Proximal-based recursive implementation for model-free data-driven fault diagnosis

Noom, Jacques; Soloviev, Oleg; Verhaegen, Michel

DOI
10.1016/j.automatica.2024.111656

Publication date
2024

Document Version
Final published version

Published in
Automatica

Automatica 165 (2024) 111656

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Proximal-based recursive implementation for model-free data-driven fault diagnosis☆

Jacques Noom [a,*], Oleg Soloviev [a,b], Michel Verhaegen [a]

[a] Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
[b] Flexible Optical BV, Polakweg 10–11, 2288 GG Rijswijk, The Netherlands

ARTICLE INFO

Article history:
Received 15 August 2023
Received in revised form 22 November 2023
Accepted 11 February 2024
Available online xxxx

Keywords:
Fault detection and isolation
System identification

ABSTRACT

We present a novel problem formulation for model-free data-driven fault diagnosis, in which possible faults are diagnosed simultaneously to identifying the linear time-invariant system. This problem is practically relevant for systems whose model cannot be identified reliably prior to diagnosing possible faults, for instance when operating conditions change over time, when a fault is already present before system identification is carried out, or when the system dynamics change due to the presence of the fault. A computationally attractive solution is proposed by solving the problem using unconstrained convex optimization, where the objective function consists of three terms of which two are non-differentiable. An additional recursive implementation based on a proximal algorithm is presented in order to solve the optimization problem online. The numerical results on a buck converter show the application of the proposed solution both offline and online.

1. Introduction

With the increase of complexity of automated systems, timely and accurate fault diagnosis is essential for preventing catastrophic failures. Accordingly, fault detection and identification has recently been considered within the top-three of control technologies with high future impact in industrial applications (Samad et al., 2020). Whereas model-based and signal-based methods (Gao, Cecati, & Ding, 2015a) require human expertise on modeling of the specific system or designing its characteristic signal shapes, knowledge-based methods rely on identifying the system and the possible faults from past data (Gao, Cecati, & Ding, 2015b; Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003). This makes knowledge-based fault diagnosis attractive specifically for large-scale industrial systems for which modeling is burdensome.

For its use of large amounts of historical data, knowledge-based fault diagnosis is often referred to as *data-driven* (Dai & Gao, 2013; Ding, 2014; Gao et al., 2015b; Simani, 2021; Yin, Ding, Xie, & Luo, 2014). However, it can effectively be partitioned in a (data-based) model acquisition phase and a model-based fault diagnosis phase. Especially in the first phase reliable and often also labeled data is required. In complex industrial applications this is not always available, or it is desired to diagnose faults directly from the first operational run of the unidentified system. Moreover, the main limitation of existing fault diagnosis techniques is that predetermined features or predetermined models lack the capability of accommodating for changing input/output dynamics for example due to changes in the internal system dynamics or in the environment.

The main contribution of this manuscript is threefold. First, a formulation for the problem of *model-free data-driven fault diagnosis* is presented. Different from existing categories for fault diagnosis which assume separate time periods for system modeling/identification and fault diagnosis, this novel formulation includes the goal of both retrieving the system dynamics and diagnosing the faults *simultaneously*. With a fixed data window, the diagnosis involves both the determination of the active faults from a set of hypothesized faults (fault isolation) and of their corresponding sizes (fault identification). The system dynamics are assumed to be Linear and Time-Invariant (LTI) over the considered data window. The proposed problem differs from the one formulated in Chen (2017), where only the presence of a fault in an unknown LTI system is to be detected. Instead, model-free data-driven fault diagnosis focuses on fault isolation and identification, simultaneously to retrieving an up-to-date model of the system.

The second contribution of this manuscript proposes to use our earlier developed solution (Noom, Soloviev, & Verhaegen, 2023) to the problem of model-free data-driven fault diagnosis.

☆ The material in this paper was partially presented at the 22nd IFAC World Congress (IFAC 2023), July 9–14, 2023, Yokohama, Japan. This paper was recommended for publication in revised form by Associate Editor Simone Formentin under the direction of Editor Alessandro Chiuso.

* Corresponding author.
E-mail addresses: j.noom@tudelft.nl (J. Noom), o.a.soloviev@tudelft.nl (O. Soloviev), m.verhaegen@tudelft.nl (M. Verhaegen).

https://doi.org/10.1016/j.automatica.2024.111656

By reformulating the problem as a convex optimization problem, the proposed solution is computationally attractive. As in Zhang (2021) a dictionary of hypothesized faults is constructed, after which sparsity is employed by the natural assumption that only a few of the hypothesized faults are concurrently active. The model-based approach in Zhang (2021) however assumes the availability of a predetermined model of the system. In the proposed solution for model-free data-driven fault diagnosis such a predetermined model is not required, but is identified simultaneously to diagnosing the faults. The simultaneous goal is achieved using results from the field of blind system identification (Scobee et al., 2015), which aims for identifying a system with unknown inputs. Different from Scobee et al. (2015), the proposed solution in this manuscript considers multiple non-differentiable optimization terms in order to apply it to model-free data-driven fault diagnosis. Hereby this contribution establishes a link between blind system identification and model-free data-driven fault diagnosis. In addition to our earlier work (Noom et al., 2023), conditions are introduced on the identifiability and diagnosability of the system and faults, respectively.

