Data-driven estimation of Gramian based interaction measures for control structure selection

Andreé Carvalho Bittencourt
Division of Automatic Control
E-mail: andre.carvalho.bittencourt@liu.se

8th December 2016

Report no.: LiTH-ISY-R-3094

Address:
Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

WWW: http://www.control.isy.liu.se
Abstract
Interaction measures quantify the input-output relations in MIMO processes and can support the selection of control structures (CSS). Interaction measures are typically computed based on an existing process models. The study of input-output interactions based on data can complement missing information on a model, e.g., revealing unknown relations in a complex system or adjusting for a time dependent behavior. This paper presents a unified approach for data-driven estimation of Gramian based interaction measures from input-output data. Given open or closed-loop data, a high-order Vector ARX (VARX) model is identified and its parameters are used to calculate predictor Markov parameters, together with a covariance estimate. Three interaction measures based on the Hankel, Hilbert-Schmidt-Hankel and $H_2$ norms are calculated from the estimated predictor Markov parameters and uncertainty estimates are provided for the last two, allowing for robust CSS. A solution which is recursive in the data points is presented, making it practical for applications to large datasets. The approach is verified through simulations and several possible extensions are discussed. As the method is suitable for open and close-loop data and for large datasets, it opens up for data-driven control structure selection based on operational data from entire plants.

Keywords: interaction measures, control structure selection, data-driven, Gramians, VARX models
Data-driven estimation of Gramian based interaction measures for control structure selection

André Carvalho Bittencourt*

Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden

Abstract

Interaction measures quantify the input-output relations in MIMO processes and can support the selection of control structures (CSS). Interaction measures are typically computed based on an existing process models. The study of input-output interactions based on data can complement missing information on a model, e.g., revealing unknown relations in a complex system or adjusting for a time dependent behavior. This paper presents a unified approach for data-driven estimation of Gramian based interaction measures from input-output data. Given open or closed-loop data, a high-order Vector ARX (VARX) model is identified and its parameters are used to calculate predictor Markov parameters, together with a covariance estimate. Three interaction measures based on the Hankel, Hilbert-Schmidt-Hankel and $H_2$ norms are calculated from the estimated predictor Markov parameters and uncertainty estimates are provided for the last two, allowing for robust CSS. A solution which is recursive in the data points is presented, making it practical for applications to large datasets. The approach is verified through simulations and several possible extensions are discussed. As the method is suitable for open and closed-loop data and for large datasets, it opens up for data-driven control structure selection based on operational data from entire plants.

Keywords: interaction measures, control structure selection, data-driven, Gramians, VARX models

*Corresponding author
Email address: andre.carvalho.bittencourt@liu.se (André Carvalho Bittencourt)

Preprint submitted to Elsevier December 8, 2016
1. INTRODUCTION

The design process of control systems comprehends the tasks of defining the control objectives, modeling of the plant, control structure selection, controller design, verification and implementation, see e.g. [1]. Control structure selection (CSS) is a key step in the design of control systems, delimiting its achievable performance and complexity. In CSS, controlled, manipulated and measured variables are chosen as well as the control configuration, depicting the information flow used by the controller, and the type of controller used.

Aiming at supporting the decisions involved in CSS, it is common to make use of quantifiable measures of interaction between input-output variables. A large number of interaction measures can be found in the literature, from the classical relative gain array, [2], to more modern approaches based on Gramians, such as the participation matrix, [3], and many others, see [4] for a review. In general, CSS is performed before deployment, based on available models of the plant. Due to the complex nature of some MIMO processes, such as reactors, incinerators and distillation columns, a priori knowledge of the plant dynamics can be difficult. The robustness of CSS can be improved by considering modeling uncertainties as presented in [5, 6]. Despite the efforts in control design, a given control structure may not perform as desired in practice and a reconfiguration of the control structure might be needed, as illustrated in [7] for a bark boiler process. As pointed out in [8], CSS might be performed after deployment also for economical reasons, when a reduced number of actuators and/or sensors is desired, or to study how new sensors and actuators could be used to improve performance. Reconfiguration of the control structure might also be needed due to changes in the plant dynamics or in its structure as presented in [8].

In order to complement missing information about the plant dynamics, it is possible to estimate the interaction measures based on data collected from an operational plant. In [9, 10, 5, 11, 12], a number of approaches are presented for specific interaction measures which rely on dedicated identification experiments,
performed in open-loop, and verified in simulations. Given the great potential for the use of data collected from routine operation data (see e.g. [13]), it is important to develop more general methods, not dependent on dedicated excitation or an open-loop operation.

In this paper, we propose a method to estimate Gramian-based interaction measures from open or closed loop data for MIMO systems. We devise a method that can handle large datasets and provides statistical measures of uncertainty for the estimates, allowing for robust CSS. Section 2 provides an overview of Gramian-based interactions and presents a review of existing literature contributions. Section 3 presents a method to estimate Gramian-based interactions from the Markov parameters of a flexible high-order vector ARX model identified from open or closed-loop input-output data. Section 5 illustrates the method through simulations. Possible extensions are presented in Section 4 and conclusions in Section 6.

2. Gramian based interaction measures

In this section we consider a stable deterministic system

\[ x(k + 1) = Ax(k) + \begin{bmatrix} b_1 & \cdots & b_{n_u} \end{bmatrix} u(k) \]  

\[ y(k) = \begin{bmatrix} c_1^T \\ \vdots \\ c_{n_y}^T \end{bmatrix} x(k) + Du(k) \]

where \( u(k) \in \mathbb{R}^{n_u}, y(k) \in \mathbb{R}^{n_y} \), are the input (manipulated variables) and output (measured variables) vectors, \( x(k) \in \mathbb{R}^{n_x} \) are internal states. This system has transfer function \( G(q) = C(qI - A)^{-1}B + D \) composed of subsystems \( G_{ij}(q) = c_i^T(qI - A)^{-1}b_j + D_{ij} \). Gramian based interaction measures are defined based
on an index array of the form

\[ \Sigma_{ij}^p = \sum_{k,l} \bar{\Sigma}_{ij}^p, \quad \bar{\Sigma}_{ij}^p \triangleq \{G_{ij}\}_p, \] (2)

where \( \bar{\Sigma}^p \) represents the unnormalized array based on the \( \{\cdot\}_p \) measure. The matrix \( \Sigma^p \) has dimension \( n_y \times n_u \) and element \( \Sigma_{ij}^p \) relates to how much input \( j \) affects output \( i \) according to the measure \( \{\cdot\}_p \). The index array is normalized, with all its elements summing to 1 and can be used as a guideline for CSS. Essentially, one would like to choose a small subset \( \mathcal{S} \) of pairs \((i,j)\) such that their total contribution is large enough, i.e.

\[ \sum_{ij \in \mathcal{S}} \Sigma_{ij}^p > h, \]

meaning that the selection of input-output pairs capture most of the plant dynamics. In [3], it is recommended to choose \( h > 0.7 \) to cover most of the system dynamics. Simpler control structures, such as decentralized, can be attempted first and the complexity can be increased until \( h \) is reached. Pairs can be added to or removed from \( \mathcal{S} \) according to the associated value of \( \Sigma_{ij}^p \). As a rule of thumb, elements smaller than the contribution in the homogeneous case, \( 1/(n_y n_u) \), can be removed and those significantly larger can be added, see [3].

