18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Estimation of State and Measurement Noise Characteristics

J. Duník, O. Straka, M. Šimandl, O. Kost, J. Ajgl
Department of Cybernetics &
European Centre of Excellence NTIS
Faculty of Applied Sciences, University of West Bohemia
Univerzitní 8, 306 14 Pilsen, Czech Republic
Emails: {dunikj, straka30, simandl, kost, jiriajgl}@kky.zcu.cz

M. Soták, R. Baránek, Z. Kaňa
Advanced Technology Europe
Honeywell International
Tuřanka 100, 627 00 Brno, Czech Republic
Emails: {milos.sotak, radek.baranek,
zdenek.kana}@honeywell.com

***Abstract*—The paper deals with the estimation of the noise characteristics of a linear system described by the state space model. In particular, the stress is laid on the linear systems with the measurement noise correlated in time. The method, providing unbiased estimates of the state and measurement noise characteristics is analytically derived. The method extends, in principle, the concept of the correlation methods by explicit consideration of the noise time-correlation. The method is also modified into the form suitable for a sensor calibration procedure typically realized by the Allan variance method. The theoretical results are thoroughly discussed and the proposed method is validated using numerical examples.**

Keywords: State estimation; Estimation theory; Identification; Sensor calibration; Kalman filtering.

I. Introduction

The state estimation, signal processing, and control design methods require not only the complete knowledge of the functions in the system description but also of statistics of the noises affecting the system. The assumption of the known noise statistics is, however, questionable in many cases. Incorrect description of the noise statistics may cause significant worsening of estimation or control quality or even divergence of the underlying algorithm output.

From the seventies, various methods for estimation of the state-space model noise covariance matrices have been proposed. The methods can be divided into several categories; correlation methods [1]–[5], Bayesian estimation methods [6], maximum likelihood estimation methods [7], covariance matching methods [8], methods based on the minimax approach [9], subspace methods [10], prediction error methods [11], the Kalman filter working as a parameter estimator [12], or methods tied with variational Bayesian approximation [13]. Besides the noise covariance matrix estimation methods, alternative approaches directly estimating the gain of a linear estimator have been developed as well [1], [14], [15]. A characterisation of the methods with their assumptions, properties, and limitations can be found in e.g., [3], [9], [16], [17].

All the methods have been prevailingly proposed for the linear (often time-invariant) systems under the assumption of the white noises (zero-mean and uncorrelated in time) in the state and measurement equations. Such assumption might, however, be limiting in certain scenarios where sensors with a time-correlated noise description are considered. Typically, the inertial sensors belong into this category.

The goal of the paper is to propose a method providing the estimates of the noise characteristics for the linear systems if the state noise is assumed to be white and the measurement noise to be correlated in time (coloured). The proposed method, extending the family of the correlation methods, is shown to provide the unbiased estimates.

The rest of the paper is organised as follows. In Section 2, the system description is defined and the problem formulation is stated. In Section 3, the method estimating the characteristics of the state and measurement noise is proposed and thoroughly discussed. Within this section, a simplified version of the method suitable for a sensor calibration procedure is given as well. The numerical illustrations are described and evaluated in Section 4. Section 5 then concludes the paper.

II. System Definition and Problem Statement

A discrete-time linear stochastic dynamic system

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{w}_k, k = 0, 1, 2, \ldots, T, \quad (1)$$

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k, k = 0, 1, 2, \ldots, T, \quad (2)$$

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 measurement at time instant k, respectively. The system 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 supposed to be known as well as the characteristics of the initial state of the system $\mathbf{x}_0$ described by the mean $\bar{\mathbf{x}}_0$ and the covariance matrix $\mathbf{P}_0$.

The variables $\mathbf{w}_k \in \mathbb{R}^{n_x}$ and $\mathbf{v}_k \in \mathbb{R}^{n_z}$ are the state and measurement noises. The state noise is assumed to be white with the covariance matrix $\mathbf{Q} \in \mathbb{R}^{n_x \times n_x}$. The measurement noise is modelled as

$$\mathbf{v}_k = \boldsymbol{\nu}_k + \boldsymbol{\beta} + \mathbf{u}_k, \quad (3)$$

where $\mathbf{u}_k$ is a white noise with the covariance matrix $\mathbf{R}_u \in \mathbb{R}^{n_z \times n_z}$, $\boldsymbol{\beta} \in \mathbb{R}^{n_z \times 1}$ is a constant vector, and $\boldsymbol{\nu}_k$ is a

process correlated in time described by the (zero-mean) Gauss-Markov (GM) process of the first order

$$\boldsymbol{\nu}_k = \boldsymbol{\alpha}\boldsymbol{\nu}_{k-1} + \boldsymbol{\xi}_{k-1} \tag{4}$$

with $\boldsymbol{\xi}_{k-1}$ being a (white) driving noise with the covariance matrix $\mathbf{R}_\xi \in \mathbb{R}^{n_z \times n_z}$ and $\boldsymbol{\alpha} \in \mathbb{R}^{n_z \times n_z}$ being a stable constant matrix. The steady-state covariance matrix (in the vector form) of the GM process is

$$(\mathbf{R}_\nu)_{vec} = (\mathbf{I}_{n_z^2} - \boldsymbol{\alpha} \otimes \boldsymbol{\alpha})^{-1}(\mathbf{R}_\xi)_{vec}, \tag{5}$$

where $\mathbf{I}_{n_z^2}$ is the identity matrix of the indicated dimension, the symbol $\otimes$ denotes the Kronecker product, and the notation $(\mathbf{A})_{vec}$ means the columnwise stacking of a matrix $\mathbf{A}$ into a vector [18]. The overall variance of the measurement noise is then

$$\mathbf{R} = cov[\mathbf{v}_k] = \mathbf{R}_\nu + \mathbf{R}_u. \tag{6}$$

The system and measurement noise sequences $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_k\}$ are assumed to be mutually independent processes and independent of the initial state.

The characteristics of the state and measurement noises, i.e., the matrices and vectors $\mathbf{Q}$, $\mathbf{R}_\xi$, $\mathbf{R}_u$, $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$ are assumed to be unknown[1] and thus, to be estimated.

III. Estimation of Noise Characteristics

The proposed method for estimation of the noise characteristics of linear systems with time-correlated noise in the measurement equation extends the concept of the autocovariance least-squares (ALS) correlation method.

The ALS method has been designed for estimation of the noise covariance matrices of the linear systems with white state and measurement noises [3]–[5], [19]. The method is based on the analysis of the innovation sequence of a "non-optimal" linear estimator. The method is briefly introduced in Appendix A.

A. Autocovariance least-squares method for time-correlated measurement noise

In this section, the ALS method respecting the time-correlated measurement noise is designed and analysed. The method is based on analysis of the multistep predictor estimate error.

