Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 54-7 (2021) 316–321

Identifiability of Unique Elements of Noise Covariances in State-Space Models

Oliver Kost, Jindřich Duník, Ondřej Straka *

* *Department of Cybernetics and Research Centre NTIS, Faculty of Applied Sciences, University of West Bohemia, Univerzitní 8, 306 14 Pilsen, Czech Republic (e-mail: {kost,dunikj,straka30}@kky.zcu.cz)*

Abstract: This paper deals with identification of noise covariance matrices of a dynamic system described by a linear discrete-in-time time-invariant stochastic state-space model. In particular, the parametric identifiability of the correlations methods is analysed and explicit relations for determination of a number of identifiable noise covariance matrices parameters are stated. The theoretical results are thoroughly discussed and illustrated.

Keywords: Linear systems, Noise Covariance Matrices, Correlation methods, Identifiability.

1. INTRODUCTION

Knowledge of a system description is an integral part of a design of modern optimal signal processing and decisions making algorithms for state estimation, fault detection, and automatic control. An incorrect system description can lead to deteriorated behaviour of the algorithms or even in their instability.

The state-space model is designed to consistently describe a combination of deterministic and stochastic impacts affecting the system quantities and sensor properties. As such, the model can be virtually viewed as a composition of two sub-models. The first one describes deterministic (and thus predictable) system behaviour that often arises from the first principles. The second one characterises random behaviour of involved quantities, but it is usually difficult to obtain it from the first principles modelling and, instead, it is *identified* from data.

Therefore, in the last fifty years, a considerable research interest has been focused on a design of methods identifying properties of the state and measurement noises with an emphasis on the noise covariance matrices [1] (CMs) identification. In the literature, four main classes of the identification methods can be found; namely Bayesian, covariance matching, maximum likelihood, and *correlation methods*. This paper focuses on the class of the correlation methods developed in a number of paper since the seventies and reviewed in (Duník et al., 2017). The correlation methods are, compared to other classes, derived analytically with minimal assumptions on the model. As a consequence, the correlation methods may, under mild assumptions, provide consistent and unbiased estimates of the noise CMs.

Despite the long-term development of the correlation methods (e.g., design for time-invariant/time-varying models, linear/nonlinear models, white/correlated noise), the question of the maximum number of identifiable noise CMs elements remains still open. Focussing on the linear *time-invariant* [2] (LTI) models, the noise CM elements identifiability by the correlation methods has been discussed rather vaguely or under certain assumptions. In (Mehra, 1970; Bélanger, 1974; Lee, 1980), the identifiability of *all* elements of the noise CMs was assessed without any specific assumptions laid on the model. However, as it is shown later, the result is not directly extendable for identifiability assessment of the unique noise CMs elements, which are of importance. Recently, identifiability of the *unique* noise CMs elements was considered in (Arnold and Rawlings, 2018) and sufficient conditions for identifiability of the unique elements were provided in a view of one particular correlation method. However, the number of identifiable elements remains *unknown*, if the assumptions considered there are not met or another correlation method is used.

The goal of the paper is, therefore, twofold:

(1) To analyse parametric identifiability of unique elements of the noise CMs for the LTI model by the correlation methods,
(2) To provide explicit relations determining number of identifiable noise CMs elements based on the model properties.

The rest of the paper is organised as follows. In Section II, the system description is provided and task of the noise CMs estimation is particularised. Then, in Section III, the correlation methods are discussed, analysed, and illustrated. In Section IV, parametric identifiability of the noise CMs is treated. Section V summarises the identifiability results and Section VI concludes the paper.

2. SYSTEM DESCRIPTION AND NOISE COVARIANCE MATRICES IDENTIFICATION

In this paper, the system described by the following LTI discrete-in-time stochastic dynamic state-space model

[1] Typically, the noise CMs are estimated under assumption of the zero-mean distribution.

[2] From the perspective of the noise CMs parameters identifiability, the methods for the LTI models are the most challenging. For linear time-varying models, it is, generally, possible to estimate all *unique* elements of the noise CMs (Duník et al., 2018, 2020).

10.1016/j.ifacol.2021.08.378

with additive noises

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{w}_k, \quad k = 0, 1, 2, \ldots, \tau, \tag{1}$$

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k, \quad k = 0, 1, 2, \ldots, \tau, \tag{2}$$

is considered, where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$, and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the immeasurable state of the system and the available measurement at time instant k, respectively. The state and measurement matrices $\mathbf{F} \in \mathbb{R}^{n_x \times n_x}$ and $\mathbf{H} \in \mathbb{R}^{n_z \times n_x}$ are known and bounded. The system state is assumed to be *observable*, i.e., the following assumption holds:

Assumption 1: The observability matrix

$$\mathsf{rank}\Big(\Big[\mathbf{H}^T, (\mathbf{H}\mathbf{F})^T, \ldots, (\mathbf{H}\underbrace{\mathbf{F}\ldots\mathbf{F}}_{n_x-1 \text{ terms}})^T\Big]^T\Big) = n_x \tag{3}$$

is of full rank.

