ 0.3129 & -0.1649 & 1.1093 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}.$$

and consider a performance output $z(t) = [0 \ 0 \ 1]x(t)$. The objective is to determine if input state data collected from the system are informative for common $\mathcal{H}_2$ control. We consider a noise signal $e(t)$ with a uniform distribution, taking values from the closed ball $\{e \in \mathbb{R}^3 \mid \|e\|_2^2 \leq 0.35\}$. First, we consider this norm bound to be known and use it as prior knowledge. This is represented by the noise model $E_- \in \{E_- \mid E_- E_-^\top \preceq 0.35 N I\}$ as described in [4, Section VI.A], which can be represented by the noise model (5) with $R_- = I$, cf. [4, Equation (5)]. We consider the informativity analysis for various data lengths N ranging from $N = 2$ to $N = 250$. For each data length N, we generate 50 data sets. Given the prior knowledge on the noise bounds, we can verify informativity for common $\mathcal{H}_2$ control via Theorem 17 in [4]. We verify the generalized Slater condition [4, Equation (16)], which holds true for all data sets. Hence, the data are informative for common $\mathcal{H}_2$ control with performance γ if and only if the condition [4, Equation ($\mathcal{H}_2$)] is feasible. The relative number of data sets for which this necessary and sufficient condition is feasible for some $\gamma > 0$ is visualized in Figure 1a for each data length N, in red. Notice that, naturally, if the condition is not feasible for any $\gamma > 0$, the data are actually not even informative for feedback stabilization, while the true system *is* stable, i.e., A_0 is stable.

Now, we consider the quadratic cross-covariance bound (5) for the noise. We choose an instrumental variable that contains lagged versions of the input:

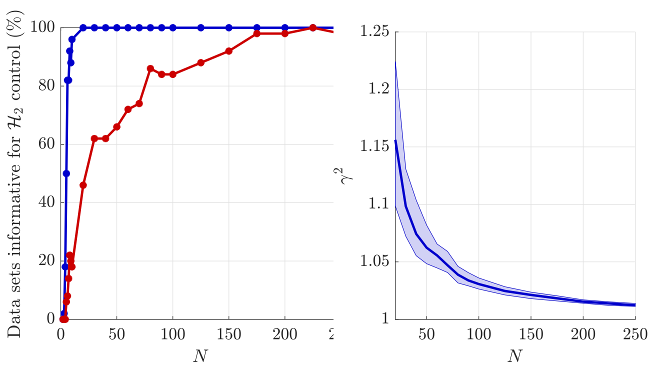
$$r(t) := \operatorname{col}(u(t), u(t-1), u(t-2), \ldots, u(t-8), u(t-9)).$$

We assume prior knowledge in the sense that $E_- \in \mathcal{E}_{RQ} = \{E_- \mid E_- R_-^\top R_- E_-^\top \preceq N H_u\}$, 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 (19) holds true for all data sets. Hence, by Remark IV.1, the data are informative for common $\mathcal{H}_2$ control with performance γ if and only if the conditions in Theorem IV.2 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 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 γ^2 are visualized in Figure 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 the true system matrices (A_0, B_0) is equal to 1.000, which therefore is a benchmark performance that cannot be outperformed by any data-based controller.

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 $T_0(q^{-1}) = A^{-1}(q^{-1})B(q^{-1})$

$$\begin{bmatrix} Y - \gamma^{-1} H_z H_z^\top - \beta I & 0 & 0 & J_1 Y + J_2 L & 0 \\ 0 & 0 & 0 & Y & 0 \\ 0 & 0 & 0 & L & 0 \\ Y^\top J_1^\top + L^\top J_2^\top & Y & L^\top & Y & F^\top \\ 0 & 0 & 0 & F & \gamma I \end{bmatrix} - \alpha \begin{bmatrix} I & H_z Y_- R_-^\top \\ 0 & -Z_- R_-^\top \\ 0 & -U_- R_-^\top \\ 0 & 0 \\ 0 & 0 \end{bmatrix} P_e \begin{bmatrix} I & H_z Y_- R_-^\top \\ 0 & -Z_- R_-^\top \\ 0 & -U_- R_-^\top \\ 0 & 0 \\ 0 & 0 \end{bmatrix}^\top \succeq 0, \quad \begin{bmatrix} Y & F^\top \\ F & \gamma I \end{bmatrix} \succ 0 \tag{22}$$

$$\begin{bmatrix} Y-\beta I & 0 & 0 & J_1Y+J_2L & 0 \\ 0 & 0 & 0 & Y & 0 \\ 0 & 0 & 0 & L & 0 \\ Y^\top J_1^\top + L^\top J_2^\top & Y & L^\top & Y & F^\top \\ 0 & 0 & 0 & F & I \end{bmatrix} - \alpha \begin{bmatrix} I & H_zY_-R_-^\top \\ 0 & -Z_-R_-^\top \\ 0 & -U_-R_-^\top \\ 0 & 0 \\ 0 & 0 \end{bmatrix} P_e \begin{bmatrix} I & H_zY_-R_-^\top \\ 0 & -Z_-R_-^\top \\ 0 & -U_-R_-^\top \\ 0 & 0 \\ 0 & 0 \end{bmatrix}^\top \succeq 0 \tag{26}$$



