Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-15 (2024) 133–138

A comparison of indirect and direct filter designs from data for LTI systems: the effect of unknown noise covariance matrices

M. Mazzoleni *, **L. Maurelli** *, **S. Formentin** **, **F.Previdi** *

* *Department of Management, Information and Production engineering University of Bergamo, via Marconi 5, 24044 Dalmine (BG), Italy.*
** *Dipartimento di Elettronica, Informazione e Biongegneria, Politecnico di Milano, via G. Ponzio 34/5, 20133 Milano, Italy*

Abstract: Existing literature on model-based filter design for stochastic LTI systems assumes complete correspondence between the system and its model. When the system is not completely known, the standard *indirect* model-based (two-steps) filtering solution consists of: (i) identify a model of the system from measured input/output data; (ii) design a Kalman filter based on the estimated model. The performance of this indirect approach are limited by the model and noise covariance matrices accuracy. To overcome such limitations, this paper investigates a *direct* (one-step) solution to the filtering problem for SISO LTI systems in the Prediction Error Method (PEM) identification framework. Simulation results indicate the effectiveness of the direct filtering approach, especially when the noise covariance matrices are misspecified.

1. INTRODUCTION

Consider a discrete-time SISO LTI stochastic system:

$$\mathcal{S}: \quad \begin{aligned} \boldsymbol{x}_{k+1} &= \mathbf{A}\boldsymbol{x}_k + \mathbf{B}u_k + \boldsymbol{w}_k, && \text{(1a)} \\ y_k &= \mathbf{C}\boldsymbol{x}_k + \mathbf{D}u_k + v_k, && \text{(1b)} \\ \boldsymbol{z}_k &= \mathbf{F}\boldsymbol{x}_k + \mathbf{G}u_k, && \text{(1c)} \end{aligned}$$

where $\boldsymbol{x}_k \in \mathbb{R}^{n_x}$ is the system state at time $k \in \mathbb{N}$, $u_k \in \mathbb{R}$ is the input, $y_k \in \mathbb{R}$ is the output and $\boldsymbol{z}_k \in \mathbb{R}^{n_z}$ is a set of variables to be estimated. Let WN denote a white noise stochastic process, with $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{F}, \mathbf{G}$ the system matrices with proper dimensions. The noises $\boldsymbol{w}_k \sim \text{WN}(\mathbf{0}_{n_x}, \mathbf{Q}), v_k \sim \text{WN}(0, R)$, with $\mathbf{Q} \in \mathbb{R}^{n_x \times n_x}, \mathbf{Q} \succeq 0$, and $R \in \mathbb{R}_{>0}$, are independent so that $\boldsymbol{w}_k \perp v_k \perp \boldsymbol{x}_0$. The notation $\mathbf{0}_{n_x} \in \mathbb{R}^{n_x}$ denotes the zero vector.

The filtering problem for the system (1) consists in providing a possibly optimal estimate $\hat{\boldsymbol{z}}_k$ of $\boldsymbol{z}_k$ given a set of past measurements $\{u_\tau, y_\tau\}_{\tau=1}^{k}$ with $\tau \leq k$. The specific meaning of $\boldsymbol{z}_k$ varies according to application. For instance, the well known full state estimation problem is recovered by setting $\mathbf{F} = I_{n_x}$, with $I_{n_x} \in \mathbb{R}^{n_x \times n_x}$ the identity matrix, and $\mathbf{G} = \mathbf{0}$ (Anderson and Moore, 1979, Chapter 3). Setting $\mathbf{F} = \mathbf{C}$ and $\mathbf{G} = \mathbf{D}$ leads to full a output estimation problem commonly employed for fault diagnosis purposes (Ding, 2013, Chapter 5). In these examples, optimal (e.g. minimum estimation error variance in the case of Kalman filtering) estimates of $\boldsymbol{z}_k$ are obtained when a model of the system is perfectly known. When the system is not completely known, a two-steps rationale, that we denote as *model-based (indirect) filtering*, is commonly employed:

(S1) a model $\hat{\mathcal{S}}$ of the system $\mathcal{S}$ in (1) is identified from a set of N input/output data $\{u_k, y_k\}_{k=1}^{N}$ collected from (1);
(S2) a Kalman filter $\mathcal{K}(\hat{\mathcal{S}})$ for $\mathcal{S}$ is designed using $\hat{\mathcal{S}}$.

The filtering performance of $\mathcal{K}(\hat{\mathcal{S}})$ degrades rapidly when the model $\hat{\mathcal{S}}$ identified in step (S1) is not well representative of the system $\mathcal{S}$ in (1), especially regarding the noise covariance matrices $\mathbf{Q}$ and R that often are *manually set*. In this case, existing approaches consider the design of filters that are robust to model/system mismatches, often assuming a bounded magnitude for the model uncertainty (Voulgaris, 1995; Shaked and Theodor, 1992; Mazzoleni et al., 2023; Boni et al., 2024).

An alternative paradigm, denoted as *direct (data-driven) filtering*, has been studied in a parametric set membership (Milanese et al., 2010; Ruiz et al., 2010) and stochastic (Novara et al., 2012) frameworks. In this rationale, the aim is to directly identify a filter $\hat{\mathcal{D}}$ from data. To this end, consider noisy measurements $\boldsymbol{\ell}_k$ of $\boldsymbol{z}_k$ as

$$\boldsymbol{\ell}_k = \boldsymbol{z}_k + \boldsymbol{e}_k = \mathbf{F}\boldsymbol{x}_k + \mathbf{G}u_k + \boldsymbol{e}_k, \tag{2}$$

where $\boldsymbol{e}_k \sim \text{WN}(\mathbf{0}_{n_z}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{n_z \times n_z}, \boldsymbol{\Sigma} \succeq 0$ $\boldsymbol{e}_k \perp (\boldsymbol{w}_k, v_k, \boldsymbol{x}_0)$. Given a dataset of N observations $\{u_k, y_k, \boldsymbol{\ell}_k\}_{k=1}^{N}$, the direct data-driven design of filters for system (1) is a system identification problem where the inputs of the filter model are (u_k, y_k) and the output is $\boldsymbol{\ell}_k$. Notice that in the direct framework the *noise covariance matrices are implicitly learnt from data* in an optimized way. Moreover, the availability of (2) is generally feasible, as a set of state measurement is usually acquired to validate traditional model-based indirect filter designs. The dataset (2) is thus here assumed to be available for offline identification purposes. In the stochastic framework, the direct filter design approach has been shown to generally possess optimal performance in terms of minimizing the variance of state estimation error, even in presence of undermodeling and for unstable systems (Novara et al., 2012). However, in (Novara et al., 2012), an investigation of the effect of uncertainty in the noise covariance matrices

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2024.08.517

on both indirect and direct filter performance has not been specifically evaluated.