The variables $\mathbf{w}_k \in \mathbb{R}^{n_x}$ and $\mathbf{v}_k \in \mathbb{R}^{n_z}$ represents the state and measurement noises, respectively. The noises are assumed to be zero-mean random variables with *unknown* but bounded joint CM defined by

$$\mathsf{E}\left[\begin{bmatrix}\mathbf{w}_k \\ \mathbf{v}_k\end{bmatrix}\begin{bmatrix}\mathbf{w}_l^T & \mathbf{v}_l^T\end{bmatrix}\right] = \begin{bmatrix}\mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R}\end{bmatrix}\delta_{kl}, \tag{4}$$

where the operator $\mathsf{E}[\cdot]$ denotes the expectation and δ_{kl} is the Kronecker delta, i.e., $\delta_{kl} = 1$ for $k = l$ and $\delta_{kl} = 0$ for $k \neq l$. Probability density functions of the noises as well as of the initial state are not assumed to be known.

The *noise CMs identification methods* aim to provide estimates of the noise CMs $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ forming the joint CM (4) on the basis of available measurements $\mathbf{z}^\tau = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_\tau]$ and known model matrices $\mathbf{F}$ and $\mathbf{H}$.

3. CORRELATION METHODS: CONCEPT AND PROPERTIES

The correlation methods represent, probably, the most developed class of the noise CMs identification methods providing estimates with well analysed properties and acceptable computational complexity (Duník et al., 2017). The methods are based on a statistical analysis of the measurement vector prediction error (MPE) generally defined and computed as

$$\mathbf{e}_k = \mathcal{Z}_k - \hat{\mathcal{Z}}_k \in \mathbb{R}^{n_e}, \tag{5}$$

where $\mathcal{Z}_k$ is a vector formed by the measurements and $\hat{\mathcal{Z}}_k$ is its prediction [3]. The construction of the measurement vector $\mathcal{Z}_k$ and definition of its prediction $\hat{\mathcal{Z}}_k$ are the key differences between particular correlation methods.

Regardless of the considered correlation method, the MPE $\mathbf{e}_k$ (5) is set to be a wide-sense stationary zero-mean random process, thus, with the following property (Duník et al., 2017):

Assumption 2: Considering $k \to \infty$, the autocovariance function in the vector form

$$\mathbf{C}_k = \mathsf{E}\Big[(\mathbf{e}_k \otimes \mathbf{e}_k)^T, (\mathbf{e}_k \otimes \mathbf{e}_{k-1})^T, \ldots, (\mathbf{e}_k \otimes \mathbf{e}_{k-M})^T\Big]^T, \tag{6}$$

where $\mathbf{C}_k \in \mathbb{R}^{n_C}$, $n_C = (M+1)n_e^2$, M denotes the autocorrelation function lag, and the symbol $\otimes$ the Kronecker product, converges to the steady-state function $\mathbf{C}$, i.e.,

$$\mathbf{C}_k \to \mathbf{C} \text{ as } k \to \infty. \tag{7}$$

The autocorrelation function $\mathbf{C}$ (7) is unknown, but it can be *estimated* from the measured data by an unbiased estimator

$$\hat{\mathbf{C}} = \tfrac{1}{\tau - M + 1}\sum_{k=M}^{\tau}\Big[(\mathbf{e}_k \otimes \mathbf{e}_k)^T, \ldots, (\mathbf{e}_k \otimes \mathbf{e}_{k-M})^T\Big]^T. \tag{8}$$

Considering the estimator (8), the estimate $\hat{\mathbf{C}}$ goes to the true value $\mathbf{C}$, i.e.,

$$\widehat{\mathbf{C}} \to \mathbf{C} \text{ as } \tau \to \infty. \tag{9}$$

In the following, the limit case is considered, and the true and estimated autocovariance functions are not further strictly distinguished [4].

3.1 Concept

As the name of the correlation methods suggests, their basic idea stems from analysis of the MPE autocorrelation function (7), which can be written as a *linear* function of the sought noise CMs

$$\mathbf{C} = \mathbf{A}\boldsymbol{\theta} \in \mathbb{R}^{n_C}, \tag{10}$$

where the vector of the unique noise CMs elements is defined as

$$\boldsymbol{\theta} = \left[\mathbf{Q}_{\mathrm{U}}^T, \mathbf{R}_{\mathrm{U}}^T, \mathbf{S}_{\mathrm{U}}^T\right]^T, \tag{11}$$

$\mathbf{A} \in \mathbb{R}^{n_C \times n_\theta}$ is the known matrix constructed from the known model matrices $\mathbf{F}$, $\mathbf{H}$, and the notation $\mathbf{Q}_{\mathrm{U}}$ denoted the *unique* elements of $\mathbf{Q}$ stacked into the vector.

Generally, the matrix $\mathbf{A}$ can be rank deficient [5] and, thus, some parameters of $\boldsymbol{\theta}$ (11) have to be correctly specified [6] by the user to find the unbiased estimate of the remaining parameters. Then, (10) reads

$$\mathbf{C} - \mathbf{A}_{\mathrm{S}}\boldsymbol{\theta}_{\mathrm{S}} = \mathbf{A}_{\mathrm{E}}\boldsymbol{\theta}_{\mathrm{E}}, \tag{12}$$

with $\mathbf{A} = [\mathbf{A}_{\mathrm{E}} \ \mathbf{A}_{\mathrm{S}}]$ and $\boldsymbol{\theta} = \left[\boldsymbol{\theta}_{\mathrm{E}}^T \ \boldsymbol{\theta}_{\mathrm{S}}^T\right]^T$, where the matrix $\mathbf{A}_{\mathrm{E}}$ belongs to the *estimated* elements $\boldsymbol{\theta}_{\mathrm{E}}$ and the matrix $\mathbf{A}_{\mathrm{S}}$ belongs to the *specified* elements $\boldsymbol{\theta}_{\mathrm{S}}$. Therefore, the following assumption is considered:

Assumption 3: The noise CM element vectors $\boldsymbol{\theta}_{\mathrm{E}}$, $\boldsymbol{\theta}_{\mathrm{S}}$ are selected so that the matrix $\mathbf{A}_{\mathrm{E}}$ has full rank.