The third contribution of this manuscript is to propose a fast recursive implementation for solving the convex but non-differentiable optimization problem, enabling online monitoring including fault detection, isolation, fault identification and simultaneous system identification. The recursive implementation allows the unknown system dynamics and active faults to change over time, while being identified and diagnosed in real-time. Whereas the recursive approaches in Ding (2014) are limited to fault detection only and require an initial model of sufficient quality, our approach is fully model-free and is able to isolate and identify the faults in addition to detection. The proposed recursive implementation of model-free data-driven fault diagnosis relies on the proximal operators (Combettes & Pesquet, 2011; Parikh & Boyd, 2014) of the objective terms, of which closed-form solutions are available. Using efficient updates of the proximal operators, an established proximal algorithm (Combettes & Pesquet, 2008) is implemented recursively. Different from recursive implementation of subgradient methods such as in Angelosante and Giannakis (2009), the proposed proximal-based implementation does not involve fragile restrictions on tuning parameters for guaranteeing convergence. Other recursive implementations of proximal algorithms (Ajalloeian, Simonetto, & Dall'Anese, 2020; Dixit, Bedi, Tripathi, & Rajawat, 2019) are only able to cope with one non-differentiable objective term (or multiple terms only if the problem is block-separable), whereas the proposed approach can handle multiple non-differentiable objective terms even if the problem is not block-separable.

The manuscript is organized as follows. Section 2 presents the novel problem formulation. Section 3 presents the methodology for model-free data-driven fault diagnosis, starting with Section 3.1 introducing the structured data matrices and Section 3.2 recapping the method of Zhang (2021) for model-based fault diagnosis, neglecting the effect of the initial state. This negligence is based on developments in subspace identification (Chiuso, 2007; Verhaegen & Verdult, 2007). The proposed data-driven approach to fault diagnosis is presented in Section 3.3 with conditions on identifiability and diagnosability in Section 3.4. For the resulting convex optimization problem, Section 4 demonstrates the adoption of a proximal algorithm. Subsequently, Section 5 shows how the proximal algorithm can be implemented recursively in order to achieve online monitoring. The proposed methodology is tested numerically on a buck converter electronic circuit in Section 6 and conclusions are drawn in Section 7.

2. Problem formulation

Consider the following linear time-invariant system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Fd(k) + w(k) \\ y(k) &= Cx(k) + v(k) \end{aligned} \tag{1}$$

with $x(k) \in \mathbb{R}^{n_x}$, $u(k) \in \mathbb{R}^{n_u}$, $d(k) \in \mathbb{R}^{n_d}$ and $y(k) \in \mathbb{R}^{n_y}$ the state, input, fault signal and output; A, B, C and F the state-space matrices; and $w(k)$ and $v(k)$ the process and measurement noise, respectively. The fault signal

$$d(k) = \theta(k) z \tag{2}$$

is constructed from a known dictionary $\theta(k) \in \mathbb{R}^{n_d \times n_z}$ consisting of possible fault signal shapes and an unknown weighing vector $z = [z_1, \ldots, z_{n_z}]^\top \in \mathbb{R}^{n_z}$ which determines the active faults and their severity. Typically only a few faults out of the set of possible faults are active simultaneously.

With $\Phi = A - KC$, consider the observer (Verhaegen & Verdult, 2007) for system (1):

$$\begin{aligned} \hat{x}(k+1) &= \Phi \hat{x}(k) + Bu(k) + F\theta(k)z + Ky(k) \\ \hat{y}(k) &= C\hat{x}(k) \end{aligned} \tag{3}$$

with estimated state $\hat{x}(k)$ and estimated output $\hat{y}(k)$. Using this model, we can write the estimated output $\hat{y}(k)$ as:

$$\begin{aligned} \hat{y}(k) &= C\Phi^s \hat{x}(k-s) \\ &+ \sum_{i=1}^{s} C\Phi^{i-1} \Big(Bu(k-i) + F\theta(k-i)z + Ky(k-i) \Big) \end{aligned} \tag{4}$$

If the system is detectable and K is designed such that Φ is asymptotically stable, the effect of the state $\hat{x}(k-s)$ decreases to zero for increasing s. This leads to the following approximate Vector Auto-Regressive model with eXogenous input (VARX):

$$\hat{y}(k) \approx \sum_{i=1}^{s} B_i u(k-i) + F_i \theta(k-i) z + K_i y(k-i). \tag{5}$$

The matrices B_i, F_i and K_i of compatible size refer to the so-called observer Markov parameters (Phan & Longman, 1996). The VARX model description covers a wide range of multiple-input multiple-output (MIMO) systems and is studied comprehensively in Chiuso (2007), Lütkepohl (2005). It is identifiable in the sense that every unique set of VARX parameters leads to a unique output given that the input is persistently exciting (Ljung, 1999).

First we formulate the problem for model-based Fault Diagnosis (FD).

Problem 1 (*Model-Based FD*)**.** Given the VARX parameters B_i, F_i and K_i in (5), input $u(k)$ and output $y(k)$ sequences and the corresponding fault dictionary $\theta(k)$, isolate the faults as the nonzero elements in the unknown vector z, together with their corresponding sizes.

The problem for model-based FD within this formulation has been widely studied both for time-invariant and time-varying systems (see *e.g.* Basseville and Nikiforov (1993), Blanke, Kinnaert, Lunze, and Staroswieck (2006), Ding (2013)). Alternatively, knowledge-based (also called data-driven) FD consists of an extra preceding phase, formulated as follows.