So, this work focuses on a comparison between the indirect and direct data-driven designs of full state filters for stochastic SISO LTI systems (1), studying the effect of noise covariance estimates on the state filtering performance. To obtain such estimates for the noise covariance matrices, we rely on Autocovariance Least Square (ALS) method (Odelson et al., 2006; Kost et al., 2021). The Prediction Error Method (PEM) approach is employed for the identification of the model of the direct filter. Simulation results in this paper show how the direct approach avoids the separate steps of model and noise matrices identification, directly optimizing in a single step the filtering performance (and thus directly searching for the optimal noise covariance matrices that optimize the state prediction performance).

The remainder of the paper is as follows. Section 2 defines the problem statement for both filter design rationales. Section 3 discusses the model-based filter design. Section 4 discusses the direct filter design. Section 5 shows a simulation comparison between the two filter design rationales. Section 6 is devoted to final remarks.

2. PROBLEM STATEMENT

We start defining a set of assumptions common to both model-based and direct data-driven filter design rationales.

Assumption 1. (System order). The order n_x of (1) is known. In filtering applications, the user has often a physical knowledge of the system states, so this assumption is less critical than in identification problems.

Assumption 2. (Measured data). A set of N observations $\mathscr{D} := \{u_\tau, y_\tau, \boldsymbol{\ell}_\tau\}_{\tau=1}^N$ has been collected from (1)-(2).

Assumption 3. (Observability). $(\mathbf{C}, \mathbf{A})$ is observable.

Under Assumptions 1-3, the problem of filter design for system (1) can be stated as follows.

Problem 1. (Full state filter design). Consider the case $\mathbf{F} = I_{n_x}$, $\mathbf{G} = 0$, so that $\boldsymbol{z}_k = \boldsymbol{x}_k$ and so $n_z = n_x$. Design a causal LTI filter using measurements in $\mathscr{D}$ that, given $\{u_\tau, y_\tau\}_{\tau=1}^k$, $\tau \le k$, gives an estimate $\hat{\boldsymbol{z}}_k$ of $\boldsymbol{z}_k$.

2.1 Indirect filter design problem

The two-step design consists first in the identification of a parametric model $\hat{\mathcal{S}}$ of (1a)-(1b) and Kalman filter design $\mathcal{K}(\hat{\mathcal{S}})$ from the identified model. Consider first the identification of system model. Let $\mathcal{M}(\boldsymbol{\theta}_\mathrm{M})$ be a family of parametric models for $\mathcal{S}$, with $\boldsymbol{\theta}_\mathrm{M} \in \mathbb{R}^d$ a parameters vector. Define [1]

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_\mathrm{M} &:= [\mathrm{vec}(\hat{\mathbf{A}})^\top\, \mathrm{vec}(\hat{\mathbf{B}})^\top\, \mathrm{vec}(\hat{\mathbf{C}})^\top\, \mathrm{vec}(\hat{\mathbf{D}})^\top\; \mathrm{vec}(\hat{\mathbf{Q}})^\top\; \hat{R}]^\top \\ &= \left[\hat{\mathbf{A}}_s\; \hat{\mathbf{B}}_s\; \hat{\mathbf{C}}_s\; \hat{\mathbf{D}}_s\; \hat{\mathbf{Q}}_s\; \hat{R}\right] \end{aligned} \quad (3)$$

as the estimated parameters for the model $\mathcal{M}(\boldsymbol{\theta}_\mathrm{M})$, with $d = 2n_x^2 + 2n_x + 2$ and $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}}, \hat{\mathbf{Q}}, R$ the estimate of the respective system matrices, so that $\hat{\mathcal{S}} := \mathcal{M}(\widehat{\boldsymbol{\theta}}_\mathrm{M})$.

[1] The operator $\mathrm{vec}(\mathbf{A})$ is the vectorization of the columns of the matrix $\mathbf{A}$ into a column vector. Throughout the paper, we denote the vectorization of $\mathbf{A}$ as $\mathbf{A}_s$.

The estimates of $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ in (3) can be obtained using the dataset

$$\mathscr{D}_\mathrm{M}^\mathrm{AB} := \{u_k, \boldsymbol{\ell}_k\} \quad k = 1, \ldots, N. \quad (4)$$

Similarly, the estimates of $(\hat{\mathbf{C}}, \hat{\mathbf{D}})$ in (3) can be obtained using

$$\mathscr{D}_\mathrm{M}^\mathrm{CD} := \{(\boldsymbol{\ell}_k, u_k), y_k\} \quad k = 1, \ldots, N. \quad (5)$$

The estimates of $(\mathbf{Q}, R)$ in (3) can be obtained using

$$\mathscr{D}_\mathrm{M}^\mathrm{QR} := \{u_k, y_k\} \quad k = 1, \ldots, N. \quad (6)$$

In the filter design step, a *steady-state* Kalman filter $\mathcal{K}$ is designed from the identified model $\hat{\mathcal{S}}$ so that

$$\hat{\mathcal{K}} := \mathcal{K}(\widehat{\boldsymbol{\theta}}_\mathrm{M}). \quad (7)$$

Given $\{u_\tau, y_\tau\}_{\tau=1}^k, \tau \le k$, the filter (7) produces an estimate $\hat{\boldsymbol{z}}_k^\mathrm{M}$ of $\boldsymbol{z}_k$. Note that the structure and order of the filter have not been chosen, as they depend on the identified model.

2.2 Direct filter design problem

In the direct design of filters from data, instead of a model $\mathcal{M}(\boldsymbol{\theta}_\mathrm{M})$ for (1a)-(1b), a parametric model family $\mathcal{D}(\boldsymbol{\theta}_\mathrm{D})$ is selected for the filter to be designed. An estimate $\widehat{\boldsymbol{\theta}}_\mathrm{D}$ of $\boldsymbol{\theta}_\mathrm{D}$ can be obtained using the dataset

$$\mathscr{D}_\mathrm{D} := \{(u_k, y_k), \boldsymbol{\ell}_k\} \quad k = 1, \ldots, N. \quad (8)$$

where (u_k, y_k) are used as inputs and $\boldsymbol{\ell}_k$ as outputs. The estimated direct filter

$$\hat{\mathcal{D}} := \mathcal{D}(\widehat{\boldsymbol{\theta}}_\mathrm{D}) \quad (9)$$

takes (u_k, y_k) as inputs and provides an estimate $\hat{\boldsymbol{z}}_k^\mathrm{D}$ of $\boldsymbol{z}_k$. Differently from the model-based case, in the direct approach the filter structure must the chosen.

Note that the dataset (8) differs from (4)-(6) although they contain the same data, since different sets of signals are considered to be the inputs and the outputs of the models.