The matrices $\mathbf{A}$ and $\widehat{\mathbf{C}}$ as well as the vector $\boldsymbol{\theta}_{\mathrm{S}}$ are *known* and the the matrix $\mathbf{A}_{\mathrm{E}}$ is of full rank. Then, based on (8), (9), and (12), the sought noise CMs elements $\boldsymbol{\theta}_{\mathrm{E}}$ are estimated using the least-squares (LS) method

$$\hat{\boldsymbol{\theta}}_{\mathrm{E}} = \mathbf{A}_{\mathrm{E}}^{\dagger}(\hat{\mathbf{C}} - \mathbf{A}_{\mathrm{S}}\boldsymbol{\theta}_{\mathrm{S}}), \tag{13}$$

where $\mathbf{A}_{\mathrm{E}}^{\dagger} = (\mathbf{A}_{\mathrm{E}}^T\mathbf{A}_{\mathrm{E}})^{-1}\mathbf{A}_{\mathrm{E}}^T$ is the pseudoinverse of $\mathbf{A}_{\mathrm{E}}$.

This general concept is common for all correlation methods and for *correctly* specified $\boldsymbol{\theta}_{\mathrm{S}}$, all the methods provide

[3] The measurement prediction is typically not optimal in the mean square error (MSE), but it is unbiased or asymptotically unbiased.

[4] Number of measurements used for the estimate $\hat{\mathbf{C}}$ (8) affects the estimate variance only, but not the noise CMs identifiability.

[5] Matrix $\mathbf{J} \in \mathbb{R}^{n \times m}$ is of full rank, if $\mathsf{rank}(\mathbf{J}) = \mathsf{min}(n, m)$. Otherwise, it is rank deficient.

[6] Certain noise CMs parameters may be also known. For example, if it is known that the noises are uncorrelated, then $\mathbf{S}_{\mathrm{U}}$ is a zero vector.

estimate $\hat{\boldsymbol{\theta}}_{\mathrm{E}}$ (13) converging to the true parameters $\boldsymbol{\theta}_{\mathrm{E}}$ as $\tau \to \infty$. The key designer decision is, therefore, related to the *specification of the identifiable set* of the noise CM elements $\boldsymbol{\theta}_{\mathrm{E}}$ so that Assumption 3 is satisfied.

3.2 Identifiability of Noise CMs Elements and Paper Goal

Determination of the number of identifiable elements of the noise CMs, i.e., the identifiability, has been examined since the development of the pioneering methods and, in particular, identifiability of *all* or *unique* noise CMs elements has been investigated.

Identifiability of *all* elements of $\mathbf{Q}$ and $\mathbf{R}$ was discussed in (Mehra, 1970). It was observed that, assuming invertible dynamic matrix $\mathbf{F}$, all elements of the measurement noise CM $\mathbf{R}$ can be identified and at least $n_x \times n_z$ elements of the state noise CM $\mathbf{Q}$ can be found [7]. However, as a consequence of estimation of *all* noise CMs parameters, the noise CMs estimates need not be necessarily symmetric. Unfortunately, once the noise CMs estimation task is reformulated for the unique elements estimation, these identifiability rules are not valid anymore [8].

Identifiability of *unique* elements of the matrices $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ was treated in (Arnold and Rawlings, 2018). It was shown that

- All unique elements of $\mathbf{Q}$ and $\mathbf{R}$ can be found if the dynamic matrix $\mathbf{F}$ is invertible and measurement matrix $\mathbf{H}$ is of full (column) rank,
- It is *not* possible to identify all unique elements of $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ regardless of assumptions placed on the system matrices.

Unfortunately, the key question "*How many unique noise CMs elements can be unbiasedly estimated for a given model?*" has not been answered yet, although the answer is essential for *efficient* design and application of the correlation methods.

The goal of the paper is to *analyse parametric identifiability* of the correlation methods for the LTI model (1), (2) in a unified framework and to *find explicit relations* providing the number of identifiable elements of the noise CMs as a function of the model properties [9].

4. PROPOSED ASSESSMENT OF PARAMETRIC IDENTIFIABILITY

The proposed solution to the identified problem is to analyse the matrix $\mathbf{A}$ in (10) and to determine its (column) rank as a function of the model properties. However, as the matrix $\mathbf{A}$ is a complex nonlinear function of the model matrices and has different form for each correlation method, the general analytical solution for a class of models defined by (1), (2) has not been found yet and it is *unlikely* that a general analytical solution exists. Therefore, the following alternative solution is proposed;

(i) Selection of characteristic correlation methods and construction of the matrix $\mathbf{A}$ for selected methods,
(ii) Definition of a set of models varying in properties and dimension,
(iii) Computation of the matrix $\mathbf{A}_{\mathrm{E}}$ and its maximal rank for all correlations methods and models,
(iv) Determining explicit relations providing the number of identifiable elements of the noise CMs as a function of the model properties.