1) State predictor: The multistep predictor of the state and measurement [5] is defined as

$$\hat{\mathbf{x}}_{k+1} = \mathbf{F}\hat{\mathbf{x}}_k, \forall k, \tag{7}$$

$$\hat{\mathbf{z}}_k = \mathbf{H}\hat{\mathbf{x}}_k, \forall k, \tag{8}$$

with the initial condition $\hat{\mathbf{x}}_0 = \bar{\mathbf{x}}_0$. The state prediction error $\boldsymbol{\varepsilon}_k$ and the measurement prediction error (innovation) $\mathbf{e}_k$ are then equal to

$$\boldsymbol{\varepsilon}_{k+1} = \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1} = \mathbf{F}\boldsymbol{\varepsilon}_k + \mathbf{w}_k, \tag{9}$$

$$\mathbf{e}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k = \mathbf{H}\boldsymbol{\varepsilon}_k + \mathbf{v}_k. \tag{10}$$

Note that the state prediction error is *independent* of the measurement noise in this case.

[1]All the characteristics except of $\boldsymbol{\beta}$ are the covariance matrices. Only the vector $\boldsymbol{\beta}$ is related to the mean of the process (3).

2) State prediction error properties: Assuming the *stable* matrix $\mathbf{F}$, the state prediction error is a (white) random process with zero mean and steady-state covariance matrix given by the solution to the Lyapunov equation

$$\mathbf{P}_\varepsilon = \mathbf{F}\mathbf{P}_\varepsilon\mathbf{F}^T + \mathbf{Q}. \tag{11}$$

Using the Kronecker calculus, the solution (in vector form) can be written as

$$(\mathbf{P}_\varepsilon)_{vec} = (\mathbf{I}_{n_x^2} - \mathbf{F} \otimes \mathbf{F})^{-1}\mathbf{Q}_{vec}. \tag{12}$$

Note that the steady-state matrix $\mathbf{P}_\varepsilon$ is a linear function of the unknown matrix $\mathbf{Q}$.

3) Measurement prediction error properties: With respect to properties of the state prediction (9) and the measurement noise (3), the steady-state mean of the measurement prediction error (10) is

$$E[\mathbf{e}_k] = \boldsymbol{\beta}. \tag{13}$$

The steady-state covariance matrix of the measurement prediction error is, respecting the independence of $\boldsymbol{\varepsilon}_k$ (9) and $\mathbf{w}_k$ (1), equal to

$$\begin{aligned}\mathcal{C}_0 = cov[\mathbf{e}_k] &= E[(\mathbf{e}_k - \boldsymbol{\beta})(\cdot)^T] = \mathbf{H}\mathbf{P}_\varepsilon\mathbf{H}^T + \mathbf{R} \\ &= \mathbf{H}\mathbf{P}_\varepsilon\mathbf{H}^T + \mathbf{R}_\nu + \mathbf{R}_u,\end{aligned} \tag{14}$$

where the notation $E[(\mathbf{a})(\cdot)^T]$ means $E[(\mathbf{a})(\mathbf{a})^T]$. The steady-state cross-covariance matrix with a lag of length j can be computed as

$$\begin{aligned}\mathcal{C}_j = cov[\mathbf{e}_k, \mathbf{e}_{k-j}] &= E[(\mathbf{e}_k - \boldsymbol{\beta})(\mathbf{e}_{k-j} - \boldsymbol{\beta})^T] \\ &= E[(\mathbf{H}\boldsymbol{\varepsilon}_k + \boldsymbol{\nu}_k + \mathbf{u}_k)(\mathbf{H}\boldsymbol{\varepsilon}_{k-j} + \boldsymbol{\nu}_{k-j} + \mathbf{u}_{k-j})^T] \\ &= \mathbf{H}\mathbf{F}^j\mathbf{P}_\varepsilon\mathbf{H}^T + \boldsymbol{\alpha}^j\mathbf{R}_\nu\end{aligned} \tag{15}$$

with $j = 1, \ldots, N-1$ and $\mathbf{R}_\nu$ given by (5).

4) Derivation of ALS method: The proposed ALS method for estimation of the characteristics of the state and measurement noises is based, in fact, on an algorithm with the following two subsequent steps:

i. estimation of the (cross-)covariance matrices of the measurement prediction error sequence, i.e., of $\mathcal{C}_0$ (14) and $\mathcal{C}_j$ (15), on the basis of available measurements,
ii. estimation of the unknown noise characteristics on the basis of the previous step outputs and relations (11)–(15).

In the first step it is thus necessary to compute the measurement prediction error sequence according to (10), i.e., as

$$\mathbf{e}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k, \forall k. \tag{16}$$

Then, having the sequence, the (cross-)covariance matrices (14), (15) can be estimated by

$$\hat{\mathcal{C}}_0 = \tfrac{1}{T-1}\sum_{k=1}^{T}(\mathbf{e}_k - \hat{\boldsymbol{\beta}})(\cdot)^T, \tag{17}$$

$$\hat{\mathcal{C}}_j = \tfrac{1}{T-j-1}\sum_{k=1}^{T}(\mathbf{e}_k - \hat{\boldsymbol{\beta}})(\mathbf{e}_{k-j} - \hat{\boldsymbol{\beta}})^T, \forall j, \tag{18}$$

where

$$\hat{\boldsymbol{\beta}} = \frac{1}{T}\sum_{k=1}^{T}\mathbf{e}_k \tag{19}$$

is directly an estimate of the unknown mean (13) of the measurement prediction error. It is worth noting that the sample based statistics (17)-(19) represent the unbiased estimates of the respective quantities, i.e.,

$$E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta},\ E[\hat{\mathcal{C}}_0] = \mathcal{C}_0,\ E[\hat{\mathcal{C}}_j] = \mathcal{C}_j, \tag{20}$$

which converge to the true values inversely with number of data T. This conclusion can be approached by taking expectation of (17)-(19) respecting the properties of sample-based mean and covariance matrix computation [3].

In the second step, the estimate of the remaining noise characteristics, namely the covariance matrices $\mathbf{Q}$, $\mathbf{R}_\xi$, and $\mathbf{R}_u$, and the noise dynamic matrix $\boldsymbol{\alpha}$, is given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} ||\hat{\mathbf{b}} - \mathcal{G}(\boldsymbol{\theta})||_2, \tag{21}$$

where

- $\boldsymbol{\theta}$ is a vector of all unknown matrices element defined as

$$\boldsymbol{\theta} = [\mathbf{Q}_{vec}^T, (\mathbf{R}_\xi)_{vec}^T, (\mathbf{R}_u)_{vec}^T, \boldsymbol{\alpha}_{vec}^T]^T, \tag{22}$$

- $\mathcal{G}(\boldsymbol{\theta})$ is a known, generally nonlinear[2], function w.r.t. the vector of unknown parameters $\boldsymbol{\theta}$ stemming from (14), (15) and it is defined as

$$\mathcal{G}(\boldsymbol{\theta}) = \left([\mathbf{g}_0(\boldsymbol{\theta})^T, \mathbf{g}_1(\boldsymbol{\theta})^T, \ldots, \mathbf{g}_{N-1}(\boldsymbol{\theta})^T]^T\right)_{vec}, \tag{23}$$

- $\hat{\mathbf{b}}$ is a known vector of dependent variables defined as

$$\hat{\mathbf{b}} = \left(\hat{\mathcal{C}}_e(N)\right)_{vec} \tag{24}$$

with

$$\hat{\mathcal{C}}_e(N) = [\hat{\mathcal{C}}_0^T, \hat{\mathcal{C}}_1^T, \ldots, \hat{\mathcal{C}}_{N-1}^T]^T, \tag{25}$$

and $||\cdot||_2$ denotes the L_2-norm.