(a) Informativity for $\mathcal{H}_2$ control (b) $\mathcal{H}_2$ performance from data

Fig. 1: (a) Number of input-state data sets that are informative for $\mathcal{H}_2$ control versus data length N for noise-norm bounds (red) and quadratic cross-covariance bounds (blue) and (b) feasible γ^2 obtained versus data length N with quadratic cross-covariance bounds.

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 drawn uniformly from the interval $[-0.35, 0.35]$. For each choice of output, we generate 50 data sets for various 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 $H_u = 0.3$, which holds for all data sets. By Theorem IV.2, feasibility of the conditions for informativity for $\mathcal{H}_2$ 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, $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control via dynamic output feedback were derived, which are also necessary if the state is measured. The considered cross-covariance bounds can capture quadratic bounds in the literature and 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

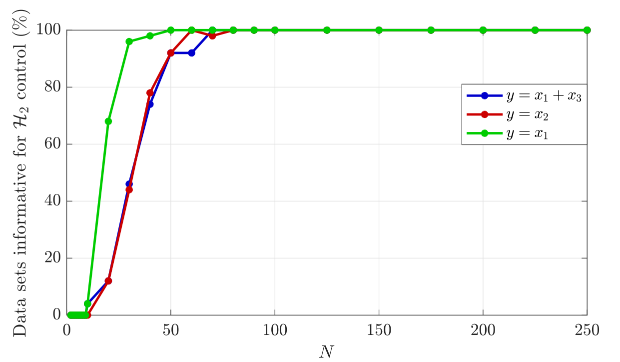


Fig. 2: Effect of the choice of output on informativity of input-output data for $\mathcal{H}_2$ control.

with magnitude bounds. Finally, we have illustrated how the choice of output affects the informativity of input-output data for control via a numerical example.

References

[1] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1999.

[2] P. M. J. Van den Hof and R. J. P. Schrama, "Identification and control: Closed-loop issues," *Automatica*, vol. 31, no. 12, pp. 1751 – 1770, 1995.

[3] H. J. van Waarde, J. Eising, H. L. Trentelman, and M. K. Camlibel, "Data informativity: A new perspective on data-driven analysis and control," *IEEE Trans. Autom. Control*, vol. 65, no. 11, pp. 4753–4768, 2020.

[4] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, "From noisy data to feedback controllers: non-conservative design via a matrix S-lemma," *IEEE Trans. Autom. Control*, 2020, early access, doi: 10.1109/TAC.2020.3047577.

[5] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, 2020.

[6] J. Berberich, C. W. Scherer, and F. Allgöwer, "Combining prior knowledge and data for robust controller design," *ArXiv e-prints*, vol. arXiv:2009.05253, 2021.

[7] T. R. V. Steentjes, M. Lazar, and P. M. J. Van den Hof, "Guaranteed $\mathcal{H}_\infty$ performance analysis and controller synthesis for interconnected linear systems from noisy input-state data," in *60th IEEE Conf. Decision and Control (CDC) (accepted)*, Austin, Texas, USA, 2021.

[8] A. Koch, J. Berberich, and F. Allgöwer, "Verifying dissipativity properties from noise-corrupted input-state data," in *59th IEEE Conf. Decision and Control (CDC)*, 2020, pp. 616–621.

[9] R. G. Hakvoort and P. M. J. Van den Hof, "Consistent parameter bounding identification for linearly parametrized model sets," *Automatica*, vol. 31, no. 7, pp. 957–969, 1995.

[10] A. Bisoffi, C. De Persis, and P. Tesi, "Trade-offs in learning controllers from noisy data," *Systems & Control Letters*, vol. 154, p. 104985, 2021.

[11] H. J. van Waarde and M. K. Camlibel, "A matrix Finsler's lemma with applications to data-driven control," *ArXiv e-prints*, vol. arXiv:2103.13461, 2021.

[12] C. Scherer and S. Weiland, "Linear matrix inequalities in control," October 2017, DISC lecture notes.