4.1 Selected Correlation Methods

Five *characteristic* and conceptually different correlation methods, designed in (Mehra, 1970; Bélanger, 1974; Lee, 1980; Bundick, 1988; Duník et al., 2018), are selected;

Method 1, developed in (Mehra, 1970), is historically the first method introducing the concept of the correlation methods for LTI models. The method requires invertible dynamics matrix $\mathbf{F}$ and design of a non-optimal in the MSE and stable linear predictor. The method estimates $\mathbf{Q}$, $\mathbf{R}$ elements.

Method 2, introduced in (Bélanger, 1974), relaxes the assumption of the invertible dynamic matrix $\mathbf{F}$. It is the first method allowing the noise CMs identification for the LTV models [10]. The method estimates $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}$ elements.

Method 3, given in (Lee, 1980), is designed for the LTI models and computes the prediction directly from the measurements without explicit estimation of the state vector. Thus there is no need for a linear predictor design and initialisation. The method estimates $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}$ elements.

Method 4, designed in (Bundick, 1988), adopts the linear predictor based concept. However, compared to Method 1, this method is based on a set of linear predictors, each with a different gain. The method estimates $\mathbf{Q}$, $\mathbf{R}$ elements.

Method 5, developed in (Duník et al., 2018; Kost et al., 2018), generalises the concept of Method 3 and it is extended for the LTV models.

Note 1: Remaining correlation methods can be understood as modifications of the selected methods.

4.2 Set of Models

To analyse the parametric identifiability, a wide set of models of the form (1), (2) is considered. The models differ in properties of matrices $\mathbf{F}$ and $\mathbf{H}$ (namely, full rank vs. rank deficient, stable vs. unstable) and in the dimensions of the state n_x and measurement n_z. Because of the space constraints, the following five exemplary models were selected, out of full set of models, for this paper with state and measurement vector dimensions $n_x = 1, 2, 3, 4$ and $n_z = 1, 2, 3, 4$.

Model 1: Matrix $\mathbf{F}$ is in the Frobenius form with the poles equal to $[0.5, 0.4, \ldots, 0.5-0.1(n_x-1)]$ and the matrix

$$\mathbf{H} = \begin{cases} [\mathbf{1}_{n_z\times 1}, \mathbf{I}_{n_z\times(n_x-1)}], & \text{if } n_x \le n_z, \\ [\mathbf{0}_{n_z\times 1}, \mathbf{I}_{n_z\times(n_x-1)}], & \text{otherwise.} \end{cases} \tag{14}$$

Model 2: Matrix $\mathbf{F}$ is the same as in Model 1, and $\mathbf{H} = \left[\mathbf{1}_{n_z\times 1},\ \mathbf{0}_{n_z\times(n_x-1)}\right]$.

[7] This observation is a motivation for direct estimation of the Kalman filter gain, which has $n_x \times n_z$ elements (Mehra, 1970; Carew and Bélanger, 1973).
[8] Identifiability of all and unique elements is illustrated in Section V.
[9] The model properties include the dimensions of the state and measurement vectors and ranks of the model matrices $\mathbf{F}$ and $\mathbf{H}$.
[10] This method modified for the LTI model is identical with the autocovariance least-squares (ALS) method (Odelson et al., 2006).

Model 3: Matrices $\mathbf{F}$ and $\mathbf{H}$ are constructed as in Model 1, but the first row of $\mathbf{F}$ is zeroed.

Model 4: Matrix $\mathbf{F}$ is the same as in Model 3 and matrix $\mathbf{H}$ is constructed as in Model 1, but both the last two rows and columns are identical (if $n_z > 1$ and $n_x > 1$).

Model 5: Matrix $\mathbf{F}$ is in the Frobenius form with the poles equal to $[1.1, 1, \ldots, 1.1 - 0.1(n_x - 1)]$ and matrix $\mathbf{H}$ is defined as a sum of (14) and $\mathbf{1}_{n_z \times n_x}$.

Notation $\mathbf{I}_{n\times m}$ stands for the $n \times m$ matrix with the element in i-th row and j-th column given by the Kronecker delta $\delta_{i,j}$ and $\mathbf{1}_{n\times m}$ stands for the $n \times m$ matrix of ones.

According to Assumption 1, all models are observable. Model 1 has full rank (and thus invertible) and stable matrix $\mathbf{F}$ and full rank $\mathbf{H}$. Model 2 has full rank matrix $\mathbf{F}$ only, matrix $\mathbf{H}$ is of rank 1. Model 3 has full rank matrix $\mathbf{H}$ only, matrix $\mathbf{F}$ has rank $n_x - 1$ (thus, not invertible). Model 4 has both matrices rank deficient, i.e., $\mathbf{F}$ has rank $n_x - 1$ and $\mathbf{H}$ has rank $\mathsf{min}(n_x, n_z) - 1$. Finally, Model 5 has full rank, but unstable, matrix $\mathbf{F}$ and full rank $\mathbf{H}$.

Note 2: Although the selected models are rather "academic", they are defined to cover a wide range of model types. The results for the wide set of models (including realistic models available in literature) are consistent with the presented results for exemplary models.