Equation (21) cannot be minimised analytically and an iterative minimisation method for the solution to the nonlinear least-squares (NLS) needs to be employed. Among these methods, the Gauss-Newton method and other gradient based methods are often used [20], [21].

It should be also mentioned that with respect to the form of (5) appearing in (15) it might be favourable to estimate by the NLS the matrices $\mathbf{Q}$, $\mathbf{R}_\nu$, $\mathbf{R}_u$, $\boldsymbol{\alpha}$ instead of the matrices $\mathbf{Q}$, $\mathbf{R}_\xi$, $\mathbf{R}_u$, $\boldsymbol{\alpha}$ which were considered heretofore. Direct estimation of $\mathbf{R}_\nu$ instead of $\mathbf{R}_\xi$ (seemingly) reduces the "nonlinearity" of the considered minimisation problem solved by the NLS as the only nonlinearity is given by the product $\boldsymbol{\alpha}^j\mathbf{R}_\nu$ in (15) and no rational function stemming from (5) is involved. Such an approach basically follows the idea of the transformably linear models discussed e.g., in [20], [21].

[2]It should be mentioned that the matrices $\mathbf{Q}$, $\mathbf{R}_u$ appears in (17), (18) in a linear fashion.

5) Algorithm of ALS method (a scalar case): The relations provided in the previous section are rather general. The particular forms to be minimised depend among others on the assumed properties of sought covariance matrices (for example, matrix diagonality, equality of certain elements).

In this section, the proposed ALS method is detailed for a scalar system, i.e., for (1), (2) with $n_x = n_z = 1$ assuming the usage of the Gauss-Newton minimisation technique.

The *algorithm of the ALS method for linear system noise characteristics estimation* can be written in the following steps:

i. Compute measurement prediction error sequence $\{e_k\}_{k=0}^T$ according to (16).
ii. Compute sample-based estimate of the mean β of the sequence according to (19).
iii. Compute sample-based estimates of the variance $\mathcal{C}_0$ and covariances $\mathcal{C}_j$ according to (17) and (18), respectively, for $j = 1, \ldots, N-1$ with $N \geq 4$.
iv. Compute estimates of the *unknown* variances and the parameter gathered in the vector

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4] = [Q, R_\nu, R_u, \alpha]^T, \tag{26}$$

by recursive computation of the following steps of the NLS

$$\hat{\mathbf{b}} = [\hat{\mathcal{C}}_0, \hat{\mathcal{C}}_1, \ldots, \hat{\mathcal{C}}_{N-1}]^T, \tag{27}$$

$$\mathcal{G}(\hat{\boldsymbol{\theta}}_i) = \begin{bmatrix} g_0(\hat{\boldsymbol{\theta}}_i) \\ g_1(\hat{\boldsymbol{\theta}}_i) \\ \vdots \\ g_{N-1}(\hat{\boldsymbol{\theta}}_i) \end{bmatrix} = \begin{bmatrix} \frac{H^2}{1-F^2}\hat{\theta}_{i,1} + \hat{\theta}_{i,2} + \hat{\theta}_{i,3} \\ \frac{H^2F}{1-F^2}\hat{\theta}_{i,1} + \hat{\theta}_{i,4}\hat{\theta}_{i,2} \\ \vdots \\ \frac{H^2F^{N-1}}{1-F^2}\hat{\theta}_{i,1} + \hat{\theta}_{i,4}^{(N-1)}\hat{\theta}_{i,2} \end{bmatrix} \tag{28}$$

$$\delta\hat{\mathbf{b}}_i = \hat{\mathbf{b}} - \mathcal{G}(\hat{\boldsymbol{\theta}}_i), \tag{29}$$

$$\mathbf{G}(\hat{\boldsymbol{\theta}}_i) = \frac{\partial\mathcal{G}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\hat{\boldsymbol{\theta}}_i} = \mathbf{G}_i = \begin{bmatrix} \frac{H^2}{1-F^2} & 1 & 1 & 0 \\ \frac{H^2F}{1-F^2} & \hat{\theta}_{i,4} & 0 & \hat{\theta}_{i,2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{H^2F^{(N-1)}}{1-F^2} & \hat{\theta}_{i,4}^{(N-1)} & 0 & (N-1)\hat{\theta}_{i,2}\hat{\theta}_{i,4}^{(N-2)} \end{bmatrix} \tag{30}$$

$$\delta\hat{\boldsymbol{\theta}}_i = (\mathbf{G}_i^T\mathbf{G}_i)^{-1}\mathbf{G}_i^T\delta\hat{\mathbf{b}}_i \tag{31}$$

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \delta\hat{\boldsymbol{\theta}}_i, \tag{32}$$

where $\mathbf{G}(\hat{\boldsymbol{\theta}}_i)$ is the Jacobian matrix of the function $\mathcal{G}(\cdot)$ evaluated at $\hat{\boldsymbol{\theta}}_i$, $\hat{\boldsymbol{\theta}}_0$ is the user-defined initial condition, and $i = 0, 1, \ldots, T_{NLS}$. The recursion ends once the error-variable $\delta\hat{\boldsymbol{\theta}}_i$ is negligible [20], [21].

B. Allan variance and autocovariance least-squares method

In the previous section, estimation of the noise characteristics of the linear dynamic systems with the time-correlated measurement noise was discussed. In this section, a conceptually simpler formulation is treated, namely estimation of the measurement noise characteristics under

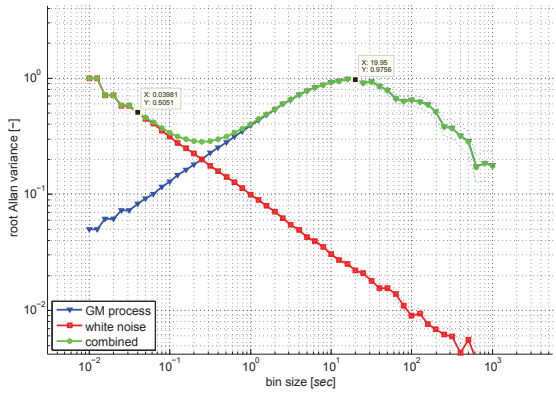

Figure 1. Root Allan variance plot of diffferent processes.

the assumption of the precise knowledge of the state. Such a formulation is typical in the area of a sensor calibration such as calibration of accelerometers or gyroscopes [22]–[25].