Because of limited knowledge of the plant, it is also common to propagate and represent uncertainty also in the elements of the index array \( \Sigma_{ij}^p \). The uncertainties are often represented by confidence intervals and the resulting analysis for CSS is denoted as robust, see e.g. [5, 10].

2.1. Gramians, Markov parameters and Hankel matrix

The measure used \( \{\cdot\}_p \) directly affects the index array and thus CSS. Before we specify commonly used measures, we introduce some definitions as they are relevant for the discussion. For system [1], the controllability and observability
Gramians, are defined respectively as
\[ P = \sum_{k=0}^{\infty} A^k B B^T (A^k)^T, \quad Q = \sum_{k=0}^{\infty} (A^k)^T C^T C A^k, \]  
and satisfy the discrete Lyapunov equations
\[ APA^T - P + B B^T = 0, \quad A^T QA - Q + C^T C = 0. \]

We also define \( P_j \) and \( Q_i \) as the Gramians for subsystem \((A, b, c, \ast)\) and the cross Gramian \( R_{ij} \), defined as
\[ R_{ij} = \sum_{k=0}^{\infty} A^k b_j c_i A^k, \]
which is given by the solution to the Sylvester equation
\[ A R_{ij} A - R_{ij} + b_j c_i = 0. \]

As shown in [3] it holds that
\[ P = \sum_{j=1}^{n_u} P_j, \quad Q = \sum_{i=1}^{n_y} Q_i, \]
and from the definition of cross-Gramian for subsystem \( G_{ij} \) it relates to \( P_j \) and \( Q_i \) through (see [14])
\[ R_{ij}^2 = P_j Q_i. \]

The Gramians have important physical interpretations which have been used to motivate the choice of interaction measures. The quantity \( x^T P^{-1} x \) describes how much control energy is needed to steer the system from rest to a given state \( x \) while \( x^T Q x \) gives how much energy is transfered to the output for an uncontrolled system starting at \( x \). Naturally, for control pairing, one would like to choose inputs which, with little effort, can have large effects in the system.
states as well as outputs that contain most information about the system states, and this joint effect is captured by the cross-Gramian.

The zero-state, $x(0) = 0$, impulse response, $u(k) = I_n \delta(0)$, for the system in (13) is given by

$$g(k) = \begin{cases} 
  D, & k = 0 \\
  CA^{k-1}B, & k > 0 
\end{cases}$$

(7)

the coefficients $g(k) \in \mathbb{R}^{n_y \times n_u}$ are known as the Markov parameters. Notice that the impulse response coefficients from input $j$ to output $i$ are given by $g_{ij}(k)$.

For a system with no direct term $(A,B,C,0)$, the effects of past inputs to present and future outputs are given by

$$\begin{bmatrix}
y(0) \\
y(1) \\
\vdots
\end{bmatrix} = \begin{bmatrix} 
  g(1) & g(2) & \cdots \\
  g(2) & g(3) & \cdots \\
  \vdots & \vdots & \ddots
\end{bmatrix} \begin{bmatrix} 
  u(-1) \\
  u(-2) \\
  \vdots
\end{bmatrix},$$

(8)

the matrix $H$ is known as the Hankel matrix. The Hankel matrix from input $j$ to output $i$ is given by

$$H_{ij} = \begin{bmatrix} 
  g_{ij}(1) & g_{ij}(2) & \cdots \\
  g_{ij}(2) & g_{ij}(3) & \cdots \\
  \vdots & \vdots & \ddots
\end{bmatrix},$$

(9)

which is a partition of $H$. 

2.2. Measures used and related norms

For a system \((A, B, C, 0)\), the Hankel norm is a measure of the maximum possible influence of past inputs to future inputs, given by

\[
\|G\|_H^2 = \max_{u(k)} \sum_0^\infty \|G(q)u(k)\|_2^2 = \rho(H^TH) = \rho(PQ),
\]

where \(\rho(\cdot)\) denotes the largest eigenvalue (spectral radius). The singular values of the Hankel matrix (HSV) coincide with the eigenvalues of the product \(PQ\) and the Hankel norm is also given by the largest HSV. The Hankel Interaction Index Array (HIIA), introduced in [15], is an interaction measure based on the Hankel norm. Here we define HIIA based on the squared Hankel norm, squared norms are used also for the other interaction measures presented,

\[
\Sigma_{ij}^H \triangleq \|G_{ij}\|_H^2 = \rho(H_{ij}^TH_{ij}) = \rho(P_jQ_i) = \max_l |\lambda_l(R_{ij})|
\]

where the last relation follows from (6).

The Hankel norm captures the effect of only the largest HSV. An aggregated measure would be to consider the effects of all HSVs. The sum of all HSVs relates to the squared Hilbert-Schmidt-Hankel norm. An index array defined based on this measure was suggested in [3] and is known as the participation matrix,

\[
\Sigma_{ij}^{PM} \triangleq \|G_{ij}\|_{HS}^2 = \text{tr}(H_{ij}^TH_{ij})
\]

\[
= \sum_{k=1}^\infty k g_{ij}(k)^2 = \begin{bmatrix} g_{ij}(1) & g_{ij}(2) & \cdots \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \ddots \end{bmatrix} \begin{bmatrix} g_{ij}(1) \\ g_{ij}(2) \\ \vdots \end{bmatrix} \triangleq W \triangleq g_{ij}(1 : \infty)
= \text{tr}(P_jQ_i) = \text{tr}(R_{ij}^2)
\]
\[ \frac{1}{\pi} \int_{0}^{2\pi} |G_{ij}(e^{j\nu})|^2 d \text{arg} G(e^{j\nu}) d\nu. \]  

(11c)

The last relationship denotes the area enclosed by the oriented Nyquist diagram and we introduced the notation \( W \) for a diagonal weighting matrix and \( g_{ij}(1 : \infty) \) for the vector of Markov parameters for subsystem \( G_{ij} \).