4.3 Proposed Definition and Evaluation of Identifiability

The noise CMs parameter identifiability analysis starts with the computation of the rank of the matrix $\mathbf{A}$ in (10) defining the *maximum number of identifiable unique noise CMs elements* further denoted as

$$r = \mathsf{rank}(\mathbf{A}). \tag{15}$$

If $r < n_\theta$, it is necessary to split $\boldsymbol{\theta}$ into two parts as in (12); namely into the estimated part $\boldsymbol{\theta}_\mathrm{E} \in \mathbb{R}^{n_{\theta\mathrm{E}}}$ with (maximally) $n_{\theta\mathrm{E}} = r$ and the user-defined selected part [11] $\boldsymbol{\theta}_\mathrm{S} \in \mathbb{R}^{n_{\theta\mathrm{S}}}$ with $n_{\theta\mathrm{S}} = n_\theta - n_{\theta\mathrm{E}}$. Analogously, the known matrix $\mathbf{A}$ is split into two corresponding parts $\mathbf{A}_\mathrm{E}$ and $\mathbf{A}_\mathrm{S}$. The matrix $\mathbf{A}_\mathrm{E}$ is composed as a combination of $n_{\theta\mathrm{E}}$ columns of the matrix $\mathbf{A}$, where no column occurs more than once. There are c realisable matrices, where c is the binomial coefficient

$$c = \binom{n_\theta}{n_{\theta\mathrm{E}}}, \tag{16}$$

and a realisable matrix is given by $\mathbf{A}^i_\mathrm{E} = [\mathbf{A}_{*,i1}, \mathbf{A}_{*,i2}, \ldots, \mathbf{A}_{*,ir}]$ with $i = [i1, i2, \ldots, ir] \in \mathbb{N}^r$ being a combination of r integers of the set $\{1, 2, \ldots, n_\theta\}$. The symbol $\mathbf{A}_{*,i}$ denotes the i-th column of $\mathbf{A}$. The set of the realisable matrices with cardinality c can be formalised as follows

$$\mathcal{A} = \{\mathbf{A}^i_\mathrm{E} | i \in \mathcal{I}\} \tag{17}$$

where $\mathcal{I}$ represents all possible r-combinations of n_θ.

Unfortunately, as illustrated later, the rank of the matrix $\mathbf{A}^i_\mathrm{E}$, which varies over $\mathcal{A}$, *depends* on the selected columns combination. Naturally, maximal rank of matrices $\mathbf{A}^i_\mathrm{E}$ in $\mathcal{A}$ is r, but it can be lower. Therefore, besides r, the *minimum number of identifiable unique elements* is analysed as well, which is defined as

$$m = \mathsf{min}_\mathcal{I}\mathsf{rank}(\mathbf{A}^i_\mathrm{E}). \tag{18}$$

[11] To solve (13), $\boldsymbol{\theta}_\mathrm{S}$ need to be specified by the user.

Note 3: In the view of (18), the maximum number of identifiable elements (15) can be written as $r = \mathsf{max}_\mathcal{I}\mathsf{rank}(\mathbf{A}^i_\mathrm{E})$.

Note 4: As $\mathbf{A}$ is function of the known model matrices, the identifiability analysis does not depend on the measurements.

5. SOLUTION TO PARAMETRIC IDENTIFIABILITY AND RESULTS

The results are split into two parts; (i) identification of $\mathbf{Q}$, $\mathbf{R}$ unique elements, (ii) identification of $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}$ unique elements. For the first part, all five correlation methods were evaluated for all models and all combinations of dimensions n_x and n_z, and the maximum and minimum numbers of identifiable elements r (15) and m (18), respectively, were computed. Identification of $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}$ unique elements was analogous, with the only exception that Method 1 and Method 4 were not evaluated as they are not capable of the matrix $\mathbf{S}$ identification.

5.1 Maximum/Minimum Number of Identifiable Elements

First observation: All considered correlation methods are *identical* in terms of maximum and minimum numbers of unique identifiable elements of the noise CMs. Therefore, the number of identifiable elements is summarised for all correlation methods in single tables.

The numbers are summarised in Tables 1 and 2 for $\mathbf{Q}$, $\mathbf{R}$ identification and $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}$ identification, respectively. It can be seen, that the minimum number of identifiable elements depends on the selected vector $\boldsymbol{\theta}_\mathrm{E}$ of the noise CMs elements to be estimated. The dependence is significant especially for models with $n_x > n_z$, where it is possible to observe a substantial difference between r and m. Also, the dependence is much stronger in $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}$ identification. For the sake of completeness, the total number of unique elements of the noise CMs is summarised in Table 3.

Table 1. $\mathbf{Q}$, $\mathbf{R}$ identification: r / m values

	n_z \ n_x	1	2	3	4
Model 1	1	2 / 2	3 / 2	4 / 2	5 / 2
	2	4 / 4	6 / 6	8 / 7	10 / 7
	3	7 / 7	9 / 9	12 / 12	15 / 14
	4	11 / 11	13 / 13	16 / 16	20 / 20
Model 2	1	2 / 2	3 / 2	4 / 3	5 / 3
	2	4 / 4	5 / 4	6 / 3	7 / 4
	3	7 / 7	8 / 7	9 / 6	10 / 4
	4	11 / 11	12 / 11	13 / 10	14 / 8
Model 3	1	1 / 1	2 / 1	3 / 2	4 / 2
	2	3 / 3	5 / 5	7 / 5	9 / 5
	3	6 / 6	8 / 8	11 / 10	14 / 12
	4	10 / 10	12 / 12	15 / 14	19 / 18
Model 4	1	1 / 1	2 / 1	3 / 2	4 / 2
	2	3 / 3	4 / 3	5 / 2	6 / 3
	3	6 / 6	7 / 6	10 / 9	12 / 8
	4	10 / 10	11 / 10	14 / 13	18 / 16
Model 5	1	2 / 2	3 / 2	4 / 3	5 / 4
	2	4 / 4	6 / 6	8 / 7	10 / 7
	3	7 / 7	9 / 9	12 / 12	15 / 14
	4	11 / 11	13 / 13	16 / 16	20 / 20