The measurement can be, thus, modelled as

$$\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k, k = 0, 1, 2, \ldots, T, \tag{33}$$

where $\mathbf{x}_k$ is assumed to be a *deterministic and known* quantity and $\mathbf{v}_k$ is the measurement noise defined analogously as in (3).

1) Allan variance: Allan variance (AVAR) method is a standard tool for estimation of the parameters of the measurement noise error model (33) which is composed by usually two independent error contributors; one contributor is correlated in time whereas other is white.

Computation of the AVAR from the sampled *scalar* signal (33) with $x_k = 0, \forall k$, follows these steps [26], [27]:

i. The measurement sequence (33) of length T is divided into K bins of length M where $K = \frac{N}{M}$ and $\tau_M = \frac{M}{f_s}$ is a time length of a bin with f_s being a sampling frequency.

ii. For each bin, the sample mean is computed as

$$\bar{z}_k(M) = \frac{1}{M} \sum_{i=1}^{M} z_{(k-1)M+i}. \tag{34}$$

iii. The root Allan variance is computed as a mean of the difference of the means in the subsequent bins, i.e., as

$$\sigma_A^2(\tau_M) = \frac{1}{2(K-1)} \sum_{k-1}^{K-1} \left(\bar{z}_{k+1}(M) - \bar{z}_k(M)\right)^2. \tag{35}$$

iv. Steps i.–iii. of this algorithm are repeated for different bin lengths τ_M.

The output of the AVAR method is the Allan variance plot (in logarithmic scales) showing dependence of the variances $\sigma_A^2(\tau_M)$ (35) on the bin length τ_M.

The AVAR method is typically used for determination of the white and time-correlated noise properties. As illustrated in Fig. 1, each noise type alone produce a characteristic shape of the Allan variance plot:

- A white noise alone produces a line with negative slope of -1/2 when using logarithmic scales on both axes; red line with square markers in Fig. 1. The quantitative characteristic of the white noise (its variance) can be computed using any point coordinates corresponding to the white noise slope according to

$$R_u = \sigma_{WN}^2 = \sigma_A^2(\tau_M)\tau_M f_s \tag{36}$$

  where R_u is the white noise variance, $\sigma_A(\tau_M)$ is root Allan variance at a selected point, τ_M is a bin size length for the selected point, and f_s is sampling frequency.
- A time correlated noise (first order GM process in this case) produces a hump in the Allan variance plot; blue line with triangle markers in Fig. 1. The coordinates of the maximum point of the hump are used to compute the characteristics (correlation time and steady-state variance) of the GM process as

$$\tau_c = \frac{\tau_{\max}}{1.89}, \tag{37}$$

$$R_\nu = \sigma_{steady}^2 = \left(\frac{\sigma_{A\max}}{0.62}\right)^2, \tag{38}$$

  where τ_c is the correlation time, R_ν is the GM process steady-state variance, $\tau_{\max}$ is the bin size length in seconds for the maximum point of the hump, and $\sigma_{A\max}$ is the root Allan variance for the maximum point of the hump.

If only one type of the noise is present, then the Allan variance plot has the specific curve shape and the measurement noise characteristics can be easily computed. When two or more noises are present simultaneously, the shape is a combination (but not the sum) of the individual shapes.

2) Derivation of calibration ALS method: The ALS method derived in this paper can be easily adapted for the noise characteristic estimation for the process defined by (33). The method can be thus viewed as an alternative[3] to the Allan variance used for the sensor calibration.

Following derivation of the ALS method and assuming the precise knowledge of the state $\mathbf{x}_k$, the measurement prediction error is

$$\mathbf{e}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k = \mathbf{v}_k, \tag{39}$$

where $\hat{\mathbf{z}}_k = \mathbf{x}_k$ is the measurement prediction[4] (stemming from (33)). It means that the measurement prediction error is directly given by the measurement noise which has the following statistics

$$\boldsymbol{\beta} = E[\mathbf{e}_k], \tag{40}$$

$$\mathcal{C}_0 = cov[\mathbf{e}_k] = \mathbf{R}_\nu + \mathbf{R}_u, \tag{41}$$

$$\mathcal{C}_j = cov[\mathbf{e}_k, \mathbf{e}_{k-j}] = \boldsymbol{\alpha}^j \mathbf{R}_\nu. \tag{42}$$

[3]The name calibration ALS (cALS) method was, therefore, chosen.

[4]Note that if $\mathbf{x}_k = \mathbf{0}$, then $\hat{\mathbf{z}}_k = \mathbf{0}$ and thus, $\mathbf{e}_k = \mathbf{z}_k$, where $\mathbf{0}$ is a zero vector of the appropriate dimension.

The statistics can be estimated on the basis of the available sequence $\{\mathbf{e}_k\}_{k=0}^T$ according to (17)–(19), where $\hat{\boldsymbol{\beta}}$ represents again the unbiased estimate of (40).

Analogously to the ALS method, the estimate of the unknown matrices $\mathbf{R}_\xi$, $\mathbf{R}_u$, and $\boldsymbol{\alpha}$ is given by (21) with

$$\boldsymbol{\theta} = [(\mathbf{R}_\xi)_{vec}^T, (\mathbf{R}_u)_{vec}^T, \boldsymbol{\alpha}_{vec}^T]^T, \tag{43}$$

$$\mathcal{G}(\boldsymbol{\theta}) = \left([\boldsymbol{\gamma}_0(\boldsymbol{\theta})^T, \boldsymbol{\gamma}_1(\boldsymbol{\theta})^T, \ldots, \boldsymbol{\gamma}_{N-1}(\boldsymbol{\theta})^T]^T\right)_{vec}, \tag{44}$$

$$\hat{\mathbf{b}} = \left(\hat{\mathcal{C}}_e(N)\right)_{vec}. \tag{45}$$

3) Algorithm of calibration ALS method (a scalar case): The *algorithm of the cALS method for measurement noise characteristics estimation* for the scalar system is given by the following steps:

i. Compute measurement prediction error sequence $\{e_k\}_{k=0}^T$ according to (39).
ii. Compute sample-based estimate of the mean β of the sequence according to (19).
iii. Compute sample-based estimates of the variance $\mathcal{C}_0$ and covariances $\mathcal{C}_j$ according to (17) and (18), respectively, for $j = 1, \ldots, N-1$ with $N \geq 4$.
iv. Compute estimates of the *unknown* variances and the parameter gathered in the vector