3. INDIRECT FILTER DESIGN

3.1 Estimation of A and B

By substituting (2) into (1a) we get

$$\boldsymbol{\ell}_{k+1} = \mathbf{A}\boldsymbol{\ell}_k + \mathbf{B}u_k + \boldsymbol{\eta}_k = [\mathbf{A}\ \mathbf{B}]\,\boldsymbol{\varphi}_k + \boldsymbol{\eta}_k, \quad (10a)$$

$$\boldsymbol{\varphi}_k := \left[\boldsymbol{\ell}_k^\top\ u_k\right]^\top \in \mathbb{R}^{n_x+1}\ \boldsymbol{\eta}_k := \boldsymbol{e}_{k+1} - \mathbf{A}\boldsymbol{e}_k + \boldsymbol{w}_k \quad (10b)$$

with $\mathbb{E}[\boldsymbol{\eta}_k] = \mathbf{0}_{n_x}$, $\mathrm{Var}[\boldsymbol{\eta}_k] = \boldsymbol{\Sigma} + \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top + \mathbf{Q}$, where $\mathbb{E}[\cdot]$ and $\mathrm{Var}[\cdot]$ are the expectation and variance operators, respectively. Inspection of (13) shows a correlation between $\boldsymbol{\varphi}_k$ and $\boldsymbol{\eta}_k$. Unbiased estimates of $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ are obtained in a least squares sense starting from a multi-targets regression

$$\min_{\mathbf{A},\mathbf{B}} \sum_{k=1}^N \left\|\boldsymbol{\ell}_{k+1} - [\mathbf{A}\ \mathbf{B}]\,\boldsymbol{\varphi}_k\right\|_2^2 \quad (11)$$

using the dataset (4) and relying on an instrumental variable $\boldsymbol{\xi}_k := \left[\boldsymbol{\ell}_{k-1}^\top\ u_k\right]^\top \in \mathbb{R}^{n_x+1}$ so that

$$[\hat{\mathbf{A}}, \hat{\mathbf{B}}] = \left(\sum_{k=2}^N \boldsymbol{\ell}_{k+1}\ \boldsymbol{\xi}_k^\top\right) \cdot \left(\sum_{k=2}^N \boldsymbol{\varphi}_k\ \boldsymbol{\xi}_k^\top\right)^{-1}. \quad (12)$$

3.2 Estimation of C and D

By substituting (2) into (1b) we get

$$y_k = \mathbf{C}\boldsymbol{\ell}_k + \mathbf{D}u_k + \zeta_k = [\mathbf{C}\ \mathbf{D}]\,\boldsymbol{\varphi}_k + \zeta_k, \tag{13a}$$
$$\zeta_k := v_k - \mathbf{C}\boldsymbol{e}_k, \tag{13b}$$

with $\mathbb{E}[\zeta_k] = 0$ and $\mathrm{Var}[\zeta_k] = \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top + R$. Unbiased estimates of $(\hat{\mathbf{C}}, \hat{\mathbf{D}})$ can be obtained in a least squares sense using the dataset (5) similarly to (11)-(12) as

$$[\hat{\mathbf{C}}, \hat{\mathbf{D}}] = \left(\sum_{k=2}^{N} y_k\ \boldsymbol{\xi}_k^\top\right) \cdot \left(\sum_{k=2}^{N} \boldsymbol{\varphi}_k\ \boldsymbol{\xi}_k^\top\right)^{-1}. \tag{14}$$

3.3 Estimation of $\mathbf{Q}$ and R

The design of a full state Kalman filter $\hat{\mathcal{K}}$ for (1) under Assumptions 1-3 based on (12)-(14) requires an estimate $\hat{\mathbf{Q}}, \hat{R}$ of the noise covariances $\mathbf{Q}, R$, so that the model-based state estimate $\hat{\boldsymbol{z}}_k^{\mathrm{M}} = \hat{\boldsymbol{x}}_{k|k}$ is given by

$$\hat{\boldsymbol{x}}_{k|k-1} = \hat{\mathbf{A}}\hat{\boldsymbol{x}}_{k-1|k-1} + \hat{\mathbf{B}}u_{k-1}, \qquad \hat{\boldsymbol{x}}_{1|0} = \boldsymbol{x}_0, \tag{15a}$$
$$\hat{y}_{k|k-1} = \hat{\mathbf{C}}\hat{\boldsymbol{x}}_{k|k-1} + \hat{\mathbf{D}}u_{k-1}, \quad \varepsilon_k := y_k - \hat{y}_{k|k-1}, \tag{15b}$$
$$\mathbf{P}_{k|k-1} = \hat{\mathbf{A}}\mathbf{P}_{k-1|k-1}\hat{\mathbf{A}}^\top + \hat{\mathbf{Q}}, \qquad \mathbf{P}_{1|0} = \mathbf{P}_0, \tag{15c}$$
$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\hat{\mathbf{C}})\mathbf{P}_{k|k-1}, \qquad \hat{y}_{1|0} = y_0, \tag{15d}$$
$$\mathbf{K}_k^{\mathrm{QR}} = \mathbf{P}_{k|k-1}\hat{\mathbf{C}}^\top(\hat{\mathbf{C}}\mathbf{P}_{k|k-1}\hat{\mathbf{C}}^\top + \hat{R})^{-1}, \tag{15e}$$
$$\hat{\boldsymbol{z}}_k^{\mathrm{M}} := \hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \mathbf{K}_k^{\mathrm{QR}}\varepsilon_k, \tag{15f}$$
$$\boldsymbol{\nu}_k^{\mathrm{M}} = \boldsymbol{x}_k - \hat{\boldsymbol{z}}_k^{\mathrm{M}}, \tag{15g}$$

where $\hat{\boldsymbol{x}}_{k|k-1} \in \mathbb{R}^{n_x}$ and $\mathbf{P}_{k|k-1} \in \mathbb{R}^{n_x \times n_x}$ are the one-step state and state covariance matrix predictions at time $k-1$, $\hat{\boldsymbol{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are the filtered state estimate and state covariance matrix at time k, $\boldsymbol{x}_0$, y_0 and $\mathbf{P}_0$ are initialization values, $y_{k|k-1}$ is the predicted output at time $k-1$, ε_k is the innovation, $\mathbf{K}_k^{\mathrm{QR}} \in \mathbb{R}^{n_x}$ is the filter gain (that depends on the noise covariance matrices $\hat{\mathbf{Q}}, \hat{R}$) and $\boldsymbol{\nu}_k^{\mathrm{M}}$ is the filtering error (Anderson and Moore, 1979).