The last index array considered here is based on the square of the 2-norm of a system, defined as,

\[ \|G\|_2^2 \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} G^*(e^{j\nu})G(e^{j\nu}) d\nu = \sum_{k=0}^{\infty} \text{tr} g(k)^T g(k) \]

\[ = \text{tr}(B^T QB + D^T D) = \text{tr}(CPC^T + DD^T), \]

where the first equivalence follows from Parseval’s relation and the subsequent from the definition of Gramians. The 2-norm gives a measure of the total energy transferred to the output from an impulse input. For a SISO system, the 2-norm is the area under the amplitude Bode diagram. An interaction measure based on the squared 2-norm was proposed in [16] and is given by

\[ \Sigma_{ij}^2 \triangleq \|G_{ij}\|_2^2 = \sum_{k=0}^{\infty} g_{ij}^2(k) = g_{ij}^T(0 : \infty)g_{ij}(0 : \infty) \]

\[ = \text{tr}(b_j^T Q_i b_j + D_{ij}^2) = \text{tr}(c_i P_j c_i^T + D_{ij}^2) \]

\[ = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G_{ij}(e^{j\nu})|^2 d\nu. \]  

(12c)

2.3. Properties

As highlighted by equations (10a), (11a) and (12a), all measures used in the index arrays defined previously can be found based only on the Markov parameters. This is useful to show some of their properties. First, notice that a change of basis \( x' = Tx \) does not change the Markov parameters

\[ g'(k) = CT^{-1}(TAT^{-1})^{k-1}TB = CA^{k-1}B = g(k), \]
the measures are therefore realization independent. Furthermore, if \( G_{ij} \equiv 0 \), so are its markov parameters, yielding \( \Sigma_{ij}^* = 0 \) and the measures preserve the structure of the system, this is in contrast for instance with RGA. All measures are scaling dependent and an appropriate choice of scaling is recommended before CSS. The measures are however independent of frequency scaling, see [17].

Among the measures defined, \( \Sigma^2 \) is the only that relates to the presence of a direct term. From [12a], it is also easy to see that \( \Sigma^2 \) is not affected by a time delay, which is not the case for \( \Sigma^H \) and \( \Sigma^{PM} \). For a system with no direct term, and impulse response \( g'(k) = g(k - d) \), i.e. with a delay \( d \), it follows from [11a] that

\[
\|G_{ij}'\|_{HSH}^2 = \sum_{k=1}^{\infty} (k + d)g_{ij}^2(k) = d\|G_{ij}\|_2^2 + \|G_{ij}\|_{HSH}^2.
\]

The presence of delays is important for CSS, however it has been noted in [17] that the use of delay sensitive measures does not necessarily provide the best CSS recommendation and it may be advised to instead consider a combination of the interaction measures when performing CSS.

2.4. Computational paths and uncertainty propagation

As we seek for methods to compute the interaction measures from data and to quantify their uncertainties, we review next the main computational paths and existing approaches found in literature. The subequations found in [10], [11] and [12] already reveal the three main computational paths for the interaction measures, based on: a state-space realization, the Markov parameters and the frequency response. For robust CSS, the chosen representation of uncertainty is also important. We present next an overview of the different approaches; Table 1 summarizes the contributions found in the literature.

2.4.1. Based on a state space realization

The interactions measures are computed from the system matrices \((A, B, C, D)\) as in [10b], [11b] and [12b]. This approach is analytical, giving an exact solu-
tion but requires availability of a complete model. The measures $\Sigma^H$ and $\Sigma^{PM}$ can be found based on the product $P_jQ_i$, which requires solution of $2n_y n_u$ Lyapunov equations. An approach based on the cross-Gramian $R_{ij}$ requires solution of $n_y n_u$ Sylvester equations and is thus more attractive computationally. $\Sigma^2$ can be found from either $P_j$ or $Q_i$, requiring $n_y n_u$ Lyapunov equations. The uncertainties in a state-space realization are typically represented by additive matrix perturbations to the system matrices $(A + \Delta A, B + \Delta B, C + \Delta C, D + \Delta D)$, which are propagated through the measures, see, e.g., [14, 18].

2.4.2. Based on the Markov parameters

The interaction measures are computed from the Markov parameters, as in (10a), (11a) and (12a). For practical reasons, the Markov parameters are truncated as a finite series, which can be motivated from the stability of the process. This approach is most suitable for data-driven CSS, when the Markov parameters are found from input-output data. In [14, 10] simulations studies were carried out for dedicated open-loop identification experiments using white noise input applied to each input channel separately. The stochastic uncertainties from the estimated Markov parameters are propagated through the estimator of the measures to find confidence intervals as described in [10, 12].

2.4.3. Based on the frequency response

The interaction measures are computed from the frequency response functions (FRFs) of the system as in (11c) and (12c). Estimates based on the FRFs are better suited for data-driven CSS, where estimates of the FRFs are used. In [3, 11, 12], the interaction measures are found in simulation studies based on open-loop experiments using an orthogonal sinusoidal excitation. The FRFs are only evaluated at a finite number of frequencies and the integrals in (11c) and (12c) are discretized. A multiplicative uncertainty description $G_\Delta(\omega) = G(\omega)(1 + \Delta(\omega))$, has been considered in [3, 11, 12] to propagate uncertainties to the interaction measure estimates, where $\Delta(\omega)$ is associated to the uncertain in the estimated FRFs. Note that although no frequency domain rep-
Table 1: Summary of contributions for Gramian-based CSS.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Index array</th>
<th>Computation</th>
<th>Uncertainty</th>
<th>Input</th>
<th>Loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>[15]</td>
<td>$\Sigma^H$</td>
<td>SS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[16]</td>
<td>$\Sigma^H$, $\Sigma^2$</td>
<td>SS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[3]</td>
<td>$\Sigma^{PM}$</td>
<td>SS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[9]</td>
<td>$\Sigma^{PM}$</td>
<td>SS ($R_{ij}$)</td>
<td>additive</td>
<td>white noise</td>
<td>open</td>
</tr>
<tr>
<td>[14]</td>
<td>$\Sigma^{PM}$</td>
<td>SS ($R_{ij}$)</td>
<td>additive</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[18]</td>
<td>$\Sigma^{PM}$</td>
<td>SS ($R_{ij}$)</td>
<td>additive</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[10]</td>
<td>$\Sigma^{PM}$</td>
<td>MP</td>
<td>stochastic</td>
<td>white noise</td>
<td>open</td>
</tr>
<tr>
<td>[5]</td>
<td>$\Sigma^{PM}$</td>
<td>FRF</td>
<td>mult/stoch</td>
<td>ortho. sinus.</td>
<td>open</td>
</tr>
<tr>
<td>[11]</td>
<td>$\Sigma^2$</td>
<td>FRF</td>
<td>mult/stoch</td>
<td>ortho. sinus.</td>
<td>open</td>
</tr>
<tr>
<td>[12]</td>
<td>$\Sigma^{PM}$, $\Sigma^2$</td>
<td>FRF</td>
<td>mult/stoch</td>
<td>ortho. sinus.</td>
<td>open</td>
</tr>
<tr>
<td>this</td>
<td>$\Sigma^H$, $\Sigma^{PM}$, $\Sigma^2$</td>
<td>MP</td>
<td>stochastic</td>
<td>general</td>
<td>any</td>
</tr>
</tbody>
</table>