Problem 2 (*Data-Driven FD*)**.** Given the input $u(k)$ and output $y(k)$ sequences and the corresponding fault dictionary $\theta(k)$ with the true weights z, first identify the system characteristics in a training phase with known faults. Afterwards, diagnosis can be performed in a subsequent phase where the faults are unknown. Based on the identified system characteristics and given the input $u(k)$ and output $y(k)$ sequences and the corresponding fault

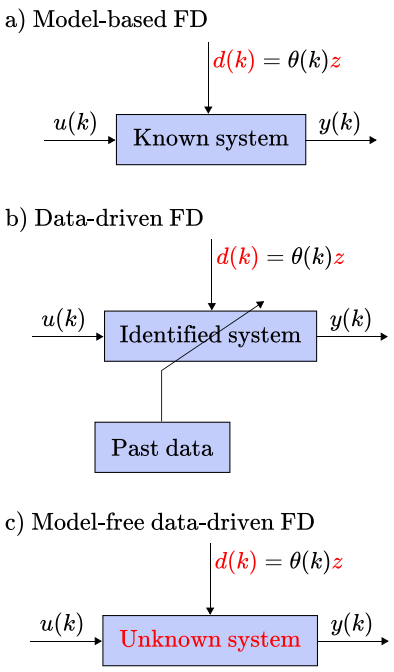

Fig. 1. (a) Model-based, (b) data-driven and (c) model-free data-driven FD.

dictionary $\theta(k)$, isolate the faults as the nonzero elements in the unknown vector z, together with their corresponding sizes.

Some well-known approaches for solving the data-driven FD problem for dynamical systems are summarized in Ding (2014), Gao et al. (2015b), Yin et al. (2014). These approaches require separate time periods for system identification prior to the fault diagnosis experiment. However, the system dynamics may already have changed in between the system identification and the fault diagnosis experiment, or for some applications it is not even possible to perform system identification due to costly data acquisition prior to productive system operation. In such cases it is desirable to perform fault diagnosis without requiring a previously identified model. This problem is formulated below.

Problem 3 (*Model-Free Data-Driven FD*)**.** Given the input $u(k)$ and output $y(k)$ sequences and the corresponding fault dictionary $\theta(k)$, simultaneously identify the system characteristics and isolate the faults as the nonzero elements in the unknown vector z, together with their corresponding sizes.

Problems 1–3 are visualized in Fig. 1. Naturally, the problem for model-free data-driven FD imposes some assumptions on system observability, the input sequence and the fault dictionary. The assumptions on input sequence and the fault dictionary will be stipulated in Section 3.4.

In the methodology we will make use of the following matrix norm notation. The Frobenius norm is defined as $\|X\|_F = \sqrt{\sum_{i,j} |x_{ij}|^2}$, the (1, 1)-norm as $\|X\|_{1,1} = \sum_{i,j} |x_{ij}|$ and the nuclear norm as $\|X\|_* = \sum_i \sigma_i(X)$, where x_{ij} are the (i, j)th elements of the matrix X, and $\sigma_i(X)$ the ith singular value.

3. Model-free data-driven fault diagnosis

3.1. VARX model identification

To introduce the structured data matrices, we first consider the fault-free identification problem with $d(k) = 0$. It should be noted that this fault-free identification step is not required for the execution of the proposed approach introduced later, however essential for building up the relevant knowledge.

Regard the available information

$$\begin{bmatrix} u(k) & u(k+1) & \dots & u(k+N-1) \end{bmatrix},$$
$$\begin{bmatrix} y(k) & y(k+1) & \dots & y(k+N) \end{bmatrix}.$$

Then with

$$Y = \begin{bmatrix} y^\top(k+s) \\ y^\top(k+s+1) \\ \vdots \\ y^\top(k+N) \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} B_1^\top \\ B_2^\top \\ \vdots \\ B_s^\top \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} K_1^\top \\ K_2^\top \\ \vdots \\ K_s^\top \end{bmatrix}, \tag{6}$$

and the Toeplitz matrices

$$T_u = \begin{bmatrix} u^\top(k+s-1) & u^\top(k+s-2) & \dots & u^\top(k) \\ u^\top(k+s) & u^\top(k+s-1) & \dots & u^\top(k+1) \\ \vdots & \vdots & \ddots & \vdots \\ u^\top(k+N-1) & u^\top(k+N-2) & \dots & u^\top(k+N-s) \end{bmatrix}$$
$$T_y = \begin{bmatrix} y^\top(k+s-1) & y^\top(k+s-2) & \dots & y^\top(k) \\ y^\top(k+s) & y^\top(k+s-1) & \dots & y^\top(k+1) \\ \vdots & \vdots & \ddots & \vdots \\ y^\top(k+N-1) & y^\top(k+N-2) & \dots & y^\top(k+N-s) \end{bmatrix} \tag{7}$$

the following least-squares problem aims at finding the system parameters $\mathbf{B}$ and $\mathbf{K}$ for the 1-step ahead predictor:

$$\min_{\mathbf{B},\mathbf{K}} \left\| Y - \begin{bmatrix} T_u & T_y \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{K} \end{bmatrix} \right\|_F^2. \tag{8}$$

The solution to this problem is unique if the matrix $\begin{bmatrix} T_u & T_y \end{bmatrix}$ has full column rank. This condition requires the input to be persistently exciting (Verhaegen & Verdult, 2007) and leads to identifiability of the system as in Definition 1.