Table 2. **Q**, **R**, **S** identification: r / m values

	n_z \ n_x	1	2	3	4
Model 1	1	2 / 2	3 / 2	4 / 2	5 / 2
	2	5 / 5	7 / 6	9 / 6	11 / 5
	3	9 / 9	12 / 10	15 / 12	18 / 11
	4	14 / 14	18 / 16	22 / 18	26 / 21
Model 2	1	2 / 2	3 / 2	4 / 3	5 / 3
	2	5 / 5	7 / 5	9 / 6	11 / 5
	3	9 / 9	12 / 10	15 / 10	18 / 11
	4	14 / 14	18 / 16	22 / 17	26 / 17
Model 3	1	2 / 1	3 / 2	4 / 2	5 / 2
	2	5 / 4	7 / 5	9 / 5	11 / 5
	3	9 / 8	12 / 9	15 / 11	18 / 11
	4	14 / 13	18 / 15	22 / 17	26 / 19
Model 4	1	2 / 1	3 / 2	4 / 2	5 / 2
	2	5 / 4	7 / 5	9 / 5	11 / 5
	3	9 / 8	12 / 9	15 / 10	18 / 10
	4	14 / 13	18 / 15	22 / 17	26 / 19
Model 5	1	2 / 2	3 / 2	4 / 3	5 / 4
	2	5 / 5	7 / 6	9 / 7	11 / 7
	3	9 / 9	12 / 10	15 / 12	18 / 14
	4	14 / 14	18 / 16	22 / 18	26 / 21

Table 3. Total number of unique elements for **Q**, **R** / **Q**, **R**, **S** Identification

n_z \ n_x	1	2	3	4
1	2 / 3	4 / 6	7 / 10	11 / 15
2	4 / 6	6 / 10	9/ 15	13 / 21
3	7 / 10	9/ 15	13 / 21	16 / 28
4	11 / 15	13 / 21	16 / 28	20 / 36

5.2 *Explicit Relations for Number of Identifiable Parameters*

Based on the analysis of the observed identifiability numbers given in Table 1, the *maximum* number of identifiable *unique* elements r for **Q**, **R** identification is the following function of the model properties

$$r = r_H n_x - \tfrac{r_H(r_H-1)}{2} + u_R - \tfrac{(n_x-r_F)(n_x-r_F+1)}{2}, \quad (19)$$

where $u_R = \frac{n_z(n_z+1)}{2}$ denotes the number of the unique elements of **R** and r_F and r_H denote ranks of the matrices **F** and **H**, respectively. If the matrices **F** and **H** are of full rank and $n_x > n_z$ (i.e., $r_H = n_z$), then the relation (19) simplifies to

$$r = u_Q - \tfrac{(n_x-n_z)(n_x-n_z+1)}{2} + u_R = n_z(n_x+1). \quad (20)$$

where $u_Q = \frac{n_x(n_x+1)}{2}$ is the number of the unique elements of **Q**. If $n_x \leq n_z$ (i.e., $r_H = n_x$), then (19) becomes

$$r = u_Q + u_R. \quad (21)$$

The *maximum* number of identifiable *unique* elements r for the **Q**, **R**, **S** identification is, w.r.t. Table 2, given by

$$r = n_z n_x + u_R. \quad (22)$$

Second observation: The maximum number r defined in (15) depends on the model properties (including state and measurement dimensions and rank of the model matrices). The maximum number is *independent* of the selection of $\boldsymbol{\theta}_{\mathrm{E}}$ elements. Thus, the maximum number is easily predictable. The selection of $\boldsymbol{\theta}_{\mathrm{E}}$, however, affects the minimum number m (18) in addition to the model properties. Thus, explicit relations for m can hardly be found. This is illustrated the following subsection.

5.3 *Dependence on Parameters Selection*

Let Model 1 be supposed with $n_x = 2$ and $n_z = 1$ and let all methods be selected. Then, the vector of unique unknown elements of **Q**, **R** can be defined as

$$\boldsymbol{\theta} = [Q_{11}, Q_{12}, Q_{22}, R]^T \quad (23)$$

with dimension $n_\theta = 4$, where the notation Q_{ij} stands for the element of the matrix **Q** in i-th row and j-th column. The matrix **A** in (10) has, however, rank

$$r = \mathsf{rank}(\mathbf{A}) = 3 \quad (24)$$

only, which is also the maximum number of identifiable parameters r. It means that at least $n_{\theta \mathrm{S}} = n_\theta - r = 1$ parameter must be specified, which results in $c = 4$ (16) possible combinations of identified noise CMs elements gathered in $\boldsymbol{\theta}_{\mathrm{E}}$ and realisable matrices $\mathbf{A}_{\mathrm{E}}$ in (12). Let three identification cases be considered;

- Estimated noise CMs elements are $\boldsymbol{\theta}_{\mathrm{E}} = [Q_{11}, Q_{12}, R]^T$ (thus, the remaining elements in $\boldsymbol{\theta}_{\mathrm{S}} = Q_{22}$ need to be specified by the user). Then, $\mathsf{rank}(\mathbf{A}_{\mathrm{E}}) = 3 = r$, i.e., in this case, the *maximum* number of elements *can* be estimated.
- Estimated noise CMs elements are $\boldsymbol{\theta}_{\mathrm{E}} = [Q_{11}, Q_{22}, R]^T$ (thus, the remaining elements in $\boldsymbol{\theta}_{\mathrm{S}} = Q_{12}$ need to be specified by the user). Then, $\mathsf{rank}(\mathbf{A}_{\mathrm{E}}) = 2 \neq r$, i.e., in this case, the maximum number of elements *cannot* be estimated and the user has to reduce $\boldsymbol{\theta}_{\mathrm{E}}$ further.
- Estimated elements are $\boldsymbol{\theta}_{\mathrm{E}} = [Q_{11}, Q_{12}, Q_{22}]^T$, remaining is $\boldsymbol{\theta}_{\mathrm{S}} = R$. Then, $\mathsf{rank}(\mathbf{A}_{\mathrm{E}}) = 2 \neq r$ and the user has to reduce $\boldsymbol{\theta}_{\mathrm{E}}$ further.