A presentation was given for the Hankel norm it is also possible to find estimates of the Hankel singular values from frequency data as described in [19].

In the next sections we outline approaches to find the interaction measures based on estimated Markov parameters of a predictor model from input-output data collected in open or closed-loop. Unbiased estimators are given for the $\Sigma^{PM}$ and $\Sigma^2$ and an approximate statistical description of uncertainties is also described.

3. Data-driven estimation based on predictor Markov parameters

In this section we present a framework for the estimation of interaction measures based on input-output data collected in open or closed-loop. Essentially, it is based on the estimation of the Markov parameters from a vector ARX (VARX) predictor model. In order to handle large datasets, a recursive formulation is provided. Estimators for the interaction measures are defined and their statistical properties are investigated. It is considered that the input-output relations in the data can be described by the state-space model under additive uncertainties,

$$x(k + 1) = Ax(k) + Bu(k) + w(k)$$  (13a)
\[ y(k) = Cx(k) + Du(k) + v(k), \quad (13b) \]

where \( w(k) \) and \( v(k) \) are white noise signals. When the pair \((A, C)\) is observable, the model in (13) can be represented in innovations form as

\[
\begin{align*}
\hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + K\epsilon(k) \quad (14a) \\
y(k) &= C\hat{x}(k) + Du(k) + \epsilon(k) \quad (14b)
\end{align*}
\]

where \( K \) is the Kalman gain and \( \epsilon(k) \) is a white sequence. From the innovations equation, \( \epsilon(k) = y(k) - C\hat{x}(k) - Du(k) \), the innovation form can be rewritten as the (optimal) predictor form,

\[
\begin{align*}
\hat{x}(k+1) &= \tilde{A}\hat{x}(k) + \tilde{B}u(k) + Ky(k) \quad (15a) \\
y(k) &= C\hat{x}(k) + Du(k) + \epsilon(k). \quad (15b)
\end{align*}
\]

where \( \tilde{A} = A - KC \) and \( \tilde{B} = B - KD \) are stable matrices. The Markov parameters for the predictor form are given by

\[
\tilde{g}(k) = \begin{cases} 
D, & k = 0 \\
C\tilde{A}^{k-1}\tilde{B} + \sum_{i=1}^{k}C\tilde{A}^{i-1}K\tilde{g}(k-i), & k > 0
\end{cases} \quad (16)
\]

we introduce the more compact notation \( \Theta_i \triangleq C\tilde{A}^{i-1}\tilde{B} \) with \( \Theta_0 = D \) and \( \Psi_i \triangleq C\tilde{A}^{i-1}K \) for the matrices used in the calculations of the predictor Markov parameters. For the predictor model (15), it can be shown, see e.g. [20], that the relationship between \( u(k) \) and \( y(k) \) is maintained, i.e.

\[
\tilde{G}(q) = \left( I - C(qI - \tilde{A})^{-1}K \right)^{-1} \left( C(qI - \tilde{A})^{-1}\tilde{B} + D \right) = C(qI - A)^{-1}B = G(q). \quad (17)
\]

It is therefore equivalent to analyse the input-output behavior in predictor form or in the original state-space form. An advantage of the predictor form is that
the innovation vector $\epsilon(k)$ is white even in closed-loop, this allows for straightforward identification of systems in closed-loop as outlined next.

### 3.1. Input-output description of the data

By iterating (15a) from $k$ to $k+p$, one obtain the following input-output representation

$$y(k+p) = C\tilde{A}^p x(k) + \sum_{i=0}^{p} \Theta_i u(k+p-i) + \sum_{i=1}^{p} \Psi_i y(k+p-i) + \epsilon(k+p)$$

which describes $y(k+p)$ based on the starting state through $C\tilde{A}^p x(k)$, and on the input-output data \{u(k)\}$_k^{k+p}$, \{y(k)\}$_k^{k+p-1}$ through $\Theta_i$ and $\Psi_i$. Since $\tilde{A}$ is asymptotically stable, $\tilde{A}^p \approx 0$ for a large enough $p$, i.e. the effect of $x(k)$ can be made insignificant. Defining the data vector

$$z^{(p)}(k+p) \triangleq \begin{bmatrix} u(k+p)^T & \cdots & u(k)^T & y(k+p-1)^T & \cdots & y(k)^T \end{bmatrix}^T,$$ (18)

with dimension $n \triangleq n_u(p+1) + n_y p$, for a large enough $p$, the data relation can be written as

$$y(k+p) = \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} z^{(p)}(k+p) + \epsilon(k+p).$$ (19)

This is a high-order VARX model. For a dataset with $N$ samples, this can be repeated from $k = p + 1$ to $k = p + N_p$, with $N_p \triangleq N - p$ the effective number of data points, to yield

$$Y_{p+1:p+N_p} = \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} Z^{(p)}_{p+1:p+N_p} + E_{p+1:p+N_p},$$ (20)

where $Y_{p+1:p+N_p} \triangleq \begin{bmatrix} y(p+1) & \cdots & y(p+N_p) \end{bmatrix}$ and similarly for $Z^{(p)}_{p+1:p+N_p}$ and $E_{p+1:p+N_p}$. The unknowns $\Theta_{0:p}, \Psi_{1:p}$ describe the input-output relations in the data.
3.2. Estimation of high-order VARX

Under the condition that the innovations $\epsilon(k)$ are white and that the data matrix $Z_{p+1:p+N_p}^{(p)}$ is persistently exciting, see e.g. [21], the matrices $\Theta_{0:p}, \Psi_{1:p}$ can be consistently estimated using a least squares criterion, i.e.

$$\min_{[\Theta_{0:p} \Psi_{1:p}]} \left\| Y_{p+1:N_p} - \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} Z_{p+1:p+N_p}^{(p)} \right\|_F^2. \quad (21)$$

Notice that, for consistency, a full rank data matrix is needed and thus $N_p \geq n$.