3.2. Model-based fault diagnosis under sparseness assumption

Suppose now that the fault signal $d(k)$ is nonzero and composed as in (2), and the system matrices A, B, C, F and K are known. Whereas Zhang (2021) indicates for conventional approaches that "tractable solutions are only available when a small number of possible faults are assumed", recently he proposed a computationally efficient method to diagnose from a large set of possible faults. Neglecting the effects of initial condition (given the fact that Φ is asymptotically stable (Chiuso, 2007; Verhaegen & Verdult, 2007)) the approach in Zhang (2021) can be (accurately) approximated as follows.

Let $\mathbf{F}$ be constructed from F similarly to $\mathbf{B}$ from B in (6). Consider the Kronecker product

$$\mathbf{F}(z) = \mathbf{F} \otimes z \tag{9}$$

and the Toeplitz matrix T_θ constructed as T_u in (7) with all elements $u^\top(k)$ replaced by $\mathrm{vec}(\theta^\top(k))^\top$, such that the VARX approximation becomes

$$Y \approx \begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{F}(z) \\ \mathbf{K} \end{bmatrix}. \tag{10}$$

With the sparseness assumption on z entailing only a small number of faults is active simultaneously, this results in a lasso optimization problem:

$$\min_z \left\| Y - \begin{bmatrix} T_u & T_y \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{K} \end{bmatrix} - T_\theta \mathbf{F}(z) \right\|_F^2 + \lambda \|z\|_1 \tag{11}$$

with $\mathbf{F}(z)$ given in (9). In words, the output residual is minimized over the fault weighing variables z by subtracting their corresponding fault responses $T_\theta \mathbf{F}(z)$ from the data equation. The 1-norm is included to enhance sparsity on the weighing vector. This approach has shown good performance on both time-invariant and time-variant systems by Zhang (2021), where also the (negligible) effect of the initial condition is taken into consideration.

3.3. Model-free data-driven approach to fault diagnosis

In our case of data-driven fault diagnosis the system matrices $\mathbf{B}$, $\mathbf{K}$ and $\mathbf{F}$ are unknown in addition to the fault(s). This implies (11) becomes a bilinear optimization problem, which is computationally expensive due to its nonconvexity. However, from the definition of $\mathbf{F}(z)$ it is possible to compose the matrix $\mathbf{F}^*(z)$ as

$$\mathbf{F}^*(z) := \operatorname{vec}\left(\mathbf{F}^\top\right) z^\top \tag{12}$$

which has rank one (Scobee et al., 2015). Besides, the variable $\mathbf{F}(z)$ has by construction in (9) a degree of sparsity (defined as the ratio of nonzero components) equal to that of z. As a result, the bilinear optimization problem can be replaced by the rank-constrained minimization problem

$$\begin{aligned} &\min_{\mathbf{B},\mathbf{F}(z),\mathbf{K}} \left\| Y - \begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{F}(z) \\ \mathbf{K} \end{bmatrix} \right\|_F^2 + \lambda \|\mathbf{F}(z)\|_{1,1} \\ &\text{s.t. } \operatorname{rank}\left(\mathbf{F}^*(z)\right) = 1. \end{aligned} \tag{13}$$

Note that in contrast to (11) where z is the optimization variable, in (13) the quantity $\mathbf{F}(z)$ is an explicit optimization variable. The solution of $\mathbf{F}$ and z can be found up to a multiplicative scalar from singular value decomposition (SVD) of $\mathbf{F}^*(z)$. Since no model knowledge is assumed, the optimization problem relies on minimizing the output residuals, now with respect to both model parameters as well as fault parameters.

Problem (13) can be relaxed to a convex optimization problem by replacing the rank constraint with an additive weighted nuclear norm to the objective function. The eventual unconstrained convex optimization problem is then

$$\begin{aligned} \min_{\mathbf{B},\mathbf{F}(z),\mathbf{K}} &\left\| Y - \begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{F}(z) \\ \mathbf{K} \end{bmatrix} \right\|_F^2 + \tau \left\| \mathbf{F}^*(z) \right\|_* \\ &+ \lambda \|\mathbf{F}(z)\|_{1,1}. \end{aligned} \tag{14}$$

The faults can be isolated using (14) only, however their magnitudes will be biased toward zero due to the additional penalties to the least-squares term. Also, the identified VARX matrices $\mathbf{B}$ and $\mathbf{K}$ may be affected by the bias in $\mathbf{F}(z)$. For refined estimation of the fault magnitudes and system parameters, a second optimization without the 1-norm can be performed over the nonzero estimated elements $\hat{z}$ found in (14). In practice, this means that the components of the dictionary $\theta(k)$ and the weighing variables z in (2) are in the second optimization neglected according to the 'zero' (in practice below a threshold) elements of $\hat{z}$ found in the first optimization, and in the second optimization $\lambda = 0$.

The choice of the tuning parameters may be nontrivial. However, it can be deduced that a rank one solution is encouraged by increasing τ, and the sparsity of z by increasing λ. A possible tuning strategy is to set λ to zero first and tune τ such that the predictor performance of (5) (for instance calculated as Variance Accounted For (VAF)) with parameters found in (14) is optimized on a validation data set. While fixing τ to the value found in the first step, λ can be adapted gradually by optimizing the performance of (5) on validation data, with parameters found in (14) after refinement.

3.4. Identifiability & diagnosability

The model-free data-driven approach to fault diagnosis requires both the system to be identifiable and the faults to be diagnosable.

Definition 1 (*Identifiability of a System*)**.** A system is regarded to be identifiable if there exists an input sequence such that the variables $\mathbf{B}$ and $\mathbf{K}$ in (14) can be determined uniquely.