Third observation: It *is not* always possible to estimate all diagonal elements of the state noise CM **Q**.

5.4 *Relation of All and Unique Elements Identifiability*

As mentioned in Section III.C, several identification methods have been designed and analysed under assumption of identification of all noise CMs elements Mehra (1970); Bélanger (1974); Lee (1980). This paper stresses identification of the unique elements. The goal of this section is to show the principal difference between identification of all elements and unique elements.

Given Model 1 with $n_x = 3$ and $n_z = 2$, let all methods be selected. Then, the vector of *all* unknown elements of **Q**, **R** can be defined as

$$\boldsymbol{\theta}_{\mathrm{ALL}} = [\mathbf{Q}_{*,1}^T, \mathbf{Q}_{*,2}^T, \mathbf{Q}_{*,3}^T, \mathbf{R}_{*,1}^T, \mathbf{R}_{*,2}^T]^T \quad (25)$$

with dimension $n_{\theta,\mathrm{ALL}} = n_x^2 + n_z^2 = 13$. The respective matrix $\mathbf{A}_{\mathrm{ALL}}$ in (10) for particular methods has rank

$$\mathsf{rank}(\mathbf{A}_{\mathrm{ALL,Method\ i}}) = 12, i = \{1, 3, 4, 5\} \quad (26)$$

$$\mathsf{rank}(\mathbf{A}_{\mathrm{ALL,Method\ 2}}) = 10, \quad (27)$$

which indicates that not all elements of **Q**, **R** can be estimated, but it may be possible to estimate unique elements by certain methods (in this case, the noise CMs has $n_{\theta,\mathrm{UNIQUE}} = u_Q + u_R = 9$ unique elements). However, if the vector of *unique* elements of **Q**, **R** is defined as

$$\boldsymbol{\theta}_{\mathrm{UNIQUE}} = [Q_{11}, Q_{12}, Q_{22}, \mathbf{Q}_{*,3}^T, R_{11}, \mathbf{R}_{*,2}^T]^T \quad (28)$$

with dimension $n_{\theta,\mathrm{UNIQUE}} = 9$, then the respective matrix $\mathbf{A}_{\mathrm{UNIQUE}}$ in (10) has the rank

$$\mathsf{rank}(\mathbf{A}_{\mathrm{UNIQUE}}) = 8 \quad (29)$$

for all methods, which is less than the number of sought unique elements $n_{\theta,\text{UNIQUE}}$.

Fourth observation: Known number of identifiable elements of *all* noise CMs elements *does not* help in determination of the number of identifiable *unique* elements.

5.5 Specification of Parameters

Usually, if $n_x > n_z$, it is possible to estimate only a subset of unique noise CMs elements. Then, it is necessary to specify the elements $\boldsymbol{\theta}_\text{S}$ and to estimate the remaining $\boldsymbol{\theta}_\text{E}$.

Let the Model 2 be considered with $n_x = 2$, $n_z = 1$, matrices $\mathbf{S}, R$ are known, and only the elements of the state noise covariance matrix $\mathbf{Q} = \left[\begin{smallmatrix} 1 & 1.8 \\ 1.8 & 4 \end{smallmatrix}\right]$ are unknown. In this case, it is possible to estimate only 2 elements out of three unknowns. If the diagonal elements are decided to be identified, i.e., $\boldsymbol{\theta}_\text{E} = [Q_{11}, Q_{22}]$, the off-diagonal element has to be specified $\boldsymbol{\theta}_\text{S} = Q_{12}$. However, e.g. specification of $\boldsymbol{\theta}_\text{S} = Q_{12} = 0$ leads to the noise CM estimate $\widehat{\mathbf{Q}} = \left[\begin{smallmatrix} -1 & 0 \\ 0 & 4.38 \end{smallmatrix}\right]$, which is *not* a positive semi-definite matrix, thus, it is not a covariance matrix.

Fifth observation: The elements in $\boldsymbol{\theta}_\text{S}$ have to be selected with caution because the identified CMs may not be positive semi-definite even for infinitely many measurements. Similar conclusion holds also in case of unreasonable specification of $\mathbf{S}$ and estimation of $\mathbf{Q}$ and $\mathbf{R}$.

5.6 Notes

In this section, several notes on properties observed in the noise CMs identifiability analysis are given.

Note 5: If $\mathbf{F}$ is of full rank, conditions for the identifiability of the unique elements of $\mathbf{R}$ are independent of identifiability conditions of $\mathbf{Q}$ and $\mathbf{S}$. The identifiability condition of $\mathbf{Q}$ and $\mathbf{S}$ are, however, closely related. It means, it is always possible to identify all u_R unique elements of $\mathbf{R}$ and only, the maximum number $r - u_R$ of $\mathbf{Q}$ or $\mathbf{S}$ elements.