The least squares estimate can be found from an RQ factorization as follows. Since the norm in the minimization is not affected by an orthonormal transformation $Q$, $QQ^T = I$, find the RQ factorization to

$$\begin{bmatrix} Z_{p+1:p+N_p}^{(p)} \\ Y_{p+1:p+N_p} \end{bmatrix} = RQ \quad (22)$$

where $R$ is a matrix of the form

$$R = \begin{bmatrix} R_0 & : & 0 \\ \end{bmatrix}, \quad R_0 = \begin{bmatrix} R_1 & 0 \\ R_2 & R_3 \end{bmatrix} \quad (23)$$

the matrix $R_0$ is square lower triangular, $R_1 \in \mathbb{R}^{n \times n}$ and $R_3 \in \mathbb{R}^{n_y \times n_y}$. Applying the orthonormal transformation $Q^T$ from the right in (21), gives

$$\left\| Y_{p+1:N_p}^T Q - \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} Z_{p+1:p+N_p}^{(p)} Q^T \right\|_F^2$$

$$= \left\| \begin{bmatrix} R_2 & R_3 \\ \end{bmatrix} - \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} \begin{bmatrix} R_1 & 0 \end{bmatrix} \right\|_F^2 = \left\| R_2 - \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} R_1 \right\|_F^2 + \left\| R_3 \right\|_F^2,$$

which has a minimum value at $\left\| R_3 \right\|_F^2$, achieved for the estimate of $\begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix}$ satisfying the equation

$$R_2 = \begin{bmatrix} \Theta_{0:p} & \Psi_{1:p} \end{bmatrix} R_1. \quad (24)$$

It is known that the RQ solution to the least squares presents superior numerical
conditioning compared to a direct solution to the standard normal equations, see e.g. [21]. The complexity for the RQ factorization in (22) is quadratic in \( N_p \) and cubic in \( n + n_y \). For analysis based on large datasets, for example based on operational data collected over days, months or years, the size of \( N_p \) might be prohibitive to allow for the outlined batch solution. It is possible to modify the algorithm to allow for a formulation which is recursive in the data.

3.2.1. Recursive implementation

Given the RQ solution at time instance \( k-1 \), denoted by the pair \( R(k-1), Q(k-1) \), it is possible to update the estimate at time \( k \) recursively as follows. The cost function in (21) can be written as

\[
\| \begin{bmatrix} Y_{p+1:k-1} & y(k) \end{bmatrix} - \begin{bmatrix} \Theta_0:p & \Psi_1:p \end{bmatrix} \begin{bmatrix} Z_{p+1:k-1}^{(p)} & z^{(p)}(k) \end{bmatrix} \|_F^2.
\]

Applying the orthonormal transformation

\[
\bar{Q}^T = \begin{bmatrix} Q^T(k-1) & 0 \\ 0 & 1 \end{bmatrix}
\]

to the criterion, decomposes the problem as

\[
\| \begin{bmatrix} R_2(k-1) & R_3(k-1) & y(k) \end{bmatrix} - \begin{bmatrix} \Theta_0:p & \Psi_1:p \end{bmatrix} \begin{bmatrix} R_1(k-1) & 0 & z^{(p)}(k) \end{bmatrix} \|_2^2
\]

which depends only on \( R_0(k-1), z^{(p)} \) and \( y(k) \) and thus applying the RQ factorization

\[
\begin{bmatrix} R_0(k-1) & z^{(p)}(k) \\ y(k) \end{bmatrix} = R(k)Q(k)
\]

(25)

gives the updated estimates as the solution to \( R_2(k) = \begin{bmatrix} \Theta_0:p & \Psi_1:p \end{bmatrix} R_1(k) \).

Notice that the left-hand side of (25) has fixed size \( n + n_y \times n + n_y + 1 \) for all \( k > n \) and the problem does not increase in size with the data.
3.3. Statistical properties of least squares VARX estimates

Provided that the order $p$ is large enough so that $C\hat{A}^p \approx 0$, that the residuals $\epsilon(k)$ are white (guaranteed by the Kalman filtering theory) and that the data $Z_{p+1:N_p}^{(p)}$ is informative enough (see [13] for a review on informativeness of data for process identification from operational data), the asymptotic distribution of the least squares estimates $[\hat{\Theta}_{0:p}, \hat{\Psi}_{1:p}]$ is Gaussian (see e.g. [22, section 10.3]). Denoting $\beta = \text{vec}[\Theta_{0:p}, \Psi_{1:p}]$ the vector of parameters stacked columnwise, with $\beta_0$ meaning the true values and $\hat{\beta}$ the least squares estimate, then

$$\sqrt{N_p}(\hat{\beta} - \beta_0) \overset{d}{\rightarrow} N(0, V_{\beta})$$

(26)

where

$$V_{\beta} = E \left[ \frac{1}{N_p} Z_{p+1:N_p}^{(p)} Z_{p+1:N_p}^{(p)\top} \otimes \Xi_{\epsilon}^{-1} \right]^{-1}$$

(27)

$\Xi_{\epsilon}$ is the residuals covariance and $\otimes$ denotes the Kronecker product. The estimate $\hat{\beta}$ is thus unbiased. For a solution based on the RQ factorization, it is straightforward to show that the empirical estimate of the covariance matrix is given by $\hat{V}_{\beta} = \frac{1}{N_p} \hat{V}_{\beta}$, with

$$\hat{V}_{\beta} = \left( \frac{1}{N_p} R_1 R_1^T \right)^{-1} \otimes \hat{\Xi}_{\epsilon}, \quad \hat{\Xi}_{\epsilon} = \frac{1}{N_p} R_3 R_3^T$$

(28)

and can thus be computed directly from the RQ factorization.