Amongst the methods proposed in literature to estimate the noise covariance matrices, *correlation methods* can be derived analytically with minimal assumptions on the model, and provide, under mild assumptions, consistent and unbiased estimates of the noise covariance matrices (Duník et al., 2017a). In this work we employ a specific one-step correlation approach known as the Autocovariance Least-Square (ALS) method (Odelson et al., 2006; Kost et al., 2021), leveraging the dataset (6). Consider a generic LTI state estimator for (1a)-(1b) under Assumptions 1-3, further assuming known system matrices:

$$\hat{\boldsymbol{x}}_{k|k-1}^{\mathrm{L}} = \mathbf{A}\hat{\boldsymbol{x}}_{k-1|k-1}^{\mathrm{L}} + \mathbf{B}u_{k-1}, \tag{16a}$$
$$\hat{\boldsymbol{x}}_{k|k}^{\mathrm{L}} = \hat{\boldsymbol{x}}_{k|k-1}^{\mathrm{L}} + \mathbf{L}\varepsilon_k^{\mathrm{L}}, \tag{16b}$$
$$\varepsilon_k^{\mathrm{L}} := y_k - \mathbf{C}\hat{\boldsymbol{x}}_{k|k-1}^{\mathrm{L}} - \mathbf{D}u_{k-1}, \tag{16c}$$
$$\boldsymbol{\delta}_k^{\mathrm{L}} := \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}^{\mathrm{L}}, \tag{16d}$$

where $\boldsymbol{\delta}_k^{\mathrm{L}} \in \mathbb{R}^{n_x}$ is the one-step state prediction error and $\mathbf{L} \in \mathbb{R}^{n_x}$ is the filter gain, not necessarily the optimal one that necessitates the knowledge of $\mathbf{Q}, R$. The dynamics of $\boldsymbol{\delta}_k^{\mathrm{L}}$ evolves according to

$$\boldsymbol{\delta}_k^{\mathrm{L}} = (\mathbf{A} - \mathbf{ALC})\boldsymbol{\delta}_{k-1}^{\mathrm{L}} + [I_{n_x}\ -\mathbf{AL}]\begin{bmatrix}\boldsymbol{w}_{k-1}\\ v_{k-1}\end{bmatrix}, \tag{17a}$$
$$= \bar{\mathbf{A}}\boldsymbol{\delta}_{k-1}^{\mathrm{L}} + \mathbf{H}\bar{\boldsymbol{w}}_{k-1}. \tag{17b}$$

The state prediction error $\boldsymbol{\delta}_k$ is related to the innovation $\varepsilon_k^{\mathrm{L}}$ by

$$\varepsilon_k^{\mathrm{L}} = \mathbf{C}\boldsymbol{\delta}_k^{\mathrm{L}} + v_k. \tag{18}$$

Given the steady-state covariance matrix of $\boldsymbol{\delta}_k^{\mathrm{L}}$ described by the Lyapunov equation

$$\bar{\mathbf{P}}_{\boldsymbol{\delta}} = \bar{\mathbf{A}}\bar{\mathbf{P}}_{\boldsymbol{\delta}}\bar{\mathbf{A}}^\top + \begin{bmatrix}\mathbf{Q} & 0\\ \mathbf{0}_{n_x \times n_x} & R\end{bmatrix}, \tag{19}$$

the autocovariance of the innovation sequence reads as

$$\bar{P}_{\varepsilon^{\mathrm{L}},0} := \mathbb{E}\left[(\varepsilon_k^{\mathrm{L}})^2\right] = \mathbf{C}\bar{\mathbf{P}}_{\boldsymbol{\delta}}\mathbf{C}^\top + R, \tag{20a}$$
$$\bar{P}_{\varepsilon^{\mathrm{L}},j} := \mathbb{E}\left[\varepsilon_k^{\mathrm{L}}\varepsilon_{k+j}^{\mathrm{L}}\right] = \mathbf{C}\bar{\mathbf{A}}^j\bar{\mathbf{P}}_{\boldsymbol{\delta}}\mathbf{C}^\top - \mathbf{C}\bar{\mathbf{A}}^{j-1}\mathbf{AL}R, \tag{20b}$$

where $j = 1, 2, \ldots, m-1$ and m is a user-defined parameter that defines the maximum considered lag.

The solution to (19) and its substitution into (20) gives the linear system

$$\boldsymbol{\mathcal{A}}\boldsymbol{\varrho}_{\mathrm{M}} = \mathbf{b}, \tag{21}$$

where

$$\boldsymbol{\mathcal{A}} := \left[\boldsymbol{\mathcal{D}}\ \boldsymbol{\mathcal{D}}(\mathbf{AL} \otimes \mathbf{AL}) + (I_{n_x} \otimes \boldsymbol{\Gamma})\right], \tag{22a}$$
$$\boldsymbol{\mathcal{D}} := (\mathbf{C} \otimes \boldsymbol{\mathcal{O}})\left(I_{n_x^2} - \bar{\mathbf{A}} \otimes \bar{\mathbf{A}}\right)^{-1}, \tag{22b}$$
$$\boldsymbol{\mathcal{O}} := \left[\mathbf{C}^\top\ (\mathbf{C}\bar{\mathbf{A}})^\top \ldots (\mathbf{C}\bar{\mathbf{A}}^{m-1})^\top\right]^\top, \tag{22c}$$
$$\boldsymbol{\Gamma} := \left[1\ -(\mathbf{CAL})^\top\ \ldots\ -(\mathbf{C}\bar{\mathbf{A}}^{m-2}\mathbf{AL})^\top\right]^\top, \tag{22d}$$

and

$$\boldsymbol{\varrho}_{\mathrm{M}} := [\mathbf{Q}_s^\top\ R]^\top, \quad \mathbf{b} := (\boldsymbol{\mathcal{C}}_{\varepsilon}(m))_s \tag{23a}$$

with

$$\boldsymbol{\mathcal{C}}_{\varepsilon^{\mathrm{L}}}(m) := \left[\bar{P}_{\varepsilon^{\mathrm{L}},0}\ \bar{P}_{\varepsilon^{\mathrm{L}},1}\ \ldots\ \bar{P}_{\varepsilon^{\mathrm{L}},m-1}\right]^\top. \tag{24}$$

The solution to (21) can be practically computed in a least square sense as

$$\hat{\boldsymbol{\varrho}}_{\mathrm{M}} = [\hat{\mathbf{Q}}_s^\top\ \hat{R}]^\top = (\boldsymbol{\mathcal{A}}^\top\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{A}}^\top\hat{\mathbf{b}}, \tag{25}$$

employing (12) and (14) into (22) and with $\hat{\mathbf{b}} = \hat{\boldsymbol{\mathcal{C}}}_{\varepsilon^{\mathrm{L}}}(m) = [\hat{\bar{P}}_{\varepsilon^{\mathrm{L}},0}\ \hat{\bar{P}}_{\varepsilon^{\mathrm{L}},1}\ \ldots\ \hat{\bar{P}}_{\varepsilon^{\mathrm{L}},m-1}]^\top$ estimated from

$$\hat{\bar{P}}_{\varepsilon^{\mathrm{L}},j} = \frac{1}{N-j}\sum_{k=1}^{N-j}(\varepsilon_k^{\mathrm{L}} \cdot \varepsilon_{k+j}^{\mathrm{L}}), \quad j = 0, 1, \ldots, m-1. \tag{26}$$

The design of an optimal gain $\mathbf{L}$ in (16b), in terms of a minimum upper bound of the estimate covariance matrices, has been studied in (Duník et al., 2017b), while strategies to enforce the semidefinite positiveness of $\hat{\boldsymbol{\varrho}}$ in (25) are considered in (Rajamani and Rawlings, 2009).