Note 6: Any solution $\boldsymbol{\theta}$ fulfilling (10) leading to the noise CMs estimates $\hat{\mathbf{Q}}, \hat{\mathbf{R}}, \hat{\mathbf{S}}$, for which a steady-state solution $\hat{\mathbf{P}}$ of the discrete-time algebraic equation

$$\hat{\mathbf{P}} = \mathbf{F}\hat{\mathbf{P}}\mathbf{F}^T - \hat{\mathbf{L}}(\mathbf{H}\hat{\mathbf{P}}\mathbf{H}^T + \hat{\mathbf{R}})^{-1}\hat{\mathbf{L}}^T + \hat{\mathbf{Q}} \quad (30)$$

with $\hat{\mathbf{L}} = \mathbf{F}\hat{\mathbf{P}}\mathbf{H}^T + \hat{\mathbf{S}}$ exists, results in the linear filter gain

$$\hat{\mathbf{K}} = (\mathbf{F}\hat{\mathbf{P}}\mathbf{H}^T + \hat{\mathbf{S}})(\mathbf{H}\hat{\mathbf{P}}\mathbf{H}^T + \hat{\mathbf{R}})^{-1}, \quad (31)$$

that is identical with the steady-state Kalman filter gain $\mathbf{K}$ computed on the basis of true noise CMs $\mathbf{Q}, \mathbf{R}, \mathbf{S}$, i.e.,

$$\hat{\mathbf{K}} = \mathbf{K}. \quad (32)$$

The equality (32) holds even if the noise CMs estimates are biased because of unsuitable choice of $\boldsymbol{\theta}_\text{S}$ for models with $n_x > n_z$. This conclusion is consistent with the observation that the Kalman gain $\mathbf{K} \in \mathbb{R}^{n_x \times n_z}$ can always be estimated (contrary to the noise CMs) (Mehra, 1970; Carew and Bélanger, 1973; Arnold and Rawlings, 2018).

Note 7: Assume identification of $\mathbf{Q}, \mathbf{R}$ and the presence of the known state noise shaping matrix $\mathbf{G} \in \mathbb{R}^{n_x \times n_{\tilde{w}}}$, $n_x \geq n_{\tilde{w}}$ defining the state noise

$$\mathbf{w}_k = \mathbf{G}\tilde{\mathbf{w}}_k \quad (33)$$

with $\mathsf{E}\left[\tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_l^T\right] = \tilde{\mathbf{Q}}\delta_{kl} \in \mathbb{R}^{n_{\tilde{w}} \times n_{\tilde{w}}}$. Then the maximum rank of $\mathbf{A}$ is $\min(r, u_{\tilde{\mathbf{Q}}} + u_r)$.

Note 8: Presence of the known input signal in the state-space model (1), (2) does not have any impact on the method parameter identifiability.

Note 9: Our simulations indicate that other classes of the noise CMs estimation methods (i.e., Bayesian and maximum likelihood) suffer from the same identifiability constraints.

6. CONCLUDING REMARKS

This paper dealt with the identification of the noise covariance matrices of the linear-time invariant state-space model with the stress on the correlation methods. Explicit relations for the maximum number of the identifiable unique noise CMs elements were derived as functions of model properties. It was shown, that the maximum identifiability number is the same for all correlation methods. Such relations give an idea about capability of the correlation methods to identify the CMs elements for a given model. Thus, derived relations may significantly simplify design and application of the correlation methods.

REFERENCES

Arnold, T.J. and Rawlings, J.B. (2018). Uniqueness conditions for ALS problems. In *6th IFAC Conf. on Nonlinear Model Predictive Control*. Madison, WI, USA.

Bélanger, P.R. (1974). Estimation of noise covariance matrices for a linear time-varying stochastic process. *Automatica*, 10(3), 267–275.

Bundick, W. (1988). Development of a technique for estimating noise covariances using multiple observers. Technical report, NASA Technical Memorandum.

Carew, B. and Bélanger, P.R. (1973). Identification of optimum steady-state gain for systems with unknown noise covariances. *IEEE Transactions on Automatic Control*, 18(6), 582–587.

Duník, J., Kost, O., and Straka, O. (2018). Design of measurement difference autocovariance method for estimation of process and measurement noise covariances. *Automatica*, 90, 16–24.

Duník, J., Kost, O., Straka, O., and Blasch, E. (2020). Covariance estimation and gaussianity assessment for state and measurement noise. *Journal of Guidance, Control, and Dynamics*, 43(1), 132–139.

Duník, J., Straka, O., Kost, O., and Havlík, J. (2017). Noise covariance matrices in state-space models: A survey and comparison - part I. *Inter. Journal of Adaptive Control and Signal Processing*, 31(11), 1505–1543.

Kost, O., Duník, J., and Straka, O. (2018). Correlated noise characteristics estimation for linear time-varying systems. In *57th Conference on Decision and Control*. Miami, FL, USA.

Lee, T. (1980). A direct approach to identify the noise covariances of Kalman filtering. *IEEE Transactions on Automatic Control*, 25(4), 841–842.

Mehra, R.K. (1970). On the identification of variances and adaptive filtering. *IEEE Transactions on Automatic Control*, 15(2), 175–184.

Odelson, B.J., Rajamani, M.R., and Rawlings, J.B. (2006). A new autocovariance least-squares method for estimating noise covariances. *Automatica*, 42(2), 303–308.