3.4. Unbiased estimates of the predictor Markov parameters

The identified parameters $\hat{\beta}$ can be used to find estimates of the predictor Markov parameters $\hat{g}(0:p)$ through equation (16). Even though $\hat{\beta}$ is Gaussian distributed and unbiased, the plug-in estimate of $\hat{g}(k)$ is biased and not Gaussian. Denoting the plug-in estimate by $\hat{g}(k)$, take for example the estimate for
the second term, even for the case of a diagonal $\Xi_\beta$, it gives

$$
\hat{g}(2) = \Theta_2 + \Psi_1 \hat{g}(1) + \Psi_2 \hat{g}(0) = \Theta_2 + \Psi \Theta_1 + (\Psi_1^2 + \Psi_2)\Theta_0,
$$

$$
E\{\hat{g}_2\} = E\{\Theta_2\} + E\{\Psi_1\}E\{\Theta_1\} + (E\{\Psi_1\}^2 + E\{\Psi_2\})E\{\Theta_0\} + \text{Var}\{\Psi_1\}E\{\Theta_0\},
$$

which is biased by $\text{Var}\{\Psi_1\}E\{\Theta_0\}$. It is possible to find the bias term, $b(k)$, for each $\hat{g}(k)$ analytically. Unfortunately, this requires long and tedious derivations which are better suited for a symbolic scientific computing tool. The Mathematica code snippet found in the Appendix was used here to calculate $b(k)$, the resulting unbiased estimate is then given by

$$
\hat{\hat{g}}(k) = \hat{g}(k) - b(k)
$$

where $\hat{g}(k)$ is the plug-in estimate. The covariance of the resulting estimate, $\Xi_{\hat{g}(0:p)}$, can also be found analytically, the code in the Appendix includes its analytical calculation.

### 3.5. Estimates of the interaction measures

Since for large enough $k$, $\hat{A}_k^{-1} \approx 0$, the Markov parameters $\hat{g}(k)$ will also tend to zero with $k$. The truncated Markov parameters $\hat{g}(0:p)$ can thus be used to estimate the interaction measures for a large enough $p$. We start by noticing that the unnormalized interaction measures defined in Section 2.2 depend on inner products of the Markov parameters as given by the rightmost expressions in equations (10a), (11a) and (12a). For a truncated series of Markov parameters, the relevant quantities are given by

$$
\Sigma^H : [H_{ij}^T H_{ij}]_{kl} = \begin{bmatrix} \hat{g}_{ij}^T(k;p) & 0^T \end{bmatrix} \begin{bmatrix} \hat{g}_{ij}^T(l;p) & 0^T \end{bmatrix}^T
$$

$$
\Sigma^{PM} : \hat{g}_{ij}^T(1:p) W \hat{g}_{ij}(1:p),
$$

$$
\Sigma^2 : \hat{g}_{ij}^T(0:p) \hat{g}_{ij}(0:p),
$$

(29a) (29b) (29c)
with an adequate number of trailing zeros in (29a). Each of these can be written as a quadratic form of random variables. For random vectors \( v, w \), the quadratic form \( v^T A w \) has expected value

\[
E\{v^T A w\} = \text{tr}(A \text{Cov}\{v, w\}) + E\{v\}^T A E\{w\} \tag{30}
\]

and unbiased estimates for the quantities in (29) are thus possible by subtracting traces of covariances, variances and weighted variances from the quadratic forms.

For instance, this directly gives unbiased estimates of \( \hat{\Sigma}^{PM} \) and \( \hat{\Sigma}^2 \),

\[
\hat{\Sigma}_{ij}^* = \hat{g}_{ij}^T(0:p) D \hat{g}_{ij}(0:p) - \text{tr} \left( D \text{Var}(\hat{g}_{ij}(0:p)) \right) \tag{31}
\]

with \( D = I_p \) for \( \hat{\Sigma}^2 \) and a diagonal with \( D_{ii} = i - 1 \) for \( \hat{\Sigma}^{PM} \). An unbiased estimate for each element of \( H^T_i \mathcal{H}_{ij} \) used for \( \Sigma^H \) is also available from these relations but an unbiased estimate of its spectral radius depends on the chosen method to compute eigenvalues, e.g. the power iteration, see [23]. The resulting statistical properties arising from such estimate of \( \hat{\Sigma}^H \) are intricate and are not considered here.

A full characterization of the statistical distribution of the estimates for the unnormalized \( \hat{\Sigma}^{PM} \) and \( \hat{\Sigma}^2 \) given by the quadratic form in (31) is only possible for some special cases. Making use of a Gaussian approximation, the final distribution is tractable but not analytical, see e.g. [23, 25, 10] and references therein. Under a Gaussian approximation, the variance is however analytical. Let \( \Xi_{\hat{g}_{ij}(0:p)} \) be the covariance for \( \hat{g}_{ij}(0:p) \), then, based on a second-order (Gaussian) approximation, the variance of the estimate in (31) is given by

\[
\text{Var}\{\hat{\Sigma}_{ij}^*\} = 2 \text{tr}(D \Xi_{\hat{g}_{ij}(0:p)} D \Xi_{\hat{g}_{ij}(0:p)}) + 2 \hat{g}_{ij}^T(0:p)^T D \Xi_{\hat{g}_{ij}(0:p)} D \hat{g}_{ij}(0:p). \tag{32}
\]

We use the above expression as a measure for the uncertainty of \( \hat{\Sigma}^{PM} \) and \( \hat{\Sigma}^2 \) and standard error propagation formulas for sums and divisions, see e.g. [24], are used to compute the errors for the normalized interactions measures.
4. Extensions

We present some different extensions to the proposed framework for data-driven control selection structure.

4.1. Filtering

The interaction measures presented consider the entire range of frequencies. In many situations it is however preferable to perform CSS based on a limited frequency band, for instance where the control energy can affect the plant or where the model of the plant is reliable. There are different ways to achieve frequency selective interactive measures. When a model is available, it has been suggested to apply a filter to the input or output of the system as in [15] or use frequency weighted Gramians as in [16, 27].

For interaction measures found from input-output data, it is possible to pre-filter the input or output before estimation of the FRFs or MPs. An alternative is to filter the FRFs or MPs after their estimation from unfiltered data. For MP based estimates, given unfiltered MPs $g(k)$ and a frequency selective filter in finite impulse response form, $f(k)$, the filtered MPs are given by the convolution $g_f(k) = (g * f)(k)$. It is therefore simple to check for the effect of different filters as a new identification procedure is avoided.