4. DIRECT FILTER DESIGN

4.1 Identification approach with Prediction Error Methods

The aim of direct data-driven design of filters for the system (1) is to identify, in a one-step fashion, a stable LTI model $\mathcal{D}(\boldsymbol{\theta}_{\mathrm{D}})$ that given (u_k, y_k) in input produces an estimate $\hat{\boldsymbol{z}}_k$ of $\boldsymbol{z}_k$, using the dataset (8). Clearly, identifying a model with guaranteed stability is not trivial. The identification problem can be solved resorting to a PEM formulation

$$\widehat{\boldsymbol{\theta}}_{\mathrm{D}} = \arg\min_{\boldsymbol{\theta}_{\mathrm{D}}} \frac{1}{N}\sum_{k=1}^{N}\left\|\boldsymbol{\ell}_k - \widehat{\boldsymbol{\ell}}_{k|k-1}(\boldsymbol{\theta}_{\mathrm{D}})\right\|_2^2 \tag{27}$$

where $\widehat{\boldsymbol{\ell}}_{k|k-1}$ is the one-step prediction from model $\mathcal{D}(\boldsymbol{\theta}_{\mathrm{D}})$. Under Assumptions 1, the system (and thus the full state filter) order is known. Thus, we are left selecting an appropriate parametric structure for $\mathcal{D}(\boldsymbol{\theta}_\mathrm{D})$.

4.2 Choice of the model structure for the direct filter

Assume that the system (1) is completely known. In this case, the Kalman filter recursions (15) provide the best linear unbiased estimator, with minimum variance of the state prediction error. Consider the steady-state filter with $\bar{\mathbf{P}}$ the solution of

$$\bar{\mathbf{P}}^\circ = \mathbf{A}\bar{\mathbf{P}}^\circ\mathbf{A}^\top + \mathbf{Q} - \mathbf{A}\bar{\mathbf{P}}^\circ\mathbf{C}^\top(\mathbf{C}\bar{\mathbf{P}}^\circ\mathbf{C}^\top + R)^{-1}\mathbf{C}\bar{\mathbf{P}}^\circ\mathbf{A}^\top$$

and $\bar{\mathbf{K}}^{\mathrm{QR}}$ the steady state filter gain

$$\bar{\mathbf{K}}^{\mathrm{QR}} := \bar{\mathbf{P}}^\circ\mathbf{C}^\top(\mathbf{C}\bar{\mathbf{P}}^\circ\mathbf{C}^\top + R). \tag{28}$$

Substitution of (15a) into (15f) leads to a dynamic expression for the Kalman filtered state, with $^\circ$ denoting the fact the perfect model is used

$$\begin{aligned}\hat{\boldsymbol{x}}^\circ_{k|k} = (\mathbf{A} - \bar{\mathbf{K}}^{\mathrm{QR}}\mathbf{C}\mathbf{A})\boldsymbol{x}^\circ_{k-1|k-1} + \\ + (\mathbf{B} - \bar{\mathbf{K}}^{\mathrm{QR}}\mathbf{D} - \bar{\mathbf{K}}^{\mathrm{QR}}\mathbf{C}\mathbf{B})u_{k-1} + \bar{\mathbf{K}}y_k. \end{aligned} \tag{29}$$

Writing (29) in operational form and denoting z^{-1} the lag operator so that $y_{k-1} = z^{-1}y_k$, we get

$$\hat{\boldsymbol{x}}^\circ_{k|k} = A(z)^{-1}B_u(z)u_k + A(z)^{-1}B_y(z)y_k, \tag{30a}$$
$$= G^\circ_u(z)u_k + G^\circ_y(z)y_k, \tag{30b}$$
$$A(z) := I_{n_x} - (\mathbf{A} - \bar{\mathbf{K}}^{\mathrm{QR}}\mathbf{C}\mathbf{A})z^{-1},$$
$$B_u(z) := (\mathbf{B} - \bar{\mathbf{K}}^{\mathrm{QR}}\mathbf{D} - \bar{\mathbf{K}}^{\mathrm{QR}}\mathbf{C}\mathbf{B})z^{-1},\ B_y(z) := \bar{\mathbf{K}}^{\mathrm{QR}}.$$

Equation (30a) can be related to state filtering error $\boldsymbol{\nu}^\circ_k := \boldsymbol{x}_k - \hat{\boldsymbol{x}}^\circ_{k|k}$ and the state measurements (2) as

$$\boldsymbol{\ell}_k = \hat{\boldsymbol{x}}^\circ_{k|k} + \boldsymbol{\nu}^\circ_k + \boldsymbol{e}_k = \hat{\boldsymbol{x}}^\circ_{k|k} + \boldsymbol{\gamma}_k, \tag{31a}$$
$$= G^\circ_u(z)u_k + G^\circ_y(z)y_k + \boldsymbol{\gamma}_k, \tag{31b}$$
$$\boldsymbol{\gamma}_k := \boldsymbol{\nu}^\circ_k + \boldsymbol{e}_k, \tag{31c}$$

where $G^\circ_u(z), G^\circ_y(z)$ are $n_x \times 1$ transfer function matrices. Note that $\boldsymbol{\gamma}_k \in \mathbb{R}^{n_x}$ in (31c) is not a white process as $\boldsymbol{\nu}^\circ_k$ contains the terms $\boldsymbol{w}_k$ and $\boldsymbol{w}_{k-1}$. Thus, the model $\mathcal{D}(\boldsymbol{\theta}_\mathrm{D})$ for the data-driven filter must be parametrized with both exogenous and noise models as in Figure 1:

$$\boldsymbol{\ell}_k = G_u(z;\boldsymbol{\theta}_\mathrm{D})u_k + G_y(z;\boldsymbol{\theta}_\mathrm{D})y_k + H(z;\boldsymbol{\theta}_\mathrm{D})\rho_k \tag{32}$$

where ρ_k is a white noise process and $G_u(z;\boldsymbol{\theta}_\mathrm{D}), G_u(z;\boldsymbol{\theta}_\mathrm{D})$, $H(z;\boldsymbol{\theta}_\mathrm{D}) \neq \mathbf{1}_{n_x}$ are $n_x \times 1$ transfer matrices that represent the parameterized models for the exogenous inputs u_k, y_k and noise ρ_k, respectively.