A sufficient condition for the solution to $\mathbf{B}$, $\mathbf{K}$ and also $\mathbf{F}(z)$ in (14) to be unique is that the matrix $\begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix}$ has full column rank. This is now a condition on the persistency of excitation of the joint input $(u^\top(k), \operatorname{vec}(\theta^\top(k))^\top)$. However, due to the regularization terms, full column rank of $\begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix}$ is not a necessary condition for uniqueness of the solution to (14). Moreover, with $\begin{bmatrix} T_u & T_y \end{bmatrix}$ full column rank and $\tau = 0$, the solution may still be unique (Ali & Tibshirani, 2019). The fact that $\tau > 0$ in (14) further increases the probability of a unique solution. It must be noted that a unique solution does not guarantee diagnosability of each fault, and vice versa.

Definition 2 (*Diagnosability of a Fault*)**.** Given an input/output sequence $u(k)$ and $y(k)$ and a fault dictionary $\theta(k)$, a fault $z_j \neq 0$ is regarded to be diagnosable if all possible solutions to (14) satisfy $z_j \neq 0$.

The following lemma states a condition necessary for a fault to be diagnosable.

Lemma 1 (*Necessary Condition for Diagnosability*)**.** *Consider a fault z_j with its corresponding dictionary signal $\theta_j(k)$. In order for the fault z_j to be diagnosable, it is necessary that at least one column of its corresponding dictionary signal*

$$\begin{bmatrix} \operatorname{vec}(\theta_j^\top(k))^\top \\ \operatorname{vec}(\theta_j^\top(k+1))^\top \\ \vdots \\ \operatorname{vec}(\theta_j^\top(k+N-1))^\top \end{bmatrix}$$

is linearly independent of the columns of

$$\begin{bmatrix} u^\top(k) & y^\top(k) \\ u^\top(k+1) & y^\top(k+1) \\ \vdots & \vdots \\ u^\top(k+N-1) & y^\top(k+N-1) \end{bmatrix}.$$

Proof. In the case of linear dependence, the dictionary signal can be written as

$$\operatorname{vec}(\theta_j^\top(k)) = L_u u(k) + L_y y(k) \tag{15}$$

with L_u and L_y time-invariant matrices representing the linear dependence. Then,

$$\begin{aligned} &\left\| Y - \begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{F}(z) \\ \mathbf{K} \end{bmatrix} \right\|_F^2 = \\ &\left\| Y - \begin{bmatrix} T_u & T_\theta & T_y \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{B}} \\ \tilde{\mathbf{F}}(z) \\ \tilde{\mathbf{K}} \end{bmatrix} \right\|_F^2 \end{aligned} \tag{16}$$

where

$$\tilde{\mathbf{B}} = \mathbf{B} + \begin{bmatrix} L_u^\top F_1^\top \\ \vdots \\ L_u^\top F_s^\top \end{bmatrix} z_j, \qquad \tilde{\mathbf{F}}(z) = \mathbf{F} \otimes \begin{bmatrix} z_1 \\ \vdots \\ \tilde{z}_j = 0 \\ \vdots \\ z_{n_z} \end{bmatrix}$$

$$\tilde{\mathbf{K}} = \mathbf{K} + \begin{bmatrix} L_y^\top F_1^\top \\ \vdots \\ L_y^\top F_s^\top \end{bmatrix} z_j,$$

and $(\mathbf{B}, \mathbf{F}(z), \mathbf{K})$ the actual situation with $z_j \neq 0$. With optimization problem (14) regularizing on $\mathbf{F}(z)$ but not on $\mathbf{B}$ and $\mathbf{K}$, the global optimum is achieved with $\tilde{z}_j = 0$. Following Definition 2, this fault is not diagnosable. □

Lemma 1 stresses the importance of designing a healthy combination of the fault dictionary, system inputs and outputs. As long as it satisfies Lemma 1, possible dictionary signals include sinusoidal, triangular or square waveforms with various frequencies, random Fourier expansions as in Bliek, Verstraete, Verhaegen, and Wahls (2018), unit steps with various starting points, user-defined fault progressions and nonlinear relations of $u(k)$ and/or $y(k)$. Practical examples of hypothesized fault patterns include varying load resistances in a buck converter (as illustrated in Section 6) and blockage of air data sensors modeled as additive sinusoidal pressure changes (Freeman, Seiler, & Balas, 2013).

4. Adoption of proximal algorithm

The data-driven fault diagnosis problem can be recast to a batch optimization problem consisting of $m = 3$ convex but possibly non-differentiable functions f_i:

$$\min_{\mathbf{x}} f(\mathbf{x}) = \min_{\mathbf{x}} \sum_{i=1}^{m} f_i(\mathbf{x}) \tag{17}$$

$$= \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - H\mathbf{x}\|_2^2 + \tau\|\mathbf{x}_L\|_* + \lambda\|\mathbf{x}_S\|_1, \tag{18}$$

where $\mathbf{y} \in \mathbb{R}^{n_\mathbf{y}}$ are the vectorized measurements, $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$ the optimization variables with $\mathbf{x}_L$ and $\mathbf{x}_S$ constructed from $\mathbf{x}$ by selecting and rearranging its elements to a low-rank matrix and sparse vector, respectively. For the data-driven fault diagnosis problem, identical elements from $\mathbf{x}$ are selected for constructing both $\mathbf{x}_L$ and $\mathbf{x}_S$.

Problem (18) is convex but non-differentiable due to both the nuclear norm and the 1-norm. Therefore, conventional gradient-based algorithms provide limited convergence properties. Alternatively, proximal algorithms have recently shown their potential for solving large-scale non-smooth problems (Combettes & Pesquet, 2011; Parikh & Boyd, 2014).