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3] = [R_\nu, R_u, \alpha]^T, \tag{46}$$

by recursive computation of the following steps of the NLS

$$\hat{\mathbf{b}} = [\hat{\mathcal{C}}_0, \hat{\mathcal{C}}_1, \ldots, , \hat{\mathcal{C}}_{N-1}]^T, \tag{47}$$

$$\mathcal{G}(\hat{\boldsymbol{\theta}}_i) = \begin{bmatrix} \boldsymbol{\gamma}_0(\hat{\boldsymbol{\theta}}_i) \\ \boldsymbol{\gamma}_1(\hat{\boldsymbol{\theta}}_i) \\ \vdots \\ \boldsymbol{\gamma}_{N-1}(\hat{\boldsymbol{\theta}}_i) \end{bmatrix} = \begin{bmatrix} \hat{\theta}_{i,1}+\hat{\theta}_{i,2} \\ \hat{\theta}_{i,3}\hat{\theta}_{i,1} \\ \vdots \\ \hat{\theta}_{i,3}^{(N-1)}\hat{\theta}_{i,1} \end{bmatrix} \tag{48}$$

$$\delta\hat{\mathbf{b}}_i = \hat{\mathbf{b}} - \mathcal{G}(\hat{\boldsymbol{\theta}}_i), \tag{49}$$

$$\mathbf{G}_i = \frac{\partial \mathcal{G}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}|_{\hat{\boldsymbol{\theta}}_i} = \begin{bmatrix} 1 & 1 & 0 \\ \hat{\theta}_{i,3} & 0 & \hat{\theta}_{i,1} \\ \vdots & \vdots & \vdots \\ \hat{\theta}_{i,4}^{(N-1)} & 0 & (N-1)\hat{\theta}_{i,1}\hat{\theta}_{i,3}^{(N-2)} \end{bmatrix} \tag{50}$$

$$\delta\hat{\boldsymbol{\theta}}_i = (\mathbf{G}_i^T\mathbf{G}_i)^{-1}\mathbf{G}_i^T\delta\hat{\mathbf{b}}_i \tag{51}$$

$$\hat{\boldsymbol{\theta}}_{i+1} = \hat{\boldsymbol{\theta}}_i + \delta\hat{\boldsymbol{\theta}}_i. \tag{52}$$

The recursion ends once the error-variable $\delta\hat{\boldsymbol{\theta}}_i$ is negligible [20], [21].

C. Notes

Note 1: Number of considered equations N is a user-defined parameter and it affects the number of ALS estimable parameters. However, even if a sufficient number of equations is considered, not all unknown parameters are generally estimable (especially those tied with the matrix $\mathbf{Q}$ in (1)). It generally depends on the dimension of the state and measurement vectors, i.e., on n_x and n_z. A related discussion for linear systems with white state and measurement noises can be found in [19].

Note 2: Heretofore, the time-correlated measurement noise error contributor $\boldsymbol{\nu}_k$ (4) was considered in the form of the first order GM process. Generally, a higher order GM process can be considered as well.

Note 3: In the problem formulation, the knowledge of the initial condition was assumed. However, because of the stable system assumption, the state prediction error (9) remains stable as well independently of the chosen initial condition of the predictor.

Note 4: The proposed ALS and cALS methods are, in fact, based on the analysis of the properties of the error of the *multistep* predictor estimates. On the contrary, the original ALS method [3] for linear systems with white noises was based on an analysis of the *one-step* prediction error which is not restricted by the assumption on the stable system (1), (2). In principle, the proposed methods can be also based on the one-step predictor, however, as it is illustrated in Appendix B, the resulting relations are much more complex which makes the application of the NLS method more difficult in terms of its convergence.

Note 5: If certain properties of the unknown matrices are known (for example positive-definiteness of the covariance matrices, admitted range for the matrix $\boldsymbol{\alpha}$), then such information might be used in an NLS method design phase [3], [20], [21].

Note 6: Convergence of the minimisation technique used in the NLS solution cannot be guaranteed even if the (sample-based) unbiased estimate of covariance matrices (14), (15) or (41), (42) is available. Convergence is affected either by the particular selected minimisation technique, by a setting of the initial values, or a structure of the considered task [20], [21].

Note 7: If two or more different noises are affecting the measurement equation, then the resulting AVAR shape might be hard to analyze visually and thus to identify the significant points needed for computation of the noise characteristics. This situation usually occurs when covariance of one noise is dominant over the others or when time constant of different correlated noises are close to each other. Such a issue cannot occur if the cALS is used instead of the AVAR.

Note 8: The cALS method allows in principle estimation of the constant bias β (3) similarly as the ALS method does. Due to the definition of the AVAR method, it is not possible to identify the constant bias by the method.

Note 9: Allan variance method processes a scalar signal. Multidimensional signal can be investigated using this method but its scalar components has to be uncorrelated and processed individually as a scalar signals.

IV. Numerical Illustration

The proposed ALS method (and its derivative the cALS) is illustrated in this section using two numerical simulations.

A. Dynamic system

In the first example, the linear dynamic system (1), (2) is assumed with the following parameters

$$n_x = 1, n_z = 1, \tag{53}$$
$$F = -0.8, H = 1, \tag{54}$$
$$Q = 1.5, R_u = 0.8, \tag{55}$$
$$R_\xi = 0.5, \alpha = 0.9, \beta = 2 \tag{56}$$

using $M = 10^4$ Monte-Carlo (MC) simulations with $T = \{10^4, 10^5\}$ samples per one MC simulation and with $N = 4$ equations used in the NLS. The estimates are plotted in Fig. 2 in the form of a histogram. It can be seen that the estimates are unbiased and its variance is going to zero with increasing number of samples T.

B. Sensor calibration

The second example illustrates application of the cALS method estimating the properties of the measurement noise of the structure (33). In this example, the following parameters are considered

$$n_x = 1, n_z = 1, \tag{57}$$
$$R_u = 1, R_\xi = 0.005 \tag{58}$$
$$\alpha = 0.999, \beta = 0. \tag{59}$$

Note that the GM process (4) with the considered parameters R_ξ and α can be understood as a discrete-time alternative to a continuous-time process [22] with time constant $\tau = 10\ [sec]$, steady-state variance $R_\nu = 2.5$, and sample frequency of $f_s = 100\ [Hz]$. The white noise power spectral density (PSD) is, in this case, $N_u = R_u/f_s = 0.01\ [(-)^2/Hz]$.

The histograms of the estimates of R_u, R_ν, and α computed by $M = 10^4$ MC simulations with $T = 10^6$ samples per one MC simulation and with $N = 4$ equations used in the NLS are shown in Fig. 3.

For the sake of completeness, application of the AVAR method (as a standard tool in this area) is illustrated below.

Based on the available (sampled simulated) data, the AVAR plot is constructed using (35). The plot is shown in Fig. 1. The significant points are located as

- white noise point coordinates (slope -1/2): $\sigma_A(\tau_M) = 0.5051$, $\tau_M = 0.03981\ [sec]$,
- time-correlated noise hump maximum point coordinates: $\sigma_{A\max}(\tau_{\max}) = 0.9756$, $\tau_{\max} = 19.95\ [sec]$.