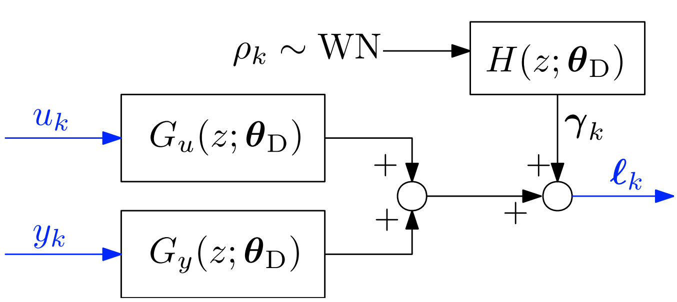

Fig. 1. Data-driven filter stochastic parametric model. Blue quantities denote the measured signals.

The state estimates provided by the direct data-driven approach can be computed by a *simulation* of the identified direct filter $\hat{\mathcal{D}}$ as

$$\hat{\boldsymbol{z}}^\mathrm{D}_k = G_u(z;\widehat{\boldsymbol{\theta}}_\mathrm{D})u_k + G_y(z;\widehat{\boldsymbol{\theta}}_\mathrm{D})y_k. \tag{33}$$

Remark 2. Model (32) can also be parameterized so that the exogenous part and the noise part are independently parameterized, so leveraging Box-Jenkins structures.

Remark 3. The direct filter estimate (33) can not be computed using the one-step model *prediction*, as this would require measurements of $\boldsymbol{\ell}_k$ that are assumed available only offline for the identification of (32).

5. NUMERICAL RESULTS

Consider a system (1) of order $n_x = 3$ sampled at $T_s = 0.01$ s and measurements (2) with

$$\mathbf{A} = \begin{bmatrix} 0.610 & 0.084 & -0.536 \\ -0.139 & 0.270 & 0.763 \\ 0.124 & 0.279 & -0.245 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.558 \\ -0.028 \\ -1.476 \end{bmatrix}, \tag{34}$$
$$\mathbf{C} = [0.259\ \ -2.018\ \ 0.199], \quad \mathbf{D} = 0, \quad R = 1,$$
$$\mathbf{Q} = \mathrm{diag}\,(0.025, 0.05, 0.1), \quad \boldsymbol{\Sigma} = \mathrm{diag}\,(0.067, 0.1, 0.2),$$

where $\mathrm{diag}(\cdot)$ indicates a diagonal matrix.

We simulate $N = 1000$ data $\{y_k, \boldsymbol{\ell}_k\}$ from (34) using a white noise input $u_k \sim \mathrm{WN}(0,1)$. These data are used to define the datasets (4)-(5)-(6). The Signal to Noise Ratio (SNR) between the *noiseless* states $\tilde{\boldsymbol{x}}_k$ of (34) and the state noise $\boldsymbol{w}_k$ is defined as

$$\mathrm{SNR}_{\tilde{x}^{(i)}} = \frac{\mathrm{Var}[\tilde{x}^{(i)}_k]}{\mathrm{Var}[w^{(i)}_k]} \approx 25, \quad i \in \{1,2,3\}, \tag{35}$$

with $\tilde{x}^{(i)}_k, w^{(i)}_k$ are the i-th elements of the noiseless and noise state vectors $\tilde{\boldsymbol{x}}_k, \boldsymbol{w}_k$ respectively. The SNR between the noisy states $\boldsymbol{x}_k$ and the their measurements $\boldsymbol{\ell}_k$ is

$$\mathrm{SNR}_{\ell^{(i)}} = \frac{\mathrm{Var}[x^{(i)}_k]}{\mathrm{Var}[e^{(i)}_k]} \approx 11, \quad i \in \{1,2,3\}, \tag{36}$$

while SNR between the *noiseless* output $\tilde{y}_k$ and the output noise v_k is

$$\mathrm{SNR}_y = \frac{\mathrm{Var}[\tilde{y}_k]}{\mathrm{Var}[v_k]} \approx 6. \tag{37}$$

We compare the following four approaches:

(1) **Nominal KF**: a Kalman Filter (15) with all system $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ and covariance matrices $(\mathbf{Q}, R)$ *known*. This is the best possible linear filtering approach for systems of type (1).
(2) **Indirect KF (Q,R)**: a Kalman Filter (15) employing estimates of the system matrices (12) and (14), but with *known* covariance matrices $\mathbf{Q}, R$.
(3) **Indirect KF**: a Kalman Filter (15) employing estimates of the system matrices (12) and (14), and estimates of covariance matrices (25). This is the *typical situation in practical scenarios.*
(4) **Direct filter**: a direct filter of a proper order and structure (32) is identified from the data (8).

We run $M_C = 100$ Monte-Carlo simulations, varying the realization of the noises $\boldsymbol{w}_k, v_k, \boldsymbol{e}_k$ at each run. For each run, the designed filters are evaluated on a test dataset (fixed for each simulation) of $N_{\mathrm{test}} = 1000$ data generated with a white noise input $u^{\mathrm{test}}_k \sim \mathrm{WN}(0,1)$. The test dataset is different from the identification one, both regarding the input and the realizations of the noises affecting the test data. Filtering performance is evaluated, for each Monte-Carlo simulation and state $x_k^{(i)}, i \in \{1,2,3\}$ of $\boldsymbol{x}_k$, by computing the root mean square error (RMSE) between noisy states $\boldsymbol{x}_k$ and the filtered ones, i.e. (15f) for the Kalman-based approaches and (33) for the direct filter, as

$$\text{Filtering RMSE}_i = \sqrt{\frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} \left(x^{(i)} - \hat{z}_k^{(i)}\right)^2}, \qquad (38)$$

where $\hat{z}_k^{(i)}$ is the i-th element of the state estimate $\hat{\boldsymbol{z}}_k$, that assumes the definition (15f) or (33) according to the considered Kalman-based of direct filter approach.

The approach reviewed in Section 3.3 is applied for the estimation of the noise covariance matrices for **Indirect KF** approach. As suggested in (Odelson et al., 2006; Duník et al., 2017b), usually a low value of the hyperparameter m in (20) is employed and gives satisfactory results. Thus, we set $m = 4$ in the simulations. A second hyperparameter regards the definition of the observer gain $\mathbf{L}$ in (16). As this gain influences the variance of the estimates (Duník et al., 2017b), we fixed the same gain $\mathbf{L}$ for all Monte-Carlo simulations. The gain $\mathbf{L}$ is set close to the optimal Kalman gain (i.e. when considering as known the system and noise matrices). Furthermore, we used the information that $\mathbf{Q}$ is diagonal in the resolution of (25) with a semidefinite program solver. Notice that the last two settings represent an advantage for the **Indirect KF** approach in the simulations, as these prior information are usually not available.