A proximal algorithm uses proximal operators of the objective terms iteratively in order to solve a convex optimization problem (Parikh & Boyd, 2014). The proximal operator of a function g is defined as

$$\text{prox}_g(\mathbf{v}) = \arg\min_{\mathbf{x}} \left(g(\mathbf{x}) + \frac{1}{2}\|\mathbf{x} - \mathbf{v}\|_2^2 \right). \tag{19}$$

For several specific functions g a closed-form expression for the proximal operator can be derived (Combettes & Pesquet, 2011), enabling fast evaluation.

For optimization problems involving a sum of two terms f_i, established proximal algorithms include forward–backward splitting (Combettes & Wajs, 2005), Douglas–Rachford splitting (Combettes & Pesquet, 2007) and FISTA (Beck & Teboulle, 2009). However, the data-driven fault diagnosis problem encompasses three terms of which two are non-differentiable, which in general cannot be handled efficiently by these algorithms. Alternatively, minimization of (18) is enabled by multiple-operator splitting schemes, such as the Parallel ProXimal Algorithm (PPXA) (Combettes & Pesquet, 2008), generalized forward–backward splitting (Raguet, Fadili, & Peyré, 2013) or the Davis–Yin algorithm (Davis & Yin, 2017). For the small number of tuning parameters, and the convergence being robust against the choice of these tuning parameters, we select PPXA for solving the data-driven fault diagnosis problem. It is reproduced in Algorithm 1.

Algorithm 1 Parallel ProXimal Algorithm (PPXA) (Combettes & Pesquet, 2008) for solving (17).

Initialization
- $0 < \rho < \infty$ ▷ Scalar step size
- $0 < \omega = [\omega_1, \ldots, \omega_m] \leq 1$ satisfying $\sum_{i=1}^m \omega_i = 1$
- $\Gamma = [\gamma_1, \ldots, \gamma_m] = \rho\omega$
- $\mathbf{x}_0$ ▷ Initial condition

function PPXA($\Gamma, \omega, \mathbf{x}_0$)
- $v_0 : (v_{i,0})_{1\leq i\leq m} = \mathbf{x}_0$ ▷ Auxiliary variables
- **for** $j = 0, 1, \ldots, n_p - 1$ **do** ▷ For n_p iterations
  - **for** $i = 1, \ldots, m$ **do**
    - $p_{i,j} = \text{prox}_{\gamma_i f_i}(v_{i,j}) + \varepsilon_{i,j}$ ▷ Error $\varepsilon_{i,j}$
  - **end for**
  - $p_j = \sum_{i=1}^m \omega_i p_{i,j}$
  - $\xi_j \in]0, 2[$ ▷ Tuning parameter
  - **for** $i = 1, \ldots, m$ **do**
    - $v_{i,j+1} = v_{i,j} + \xi_j(2p_j - \mathbf{x}_j - p_{i,j})$
  - **end for**
  - $\mathbf{x}_{j+1} = \mathbf{x}_j + \xi_j(p_j - \mathbf{x}_j)$
- **end for**

end function

In words, PPXA evaluates the proximal operator for the individual terms f_i in parallel, after which the outcomes are averaged and employed in next iteration. Convergence is ensured under the following conditions (Combettes & Pesquet, 2008):

- $\lim_{\|\mathbf{x}\|\to\infty} f(\mathbf{x}) = +\infty$
- $\cap_{i=1}^m \text{ri dom} f_i \neq \emptyset$ (the intersection of the relative interiors of the domains of f_i is nonempty)
- $\lim_{n_p\to\infty} \sum_{j=0}^{n_p} \xi_j(2 - \xi_j) = +\infty$
- $\forall i \in \{1, \ldots, m\} \lim_{n_p\to\infty} \sum_{j=0}^{n_p} \xi_j \|\varepsilon_{i,j}\| < +\infty$ (the possible error $\varepsilon_{i,j}$ in the computation of the ith proximal operator decreases to zero)

For the data-driven fault diagnosis problem (18) the first two conditions are naturally satisfied and the third condition by the straightforward choice $\xi_j = 1\ \forall j$. Assuming the ability for precise evaluation of the proximal operators (for instance by closed-form computation) and by assigning trivial values to the weights $\omega_i = 1/m$, the only remaining tuning parameter γ affects the speed of convergence. A small step size γ will lead to slow initial convergence, whereas a large step size decelerates convergence close to the optimum.

With the soft-thresholding operator defined entrywise as

$$[\mathcal{S}_\gamma(\mathbf{v})]_n = \text{sign}(\mathbf{v}_n)[|\mathbf{v}_n| - \gamma]_+ \tag{20}$$

closed-form expressions for the proximal operators of the objective terms in optimization problem (18) are presented in Table 1.

The evaluation of $\text{prox}_{\gamma f_1}$ is relatively computationally expensive due to the inverse $(\gamma H^\top H + I)^{-1}$, which would require $O(n_\mathbf{x}^3)$ flops for each iteration of the proximal algorithm. However, given that PPXA allows some small errors $\varepsilon_{i,j}$ in the calculation of the proximal operators, there are several possibilities to ease the evaluation of $\text{prox}_{\gamma f_1}(\mathbf{v})$. Some of these possibilities are specified in the next section.

Table 1
Proximal operators for objective terms in (18) with step size γ.