White noise variance can be computed using the respective point coordinates and equation (36) assuming $R_u = 1.0157$. Time-correlated noise characteristics can be computed using the maximum point coordinates and equations (38), (37) assuming $\tau_c = 9.9950\ [sec]$ and $R_\nu = 2.4761$. The discrete parameter α is related to the time constant τ_c by equation

$$\alpha = e^{\frac{-1}{\tau_c f_s}}, \tag{60}$$

Table I
EXAMPLE B: AVAR AND AVERAGED CALS ESTIMATES OF MEASUREMENT NOISE CHARACTERISTICS.

	true	AVAR estimates	**cALS estimates (aver.)**
R_u	1	1.0157	1.0000
α	0.999	0.9991	0.9990
R_ξ	0.005	0.0047	0.0050

which results in $\alpha = 0.9991$. The driving noise variance R_ξ is related to steady-state variance R_ν and α as

$$R_\xi = R_\nu \left(1 - \alpha^2\right) \tag{61}$$

which leads to $R_\xi = 0.0047$.

The results of the AVAR for one particular measurement sequence realisation and the averaged results of the proposed cALS method over all MC simulations (depicted in Fig. (3)) are summarized in Table I.

V. CONCLUDING REMARKS

The paper was devoted to the estimation of the noise characteristics of a linear system described by the state space model. The autocovariance least-squares correlation method, originally designed for linear systems with white noises, was extended for linear systems with measurement noise correlated in time. Then, the proposed method was modified into the form suitable for sensor calibration which is typically solved by the Allan variance method. Compared to the Allan variance method, the proposed correlation method was analytically derived and was shown to provide unbiased estimates which converge to the true values with increasing number of data. The proposed method was thoroughly discussed, analysed, and illustrated by two examples; first, considering estimation of the noise characteristics of the whole state space model, second, considering estimation of the measurement noise characteristics only.

Acknowledgement: This work was supported by the ALFA Program of the Technology Agency of the Czech Republic, project No. TA03030674 (RESPOL).

VI. APPENDIX A: AUTOCOVARIANCE LEAST-SQUARES METHOD FOR WHITE STATE AND MEASUREMENT NOISES - BRIEF REVIEW

The correlation methods have received quite considerable attention in the past as they provide unbiased estimates with acceptable computational requirements even for high-dimensional systems. The methods are based on an analysis of the innovation sequence of "non-optimal" linear estimators. The methods have been pioneered in [1], [2] and then further developed in [3]–[5], [19].

The ALS method for the system description (1), (2) with white state and measurement noises (i.e., with $\boldsymbol{\alpha} = \mathbf{0}$, $\boldsymbol{\beta} = \mathbf{0}$) is based on the analysis of the output of a "non-optimal" linear predictor (either one-step or multistep)

$$\hat{\mathbf{x}}_{k+1} = \mathbf{F}(\hat{\mathbf{x}}_k + \mathbf{K}\mathbf{e}_k), \tag{62}$$
$$\hat{\mathbf{z}}_k = \mathbf{H}\hat{\mathbf{x}}_k, \tag{63}$$

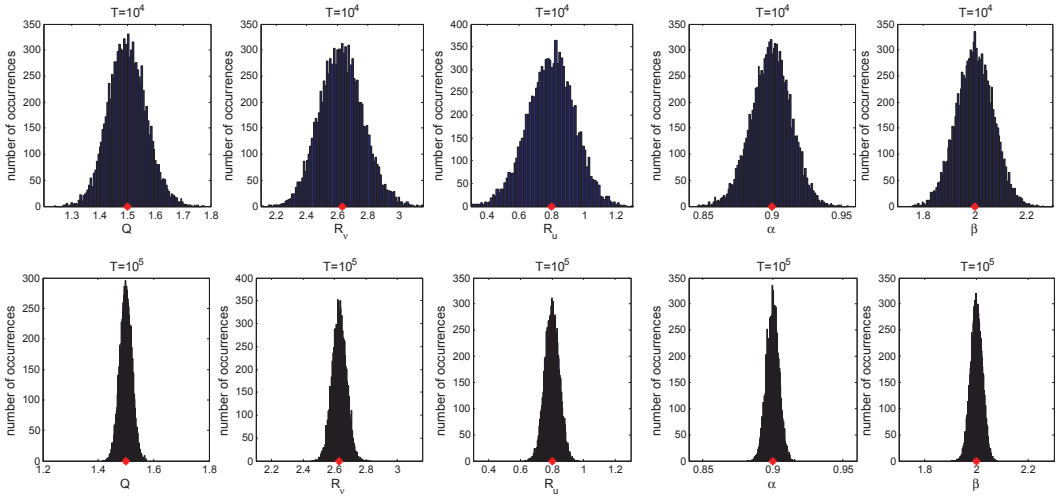

Figure 2. Example A: ALS estimates of state and measurement noise characteristics in form of histograms (true values denoted by red star).

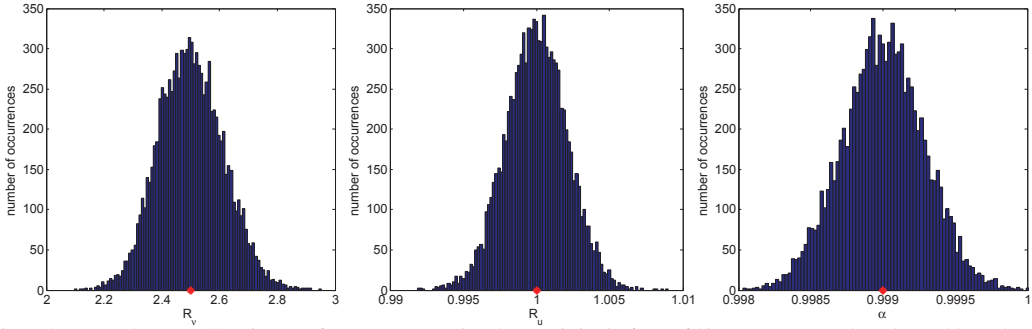

Figure 3. Example B: cALS estimates of measurement noise characteristics in form of histograms (true values denoted by red star).

with the state and measurement prediction errors given by

$$\boldsymbol{\varepsilon}_{k+1} = \bar{\mathbf{F}}\boldsymbol{\varepsilon}_k + \mathbf{G}\bar{\mathbf{w}}_k, \tag{64}$$

$$\mathbf{e}_k = \mathbf{H}\boldsymbol{\varepsilon}_k + \mathbf{v}_k. \tag{65}$$

where $\mathbf{K}$ is a (user-defined) predictor gain[5], $\bar{\mathbf{F}} = (\mathbf{F} - \mathbf{FKH})$, $\mathbf{G} = [\mathbf{G}_1, \mathbf{G}_2] = [\mathbf{I}_{n_x}, -\mathbf{FK}]$, $\bar{\mathbf{w}}_k = [\mathbf{w}_k^T, \mathbf{v}_k^T]^T$, and $\mathbf{0}$ is a zero matrix or vector of appropriate dimension.