4.2. Orthonormal basis approximations

The transfer function for the predictor model in (15) is given by

$$y(k) = C(qI - \tilde{A})^{-1} \begin{bmatrix} \tilde{B} & K \end{bmatrix} \begin{bmatrix} u(k) \\ g(k) \end{bmatrix} + Du(k) + \epsilon(k)$$

(33)

applying the Taylor expansion at the resolvent of the stable matrix $\tilde{A}$,

$$(qI - \tilde{A})^{-1} = \sum_{i=1}^{\infty} q^{-i} \tilde{A}^{i-1},$$

(34)

and truncating the expression to $p$ reduces to the VARX model in (20). It is possible to approximate this resolvent using different expansions of orthonormal
polynomials, for example based on Laguerre, \[28\], and Kautz polynomials, \[29\]. Models based on a Laguerre polynomial approximation present some advantages for systems with delays and can allow for a more compact representation, i.e. requiring smaller values of $p$. In \[13\], a mixed representation was considered for mining operational data for identification of SISO loops from routine operational data.

4.3. Relation to subspace identification

The identification of high order VARX predictor models is closely related to a class of subspace identification methods \[30, 31\]. Starting at $\hat{x}(k)$ and moving to $\hat{x}(k + p)$ and pre-multiplying by $C$, the state transition equation for the predictor form in \[15a\] gives

$$C\hat{x}(k + p) = \begin{bmatrix} \Theta_1 & \cdots & \Theta_p & \Psi_1 & \cdots & \Psi_p \end{bmatrix} \begin{bmatrix} u(k + p - 1) \\ \vdots \\ u(k) \\ y(k + p - 1) \\ \vdots \\ y(k) \end{bmatrix} + C\hat{x}(k) \approx 0$$

notice that $\bar{z}(p)(k + p)$ is readily available. The above relation can be applied for the entire dataset $\bar{Z}(p)(k + p : N_p)$ and shifted into the future $f$ times, giving

$$\bar{X}(k + p : N_p) = \begin{bmatrix} C \\ C\hat{A} \\ \vdots \\ C\hat{A}^{f-1} \end{bmatrix} \begin{bmatrix} \Theta_1 & \cdots & \Theta_p & \Psi_1 & \cdots & \Psi_p \\ \Theta_2 & \cdots & \Theta_{p+1} & \Psi_2 & \cdots & \Psi_{p+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Theta_f & \cdots & \Theta_{p+f} & \Psi_f & \cdots & \Psi_{p+f} \end{bmatrix} \bar{Z}(p)(k + p : N_p)$$

$$\triangleq \Theta\Psi$$

$$\triangleq C_f$$
where \( C_f \) is the extended observability matrix. Since \( \Theta_{k>p}, \Psi_{k>p} \approx 0 \), the right hand-side of the above equation can be computed based on the estimate \( \hat{\Theta}_{1:p}, \hat{\Psi}_{1:p} \) and the input-output data. An estimate of an \( n_x \) dimensional state vector is possible from the singular value decomposition

\[
C_f \hat{X}(k + p : N_p) = \hat{\Theta}\Phi \hat{Z}^{(p)}(k + p : N_p) = U_n s_n v_{n_x}^T \tag{35a}
\]

\[
\Rightarrow C_f = U_n, \quad \hat{X}(k + p : N_p) = S_n v_{n_x}^T, \tag{35b}
\]

From the estimated states, the \( A, B, C, D \) and \( K \) matrices can be found from the solution of the equations below,

\[
Y_{p:N_p} = \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} \hat{X}_{p:N_p} \\ U_{p:N_p} \end{bmatrix} + E_{p:N_p}, \tag{36a}
\]

\[
\hat{X}_{p+1:N_p-1} = \begin{bmatrix} A & B \\ K \end{bmatrix} \begin{bmatrix} \hat{X}_{p:N-1} \\ U_{p:N_p-1} \end{bmatrix}, \tag{36b}
\]

where (36a) is solved first and its residuals \( \hat{E}_{p:N_p} \) are used in (36b). Notice that (36b) is based on the innovations form given in (14) instead of the predictor form (15). This is possible since the states for these two forms are the same, the advantage is that \( A, B \) are found directly from the innovations form instead of \( \tilde{A}, \tilde{B} \) from the predictor form.

The procedure outlined above is known as the PBSID\textsubscript{opt} algorithm, which has been studied extensively in the literature. A recursive implementation of the algorithm was given in [32]. In [33], the asymptotic variance of the estimated matrices are studied, which is useful for robust CSS.

Notice that, while it might be tempting to estimate a state-space model from the data, the operations involved require the solution to an additional singular value decomposition and two additional least squares. For the purpose of estimating the interaction measures, those additional steps are not necessary as described in this paper.
5. Illustrative example - a bark boiler process

This example was taken from [7], where the control reconfiguration of a bark boiler system is studied. A model of the plant was achieved through dedicated identification experiments, making use of step responses in open-loop (the system is stable). The model has no direct term \( D = 0 \), \( n_x = 8 \), \( n_u = n_y = 4 \), see [7] for the values used for the \( A, B, C \) and the scaling matrices. The Gramian-based interaction measures for this system are shown in Figure 1 as intensity maps. We study the estimation of interaction measures for this system based on data generated from simulations. We compare the estimation achieved based on the predictor Markov parameters with those given by a subspace model from the PBSID method outlined in Section 4.3.

5.1. Simulation setup

The control configuration shown in Figure 2 is considered for the simulation study. The first two outputs \( y_1 \) and \( y_2 \) are operated in closed-loop through the reference signals \( r_1 \) and \( r_2 \). The controller is given by

\[
F_{11}(z) = \frac{1.4z - 0.4}{z - 1}, \quad F_{22}(z) = \frac{3z - 1.5}{z - 1}.
\]
The lower outputs $y_3$ and $y_4$ are left uncontrolled but excited directly through the inputs $u_3$ and $u_4$. This configuration allows for an independent excitation of each subsystem, which is important in order to achieve an informative dataset, recall Section 3.3. Open as well as closed loops are considered to illustrate the versatility of the method.

A sequence of stepwise changes to the references and inputs are applied to the system, generating $N = 2400$ data points. White Gaussian process and measurement uncertainties, $w(k)$ and $v(k)$, are added with covariances $0.1I_{n_x}$ and $0.1I_{n_y}$ respectively. The generated data can be seen in Figure 3.