Regarding the **Direct filter** approach, we choose a MIMO Box-Jenkins model structure, with 2 inputs (the input u_k and the output y_k of the system) and 3 outputs (the states measurements $\boldsymbol{\ell}_k$). For each of the input-output transfer functions of the direct filter model, we have the following relation

$$\begin{aligned}\ell_k^{(i)} = &\frac{b_0^{u,i} + b_1^{u,i} z^{-1} + \dots b_{n_b}^{u,i} z^{-n_b}}{1 + f_1^{u,i} z^{-1} + \dots f_{n_f}^{u,i} z^{-n_f}} u_k + \\ &+ \frac{b_0^{y,i} + b_1^{y,i} z^{-1} + \dots b_{n_b}^{y,i} z^{-n_b}}{1 + f_1^{y,i} z^{-1} + \dots f_{n_f}^{y,i} z^{-n_f}} y_k + \\ &+ \frac{1 + c_1^i z^{-1} + \dots c_{n_c}^i z^{-n_c}}{1 + d_1^i z^{-1} + \dots d_{n_d}^i z^{-n_d}} \rho_k, \quad i \in \{1,2,3\}, \qquad (39)\end{aligned}$$

where $\{b_0^{u,i}, \dots, b_{n_b}^{u,i}\}$, $\{f_1^{u,i}, \dots, f_{n_f}^{u,i}\}$ are the coefficients of the transfer function from u_k to the i-th state measurement $\ell_k^{(i)}$, $\{b_0^{y,i}, \dots, b_{n_b}^{y,i}\}$, $\{f_1^{y,i}, \dots, f_{n_f}^{y,i}\}$ are the coefficients of the transfer function from y_k to ℓ_k^i, and $\{c_1^i, \dots, c_{n_c}^i\}$, $\{d_1^i, \dots, d_{n_d}^i\}$ are the coefficients of the transfer function from the white noise ρ_k to $\ell_k^{(i)}$, for $i \in \{1,2,3\}$. The orders n_b, n_f, n_c, n_d are fixed for all the three outputs of the direct filter model. In particular, we set $n_b = n_f = n_c = n_d = n_x$. The parameters vector $\boldsymbol{\theta}_{\text{D}}$ for the direct filter model is defined as

$$\begin{aligned}\boldsymbol{\theta}_{\text{D}}^{(i)} := \big[&b_0^{u,i}, \dots, b_3^{u,i},\ f_1^{u,i}, \dots, f_3^{u,i},\ b_0^{y,i}, \dots, b_3^{y,i}, \\ &f_1^{y,i}, \dots, f_3^{y,i},\ c_1^i, \dots, c_3^i,\ d_1^i, \dots, d_3^i\big] \in \mathbb{R}^{1\times 20}, \\ &i \in \{1,2,3\}, \qquad (40a)\end{aligned}$$

$$\boldsymbol{\theta}_{\text{D}} := \left[\boldsymbol{\theta}_{\text{D}}^{(1)}\ \boldsymbol{\theta}_{\text{D}}^{(2)}\ \boldsymbol{\theta}_{\text{D}}^{(3)}\right]^\top \in \mathbb{R}^{60}. \qquad (40b)$$

The adequacy of these setting is evaluated on the test dataset by residual correlation analysis of the one-step prediction error $\boldsymbol{\ell}_k - \widehat{\boldsymbol{\ell}}_{k|k-1}(\hat{\boldsymbol{\theta}}_{\text{D}})$ of the direct filter model.

The first step in the **Indirect KF (Q,R)** and **Indirect KF** approaches is the estimation of a model for the system (34), using the instrumental variable estimates (12)-(14). Figure 3 depicts the estimates of the elements of of the matrix $\mathbf{A}$ in (34) as an example: the estimates of the other matrices have similar properties.

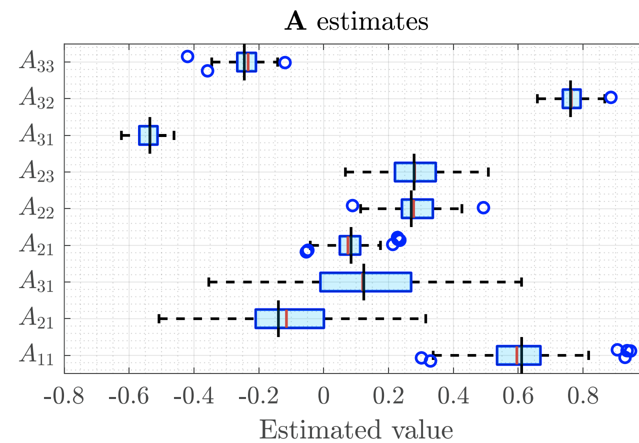

Fig. 2. Estimates of the matrix $\mathbf{A}$ in (34). Each A_{ij} is the (ij)-th element of $\mathbf{A}$. The vertical black line represents the true values of the parameters.

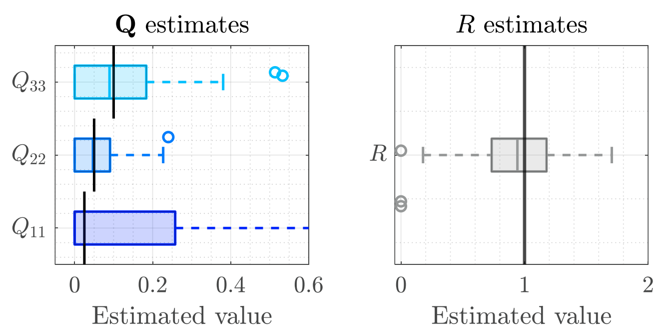

Fig. 3. Estimates of the noise covariance matrices. (Left) Estimates of $\mathbf{Q}$. The Q_{ii} represents the i-th diagonal element of $\mathbf{Q}$. (Right) Estimates of R. The vertical black line represents the true values of the parameters.

Figure 3 reports the estimation results of the noise covariance matrices $\mathbf{Q}, R$ following the approach of Section 3.3 solving (21) in a least-square sense using the MOSEK optimizer via YALMIP with semipositive definite constraints on the covariance matrices and information about the diagonality of $\mathbf{Q}$ (ApS, 2022; Löfberg, 2004).