The (cross-)covariance matrices of the measurement prediction error are

$$\mathbf{C}_{e,0} = E[\mathbf{e}_k\mathbf{e}_k^T] = \mathbf{H}\mathbf{P}_\epsilon\mathbf{H}^T + \mathbf{R}, \tag{66}$$

$$\mathbf{C}_{e,j} = E[\mathbf{e}_k\mathbf{e}_{k-j}^T] = \mathbf{H}\bar{\mathbf{F}}^j\mathbf{P}_\epsilon\mathbf{H}^T - \mathbf{H}\bar{\mathbf{F}}^{j-1}\mathbf{FKR}, \tag{67}$$

where $\mathbf{P}_\varepsilon$ is a solution to the Lyapunov equation

$$\mathbf{P}_\varepsilon = \bar{\mathbf{F}}\mathbf{P}_\varepsilon\bar{\mathbf{F}}^T + \mathbf{G}\begin{bmatrix}\mathbf{Q} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{bmatrix}\mathbf{G}^T, \tag{68}$$

can be written in a form for the least-squares method as $\mathcal{A}\boldsymbol{\theta} = \mathbf{b}$ with $\boldsymbol{\theta} = [\mathbf{Q}_{vec}^T, \mathbf{R}_{vec}^T]^T$, $\mathbf{b} = (\mathcal{C}_e(N))_{vec}$, and

$$\mathcal{A} = [\mathbf{D}, \mathbf{D}(\mathbf{FK}\otimes\mathbf{FK}) + (\mathbf{I}_{n_z}\otimes\boldsymbol{\Gamma})], \tag{69}$$

$$\mathbf{D} = (\mathbf{H}\otimes\mathcal{O})(\mathbf{I}_{n_x^2} - \bar{\mathbf{F}}\otimes\bar{\mathbf{F}})^{-1}, \tag{70}$$

$$\mathcal{O} = [\mathbf{H}^T, (\mathbf{H}\bar{\mathbf{F}})^T, \ldots, (\mathbf{H}\bar{\mathbf{F}}^{N-1})^T]^T, \tag{71}$$

$$\boldsymbol{\Gamma} = [\mathbf{I}_{n_z}, -(\mathbf{HFK})^T, \ldots, , -(\mathbf{H}\bar{\mathbf{F}}^{N-2}\mathbf{FK})^T]^T. \tag{72}$$

[5]If the gain is zero, the predictor provides multistep prediction. Otherwise, the one-step prediction is outputted.

The estimate of the vector of parameters $\boldsymbol{\theta}$ (composed by the elements of the unknown noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$) in the least-squares sense is given by [3], [19]

$$\hat{\boldsymbol{\theta}} = (\mathcal{A}^T\mathcal{A})^{-1}\mathcal{A}^T\hat{\mathbf{b}} = \mathcal{A}^\dagger\hat{\mathbf{b}}, \tag{73}$$

where the sample-based estimate of the vector $\hat{\mathbf{b}}$ is computed according to (24).

VII. Appendix B: ALS and cALS methods based on One-Step Predictor

Heretofore, the proposed ALS method for linear systems with time-correlated measurement noise was based on the multistep predictor (7), (8). The method can be, however, extended for a more general one-step predictor (62), (63).

Considering the system description (1), (2) with time-correlated measurement noise, the one-step predictor (62), (63), and the associated prediction error model of the form given by (64), (65), the (cross-)covariance matrices of the measurement prediction error can be derived[6] as

$$\begin{aligned}\mathbf{C}_{e,0} &= E[\mathbf{e}_k\mathbf{e}_k^T] = E[(\mathbf{H}\boldsymbol{\varepsilon}_k + \boldsymbol{\nu}_k + \mathbf{u}_k)(\cdot)^T]\\ &= \mathbf{H}\mathbf{P}_\varepsilon\mathbf{H}^T + \mathbf{H}\mathbf{P}_{\varepsilon v} + \mathbf{P}_{v\varepsilon}\mathbf{H}^T + \mathbf{R}_\nu + \mathbf{R}_u,\end{aligned} \tag{74}$$

[6]The constant vector $\boldsymbol{\beta}$ affecting the mean of a prediction error is not considered here for the sake of brevity. It is estimation is, however, analogous to the multistep predictor case.

$$
\begin{aligned}
\mathbf{C}_{e,j} &= E[\mathbf{e}_k \mathbf{e}_{k-j}^T] \\
&= E[(\mathbf{H}\boldsymbol{\varepsilon}_k + \boldsymbol{\nu}_k + \mathbf{u}_k)(\mathbf{H}\boldsymbol{\varepsilon}_{k-j} + \boldsymbol{\nu}_{k-j} + \mathbf{u}_{k-j})^T] \\
&= E[\Big(\mathbf{H}\bar{\mathbf{F}}^j \boldsymbol{\varepsilon}_{k-1} + \sum_{i=1}^{j} \mathbf{H}\bar{\mathbf{F}}^{(i-1)} \mathbf{w}_{k-i} \\
&+ \sum_{i=1}^{j} \mathbf{H}\bar{\mathbf{F}}^{(i-1)} \mathbf{G}_2 (\boldsymbol{\nu}_{k-i} + \mathbf{u}_{k-i}) + \boldsymbol{\nu}_k + \mathbf{u}_k \Big) \\
&\times \Big(\mathbf{H}\boldsymbol{\varepsilon}_{k-j} + \boldsymbol{\nu}_{k-j} + \mathbf{u}_{k-j}\Big)^T] \\
&= \mathbf{H}\bar{\mathbf{F}}^j \mathbf{P}_\varepsilon \mathbf{H}^T + \mathbf{H}\bar{\mathbf{F}}^j \mathbf{P}_{\varepsilon v} + \mathbf{C}\bar{\mathbf{F}}^{(j-1)} \mathbf{G}_2 \mathbf{R}_u \\
&+ \sum_{i=1}^{j} \mathbf{H}\bar{\mathbf{F}}^{(i-1)} \mathbf{G}_2 \boldsymbol{\alpha}^{(j-i)} \mathbf{P}_{v\varepsilon} \mathbf{H}^T + \boldsymbol{\alpha}^j \mathbf{P}_{v\varepsilon} \mathbf{H}^T \\
&+ \sum_{i=1}^{j} \mathbf{H}\bar{\mathbf{F}}^{(i-1)} \mathbf{G}_2 \boldsymbol{\alpha}^{(j-i)} \mathbf{R}_\nu + \boldsymbol{\alpha}^j \mathbf{R}_\nu, \qquad (75)
\end{aligned}
$$