Function $g(\mathbf{x})$	Proximal operator $\text{prox}_{\gamma g}(\mathbf{v})$
$f_1(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - H\mathbf{x}\|_2^2$	$(\gamma H^\top H + I)^{-1}(\mathbf{v} + \gamma H^\top \mathbf{y})$
$f_2(\mathbf{x}) = \tau\|\mathbf{x}_L\|_*$ (Cai, Candès, & Shen, 2010)	$U\mathcal{S}_{\gamma\tau}(\Sigma)V^\top$ with SVD $\mathbf{v}_L = U\Sigma V^\top$
$f_3(\mathbf{x}) = \lambda\|\mathbf{x}_S\|_1$	$\mathcal{S}_{\gamma\lambda}(\mathbf{v}_S)$

4.1. Opportunities for accelerated implementation

Since the inverse $(\gamma H^\top H + I)^{-1}$ and $\gamma H^\top \mathbf{y}$ remain invariable over the iterations of the proximal algorithm for the batch problem (18), a first option is to cache and reuse it in subsequent iterations of the proximal algorithm. This would require $O(n_{\mathbf{x}}^3)$ flops for the first iteration, however only $O(n_{\mathbf{x}}^2)$ flops for subsequent iterations, without introducing approximation errors.

A second option is to replace the proximal operator with a step in the direction of the negative gradient, as elaborated in the following lemma.

Lemma 2. *If γ is small enough such that*

$$\lim_{n\to\infty}(-\gamma H^\top H)^n = 0, \tag{21}$$

the proximal operator for γf_1 can be approximated by a step in the direction of its negative gradient:

$$\text{prox}_{\gamma f_1}(\mathbf{v}) \approx \mathbf{v} - \gamma H^\top H\mathbf{v} + \gamma H^\top \mathbf{y} = \mathbf{v} - \gamma\nabla f_1(\mathbf{v}). \tag{22}$$

Proof. Under condition (21), by the Neumann series:

$$\begin{aligned}(\gamma H^\top H + I)^{-1} &= \sum_{n=0}^{\infty}(-\gamma H^\top H)^n \\ &= I - \gamma H^\top H + \sum_{n=2}^{\infty}(-\gamma H^\top H)^n \\ &= I - \gamma H^\top H + e(\gamma, H)\end{aligned}$$

This approximation is substituted in the proximal operator for f_1 in Table 1:

$$\begin{aligned}\text{prox}_{\gamma f_1}(\mathbf{v}) = &\left(I - \gamma H^\top H + e(\gamma, H)\right)(\mathbf{v} + \gamma H^\top \mathbf{y}) \\ &= \mathbf{v} - \gamma(H^\top H\mathbf{v} - H^\top \mathbf{y}) - \gamma^2 H^\top HH^\top \mathbf{y} \\ &\quad + e(\gamma, H)(\mathbf{v} + \gamma H^\top \mathbf{y}) \\ &= \mathbf{v} - \gamma\nabla f_1(\mathbf{v}) - \gamma^2 H^\top HH^\top \mathbf{y} \\ &\quad + e(\gamma, H)(\mathbf{v} + \gamma H^\top \mathbf{y}),\end{aligned}$$

which concludes the proof. □

The gradient step in (22) reduces the computational complexity to $O(n_{\mathbf{x}}^2)$ flops for each iteration. This goes at the expense of trading off γ for small approximation errors and fast convergence while guaranteeing (21).

A third possibility is to approximate it with an iterative method such as conjugate gradient, provided with a warm start from a previous solution (Parikh & Boyd, 2014). The conjugate gradient method requires multiple iterations of $O(n_{\mathbf{x}}^2)$ flops in order to solve $\text{prox}_{\gamma f_1}(\mathbf{v})$. However, in contrast to approximation (22), the conjugate gradient method allows the tolerance (and thus the magnitude of the approximation error $\varepsilon_{i,j}$) to be predefined, providing more control on the convergence properties of Algorithm 1.

5. Proximal-based recursive implementation

Multiple variants are possible to solve (18) online in which each time step a new set of inputs and outputs become available. Let us consider (18) over an Infinite Window (IW), Finite moving Window (FW) or Exponentially Weighted (EW) window. These are all addressed by the following adapted cost function:

$$\hat{\mathbf{x}}_k = \arg\min_{\mathbf{x}} \frac{1}{2}\sum_{j=\ell}^{k}\beta_{k-j}\|y_j - H_j\mathbf{x}\|_2^2 + \tau\|\mathbf{x}_L\|_* + \lambda\|\mathbf{x}_S\|_1 \tag{23}$$

where $0 \le \beta_{k-j} \le 1$ is a forgetting factor and H_j is the regressor matrix corresponding to time instance j. The formulation (23) allows to make three different choices by the parameters ℓ and β_{k-j} as highlighted in Table 2. With the corresponding recursive definitions $\mathbf{R}_k$ and $\mathbf{r}_k$ in Table 2, the argument of optimization problem (23) can be found by

$$\begin{aligned}\hat{\mathbf{x}}_k &= \arg\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^\top\mathbf{R}_k\mathbf{x} - \mathbf{x}^\top\mathbf{r}_k + \tau\|\mathbf{x}_L\|_* + \lambda\|\mathbf{x}_S\|_1 \\ &= \arg\min_{\mathbf{x}} f_{r,k}(\mathbf{x}) + f_2(\mathbf{x}) + f_3(\mathbf{x}).\end{aligned} \tag{24}$$

Accordingly, the proximal operator for the least-squares term $f_{r,k}$ is

$$\text{prox}_{\gamma f_{r,k}}(\mathbf{v}) = (\gamma\mathbf{R}_k + I)^{-1}(\mathbf{v} + \gamma\mathbf{r}_k). \tag{25}$$

Together with the proximal operators for f_2 and f_3 in Table 1, PPXA in Algorithm 1 can be performed each time instance k with warm start $\mathbf{x}_0 = \hat{\mathbf{x}}_{k-1}$ in order to solve (24).