Figure 4 shows a comparison of the filtering performance of the four approaches, on the test dataset. The **Direct Filter** is superior to both the **Indirect KF (Q,R)** and **Indirect KF** approaches, as it *directly* (and implicitly) *optimizes for the unknown covariance and system matrices.* On the other side, the **Indirect KF (Q,R)** is sensitive to the estimation uncertainty of system matrices, and the **Indirect KF** approach is sensitive to the uncertainty in both system and noise covariance matrices. We recall that

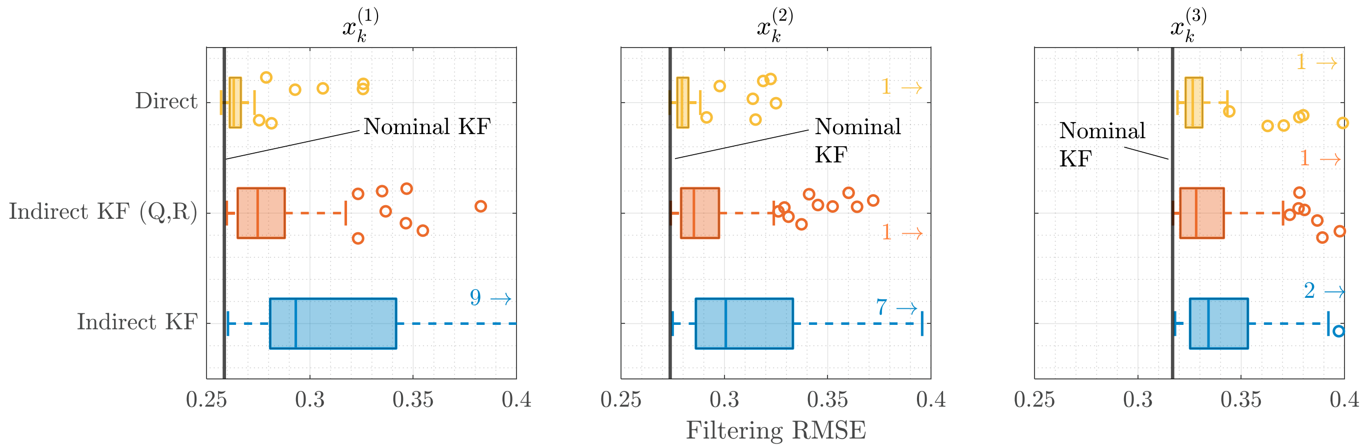

Fig. 4. Filtering error of the four approaches as measured by (38) (the lower the better). Each plot represents the filtering error of a single state of system (34). The vertical black line denotes the performance of the **Nominal KF** approach. The numbers close the plot border denote the number of outliers.

the **Indirect KF** has, in these simulations , the advantage of knowing the diagonality of the **Q** matrix and the setting of the observer gain **L** which is close to the optimal Kalman one. Both these aspects are unknown in practical scenarios: nonetheless, its results are worse than the direct filter ones, that does not make use of such information.

6. CONCLUSIONS

This paper presented a comparison of traditional indirect (two-steps) and direct (one-step) data-driven approaches for the state estimation of linear stochastic systems. In particular, is has been shown how the uncertainty in the estimation of the noise covariance matrices impacts the most on the filtering performance, even if prior knowledge about such matrices is leveraged. If a set of state measurements is available, the design of a direct filter can be beneficial with respect to the two-stage designs. Future work is devoted to study the application of the direct filter to MIMO systems and the choice of the appropriate model structure for the direct filter, and applications to real datasets.

REFERENCES

Anderson, B.D. and Moore, J.B. (1979). *Optimal filtering*. Prentice-Hall: Englewood Cliffs, NJ.

ApS, M. (2022). *The MOSEK optimization toolbox for MATLAB manual. Version 10.0.*

Boni, P., Mazzoleni, M., and Previdi, F. (2024). Robust data-driven design of a jamming detection filter for airborne electromechanical actuators. *European Journal of Control*, 75, 100926. doi:10.1016/j.ejcon.2023.100926.

Ding, S. (2013). *Model-based fault diagnosis techniques: design schemes, algorithms, and tools, 2nd ed.* Springer Science & Business Media. doi:10.1007/978-1-4471-4799-2.

Duník, J., Straka, O., Kost, O., and Havlík, J. (2017a). Noise covariance matrices in state-space models: A survey and comparison of estimation methods-part i. *International Journal of Adaptive Control and Signal Processing*, 31(11), 1505–1543. doi:10.1002/acs.2783.

Duník, J., Straka, O., and Šimandl, M. (2017b). On autocovariance least-squares method for noise covariance matrices estimation. *IEEE Transactions on Automatic Control*, 62(2), 967–972. doi:10.1109/TAC.2016.2571899.

Kost, O., Duník, J., and Straka, O. (2021). Identifiability of unique elements of noise covariances in state-space models. *IFAC-PapersOnLine*, 54(7), 316–321. doi:10.1016/j.ifacol.2021.08.378. 19th IFAC Symposium on System Identification SYSID 2021.

Löfberg, J. (2004). Yalmip : A toolbox for modeling and optimization in matlab. In *In Proceedings of the CACSD Conference*. Taipei, Taiwan.

Mazzoleni, M., Valceschini, N., and Previdi, F. (2023). Model uncertainty-aware residual generators for siso lti systems based on kernel identification and randomized approaches. In *Proceedings of the 62nd IEEE Conference on Decision and Control*.

Milanese, M., Ruiz, F., and Taragna, M. (2010). Direct data-driven filter design for uncertain lti systems with bounded noise. *Automatica*, 46(11), 1773–1784. doi:10.1016/j.automatica.2010.07.006.

Novara, C., Milanese, M., Bitar, E., and Poolla, K. (2012). The filter design from data (fd2) problem: parametric-statistical approach. *International Journal of Robust and Nonlinear Control*, 22(16), 1853–1872. doi:10.1002/rnc.1791.

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. doi:10.1016/j.automatica.2005.09.006.

Rajamani, M.R. and Rawlings, J.B. (2009). Estimation of the disturbance structure from data using semidefinite programming and optimal weighting. *Automatica*, 45(1), 142–148. doi:10.1016/j.automatica.2008.05.032.

Ruiz, F., Novara, C., and Milanese, M. (2010). Direct design from data of optimal filters for lpv systems. *Systems & Control Letters*, 59(1), 1–8. doi:10.1016/j.sysconle.2009.10.008.

Shaked, U. and Theodor, Y. (1992). H/sub infinity /-optimal estimation: a tutorial. In *Proceedings of the 31st IEEE Conference on Decision and Control*, 2278–2286 vol.2. doi:10.1109/CDC.1992.371384.

Voulgaris, P. (1995). On optimal ℓ_∞ to ℓ_∞ filtering. *Automatica*, 31(3), 489–495. doi:10.1016/0005-1098(94)00123-Z.