where $\mathbf{P}_\varepsilon$ is a solution to the Lyapunov equation

$$
\mathbf{P}_\varepsilon = \bar{\mathbf{F}} \mathbf{P}_\varepsilon \bar{\mathbf{F}}^T + \bar{\mathbf{F}}^T \mathbf{P}_{\varepsilon v} \mathbf{G}_2^T + \mathbf{G}_2 \mathbf{P}_{v\varepsilon} \bar{\mathbf{F}}^T + \mathbf{G}_2 \mathbf{R} \mathbf{G}_2 + \mathbf{Q} \qquad (76)
$$

with the cross-covariance matrices $\mathbf{P}_{\varepsilon v}$ and $\mathbf{P}_{v\varepsilon}$ computed from

$$
\mathbf{P}_{\varepsilon v} = \bar{\mathbf{F}} \mathbf{P}_{\varepsilon v} \boldsymbol{\alpha}^T + \mathbf{G}_2 \mathbf{R} \boldsymbol{\alpha}^T, \qquad (77)
$$

$$
\mathbf{P}_{v\varepsilon} = \boldsymbol{\alpha} \mathbf{P}_{v\varepsilon} \bar{\mathbf{F}}^T + \boldsymbol{\alpha} \mathbf{R} \mathbf{G}_2^T. \qquad (78)
$$

Equations (74), (75) are then used in construction of the equations for the NLS (21)–(25) analogously to the usage of (14), (15).

It can be seen that the (cross-)covariance matrices of the measurement prediction error sequence are much more complex with respect to those based just on the multistep predictor. The reason can be found in the dependence of the state prediction error $\boldsymbol{\varepsilon}_{k+1}$ and the measurement noise $\mathbf{v}_k$, i.e., in the non-zero matrices $\mathbf{P}_{\varepsilon v}$, $\mathbf{P}_{v\varepsilon}$, and $\mathbf{G}_2$.

The more complex relations usually require more careful set-up of the minimisation technique used for the NLS (21). On the other hand, considering the one step predictor, the assumption on the stable matrix $\mathbf{F}$ might be omitted. Rather, the stability of the state prediction error transition matrix $\bar{\mathbf{F}}$, as a function of the (user-defined) predictor gain $\mathbf{K}$, needs to be assumed instead. Then, the solution to the Lyapunov equation (76) for $\mathbf{P}_\epsilon$ exists [28].

REFERENCES

[1] R. K. Mehra, "On the identification of variances and adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175–184, 1970.

[2] P. R. Bélanger, "Estimation of noise covariance matrices for a linear time-varying stochastic process," *Automatica*, vol. 10, no. 3, pp. 267–275, 1974.

[3] B. J. Odelson, M. R. Rajamani, and J. B. Rawlings, "A new autocovariance least-squares method for estimating noise covariances," *Automatica*, vol. 42, no. 2, pp. 303–308, 2006.

[4] B. M. Åkesson, J. B. Jørgensen, N. K. Poulsen, and S. B. Jørgensen, "A generalized autocovariance least-squares method for Kalman filter tuning," *Journal of Process Control*, vol. 18, no. 7–8, pp. 769–779, 2008.

[5] M. Šimandl and J. Duník, "Estimation of noise covariance matrices for periodic systems," *International Journal of Adaptive Control and Signal Processing*, vol. 25, pp. 928–942, 2011.

[6] D. G. Lainiotis, "Optimal adaptive estimation: Structure and parameters adaptation," *IEEE Transactions on Automatic Control*, vol. 16, no. 2, pp. 160–170, 1971.

[7] R. L. Kashyap, "Maximum likelihood identification of stochastic linear systems," *IEEE Transactions on Automatic Control*, vol. 15, no. 1, pp. 25–34, 1970.

[8] K. A. Myers and B. D. Tapley, "Adaptive sequential estimation with unknown noise statistics," *IEEE Transactions on Automatic Control*, vol. 21, no. 8, pp. 520–523, 1976.

[9] S. Verdú and H. V. Poor, "Minimax linear observers and regulators for stochastic systems with uncertain second-order statistics," *IEEE Transactions on Automatic Control*, vol. 29, no. 6, pp. 499–511, 1984.

[10] P. van Overshee and B. de Moor, *Subspace Identification for Linear Systems: Theory-Implementation-Applications*. Kluwer Academic Publishers, 1996.

[11] L. Ljung, *System Identification: Theory for the User*. UpperSaddle River, NJ: Prentice-Hall, 1999.

[12] D. M. Wiberg, T. D. Powell, and D. Ljungquist, "An on-line parameter estimator for quick convergence and time-varying linear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1854–1863, 2000.

[13] S. Särkkä and A. Nummenmaa, "Recursive noise adaptive Kalman filtering by variational bayesian approximations," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 596–600, 2009.

[14] B. Carew and P. R. Bélanger, "Identification of optimum steady-state gain for systems with unknown noise covariances," *IEEE Transactions on Automatic Control*, vol. 18, no. 6, pp. 582–587, 1973.

[15] R. Bos, X. Bombois, and P. M. J. Van den Hof, "Designing a Kalman filter when no noise covariance information is available," in *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, July 2005.

[16] R. K. Mehra, "Approaches to adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 17, no. 10, pp. 693–698, 1972.

[17] J. Duník, M. Šimandl, and O. Straka, "Methods for estimating state and measurement noise covariance matrices: Aspects and comparison," in *Proceedings of 15th IFAC Symposium on System Identification*, Saint-Malo, France, July 2009.

[18] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp. 772–781, 1978.

[19] M. R. Rajamani and J. B. Rawlings, "Estimation of the disturbance structure from data using semidefinite programming and optimal weighting," *Automatica*, vol. 45, no. 1, pp. 142–148, 2009.

[20] D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*. John Wiley & Sons, 1988.

[21] B. P. Gibbs, *Advanced Kalman Filtering, Least-Squares and Modelling*. John Wiley & Sons, 2011.

[22] R. M. Rogers, *Applied Mathematics in Integrated Navigation Systems (2nd Edition)*. AIAA, 2003.

[23] J. A. Farrell, *Aided Navigation - GPS with High Rate Sensors*. McGraw Hill, 2008.

[24] "IEEE standard specification format guide and test procedure for single-axis interferometric fiber optic gyros," *IEEE Std. 952-1997*, 2008.

[25] "IEEE standard definitions of physical quantities for fundamental frequency and time metrology - random instabilities," *IEEE Std. 952-1997*, 1999.

[26] L. C. Ng and D. J. Pines, "Characterization of ring laser gyro performance using the Allan variance method," *Journal of Guidance, Control, and Dynamics*, vol. 20, no. 1, pp. 211–214, 1997.

[27] M. Soták, "Determining stochastic parameters using an unified method," *Acta Electrotechnica et Informatica*, vol. 9, no. 2, pp. 59–63, 2009.

[28] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006.