5.2. Estimation of interaction measures based on predictor Markov parameters

Given the input-output data, a VARX model of order $p = 10$ with $n = 84$ parameters is used. The model is identified using the recursive RQ-factorization as described in Section 3.2.1, providing also an estimate for their covariance matrix as described in Section 3.3. From the retrieved parameters, the unbiased predictor Markov parameters are calculated, together with an estimate of their covariance matrix as described in Section 3.4. Finally, the interaction measures are calculated from the estimates of the predictor Markov parameters as described in Section 3.5. The absolute values for the estimation error, $|\Sigma^* - \hat{\Sigma}^*|$, and uncertainties for $\hat{\Sigma}_P^2$ and $\hat{\Sigma}_{PM}$ are presented in the form of intensity maps in Figure 4. Notice the small estimation error and confidence bounds within the error region., indicating a successful estimation of the interaction measures.
5.2.1. Estimation of interaction measures based on a subspace model

For a comparison, we find the interaction measures also from a state-space model identified from the data. For that purpose, we make use of the predictor Markov parameters identified in the previous section and use the PBSID method described in Section 4.3 to find a state-space model. The orders \( p = f = 10 \) are used, leading to a matrix \( \Theta \Psi \) of dimensions \( (n_y f \times n_u p + n_y p) = (40 \times 80) \). The normalized cumulative sum of singular values for \( \Theta \Psi \bar{Z}^{(p)}(k+p:N_p) \) are shown in Figure 5 and the model order \( n_x = 4 \) is chosen. The interaction measures are calculated from the estimated \( (A, B, C, D) \) matrices and the absolute estimation errors \( |\Sigma^* - \hat{\Sigma}^*| \) are shown in Figure 6 as intensity maps, compare with Figure 4. Boxplots for the absolute estimation errors achieved through both methods are shown in Figure 7 for a direct comparison. The errors achieved from the PBSID method are considerably larger. As a subspace approach also requires additional computations, unless a state-space model is needed, there are no
Figure 4: Estimation error for the interaction measures found using the predictor Markov parameters (top row) and uncertainty estimate for $\Sigma^2$ and $\Sigma^{PM}$ (bottom row). Notice the different scales for each row.
Figure 5: Normalized cumulative sum of singular values of $\Theta \Psi \tilde{Z}^{(p)} (k+p: N_p)$ used for model order selection.

Figure 6: Estimation error for the interaction measures found using a subspace model from the PBSID method. The same scale is used for the errors in Figure 4.

immediate benefits in finding a state-space model to compute the interaction measures.

6. Conclusions and future work

Requiring a minimal specification from the user, only the VARX order $p$ is needed, this paper presents a flexible approach for data-driven estimation of interaction measures. The same approach can handle data in open and closed loop to estimate three important Gramian-based interaction measures and the uncertainty estimates allow for robust CSS. The effects of different filtering to the data are also easily incorporated. The very general conditions for excitation
in the data and a recursive implementation also make it possible to apply to a variety of situations.

The main computational step lies in the calculation of the covariance matrix for the VARX model. For large-scale systems, with several inputs and outputs and with a large $p$, the calculations might be prohibitive. Simplifications to the structure of the covariance can be introduced to tackle this problem.

Given the flexibility of the presented approach, it can be used as a framework to further develop methods that can find interaction measures from operational data. In this direction, additional tests can be defined for the identification procedure to check for pre-conditions on the data quality and pos-conditions on the model consistency. Considering the multivariate nature of interaction analysis and the scarcity of segments of operational data with simultaneous excitation of all input channels, it is important to devise an strategy for segmenting and merging data segments containing quality data.

Appendix

References

Algorithm 1 Finds the unbiased estimator $\hat{g}(0:p)$ and its covariance $\Xi_{\hat{g}(0:p)}$

Given the order $p$

$\rightarrow$ Defines VARX variables (true and estimated)

$\beta = \text{Flatten}[[\text{Array}[\Theta, p + 1, 0], \text{Array}[\Psi, p]]]$;
$\hat{\beta} = \text{Flatten}[[\text{Array}[\hat{\Theta}, p + 1, 0], \text{Array}[\hat{\Psi}, p]]]$;

$\rightarrow$ Defines the covariance matrix for the VARX parameters

$n = \text{Length}[\beta]; \Sigma = \text{Array}[\sigma, \{n, n\}]$;

$\rightarrow$ Markov parameters recursions (true and empirical)

$\tilde{g}[k] := \Theta[k] + \text{Sum}[\Psi[i] \ast \tilde{g}[k - i], \{i, 1, k\}]$;
$\hat{\tilde{g}}[k] := \hat{\Theta}[k] + \text{Sum}[\hat{\Psi}[i] \ast \hat{\tilde{g}}[k - i], \{i, 1, k\}]$;

$\rightarrow$ True Markov parameters and empirical estimator

$\tilde{g}(0:p) = \text{Array}[^{\tilde{g}}, p + 1, 0]$;
$\hat{\tilde{g}}(0:p) = \text{Array}[\hat{\tilde{g}}, p + 1, 0]$;

$\rightarrow$ Finds the bias of the empirical estimator

$b = \text{Mean}[\text{TransformedDistribution}[\hat{\tilde{g}}(0:p), \hat{\beta} \sim \text{MultinormalDistribution}[\beta, \Sigma]]] - \tilde{g}(0:p) \text{// FullSimplify};$

$\rightarrow$ Analytical function for the unbiased estimator

$\hat{g}(0:p) = \hat{\tilde{g}}(0:p) - b$;

$\rightarrow$ Distribution of $\hat{g}(0:p)$ and its covariance

$\hat{G} = \text{TransformedDistribution}[\hat{g}(0:p), \hat{\beta} \sim \text{MultinormalDistribution}[\beta, \Sigma]]$;
$\Xi_{\hat{g}(0:p)} = \text{Covariance}[\hat{G}]$;


Title: Data-driven estimation of Gramian based interaction measures for control structure selection

Abstract: Interaction measures quantify the input-output relations in MIMO processes and can support the selection of control structures (CSS). Interaction measures are typically computed based on an existing process model. The study of input-output interactions based on data can complement missing information on a model, e.g., revealing unknown relations in a complex system or adjusting for a time dependent behavior. This paper presents a unified approach for data-driven estimation of Gramian based interaction measures from input-output data. Given open or closed-loop data, a high-order Vector ARX (VARX) model is identified and its parameters are used to calculate predictor Markov parameters, together with a covariance estimate. Three interaction measures based on the Hankel, Hilbert-Schmidt-Hankel and $H_2$ norms are calculated from the estimated predictor Markov parameters and uncertainty estimates are provided for the last two, allowing for robust CSS. A solution which is recursive in the data points is presented, making it practical for applications to large datasets. The approach is verified through simulations and several possible extensions are discussed. As the method is suitable for open and closed-loop data and for large datasets, it opens up for data-driven control structure selection based on operational data from entire plants.

Keywords: interaction measures, control structure selection, data-driven, Gramians, VARX models