Section 4.1 provides three opportunities to accelerate the evaluation of the proximal operator for the least-squares term in the batch problem (18). These opportunities can be extended to the case when the cost function is recursive as in (24). The first opportunity of caching and reusing the inverse enables exact evaluation, however in the recursive problem the inverse $(\gamma\mathbf{R}_k + I)^{-1}$ should be updated each time step. This exact evaluation is handled in Section 5.1, whereas two approximations analogous to those in Section 4.1 are elaborated in Section 5.2.

5.1. Exact recursive evaluation of (25)

Depending on the window type, the evaluation of $\text{prox}_{\gamma f_{r,k}}$ can be accelerated using the matrix inversion lemma (for Infinite Window (IW) and Finite Window (FW)) or CG (for Exponentially Weighted (EW) window). First, we study the case of IW after which FW is studied in the following two lemmas, which are general results from recursive least-squares literature (Sayed, 2003).

Lemma 3. *In the case of IW, $\mathbf{P}_k = (\gamma\mathbf{R}_k + I)^{-1}$ can be updated recursively by*

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \gamma\mathbf{P}_{k-1}H_k^\top(I + \gamma H_k\mathbf{P}_{k-1}H_k^\top)^{-1}H_k\mathbf{P}_{k-1}. \tag{26}$$

Proof. The matrix inversion lemma is given by Verhaegen and Verdult (2007):

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}. \tag{27}$$

In the case of IW (see Table 2),

$$(\gamma\mathbf{R}_k + I)^{-1} = \left((\gamma\mathbf{R}_{k-1} + I) + \gamma H_k^\top H_k\right)^{-1}.$$

By substituting the matrices in (27) with

$$\begin{aligned}&A \leftarrow \mathbf{P}_{k-1}^{-1} = (\gamma\mathbf{R}_{k-1} + I), && B \leftarrow \gamma H_k^\top, \\ &C \leftarrow I, && D \leftarrow H_k\end{aligned}$$

one arrives at (26). □

Table 2
Basic instances of (23) by choices of ℓ and β_{k-j}, where $\mathcal{L}$ is a fixed window length and $0 \leq \beta < 1$.

Window type	ℓ	β_{k-j}	$\mathbf{R}_k$	$\mathbf{r}_k$
Infinite Window (IW)	1	1	$\mathbf{R}_{k-1} + H_k^\top H_k$	$\mathbf{r}_{k-1} + H_k^\top y_k$
Finite Window (FW)	$k - \mathcal{L}$	1	$\mathbf{R}_{k-1} + H_k^\top H_k - H_\ell^\top H_\ell$	$\mathbf{r}_{k-1} + H_k^\top y_k - H_\ell^\top y_\ell$
Exponentially Weighted (EW)	1	β^{k-j}	$\beta\mathbf{R}_{k-1} + H_k^\top H_k$	$\beta\mathbf{r}_{k-1} + H_k^\top y_k$

Lemma 4. *In the case of FW, $\mathbf{P}_k = (\gamma\mathbf{R}_k + I)^{-1}$ can be updated recursively by*

$$\begin{aligned}\mathbf{P}'_k &= \mathbf{P}_{k-1} - \gamma\mathbf{P}_{k-1}H_k^\top(I + \gamma H_k\mathbf{P}_{k-1}H_k^\top)^{-1}H_k\mathbf{P}_{k-1}\\ \mathbf{P}_k &= \mathbf{P}'_k - \gamma\mathbf{P}'_kH_\ell^\top(-I + \gamma H_\ell\mathbf{P}'_kH_\ell^\top)^{-1}H_\ell\mathbf{P}'_k\end{aligned} \tag{28}$$

Proof. By defining

$$\mathbf{P}'_k = (\gamma\mathbf{R}'_k + I)^{-1} = \left((\gamma\mathbf{R}_{k-1} + I) + \gamma H_k^\top H_k\right)^{-1},$$

the first line in (28) is obtained using Lemma 3.

In the case of FW (see Table 2),

$$\begin{aligned}(\gamma\mathbf{R}_k + I)^{-1} &= \left((\gamma\mathbf{R}_{k-1} + I + \gamma H_k^\top H_k) - \gamma H_\ell^\top H_\ell\right)^{-1}\\ &= \left((\gamma\mathbf{R}'_k + I) - \gamma H_\ell^\top H_\ell\right)^{-1}.\end{aligned}$$

By substituting the matrices in (27) with

$$\begin{aligned}A &\leftarrow (\mathbf{P}'_k)^{-1} = (\gamma\mathbf{R}'_k + I), & B &\leftarrow \gamma H_\ell^\top,\\ C &\leftarrow -I, & D &\leftarrow H_\ell\end{aligned}$$

one arrives at the second line in (28). □

The computational complexity for evaluating the proximal operator (25) using (26) or (28) is $O(n_{\mathbf{x}}^2)$ flops for each iteration of the proximal algorithm.

5.2. Approximate evaluation of (25)

An exact recursive update of $(\gamma\mathbf{R}_k + I)^{-1}$ is not for each window type available. For EW some possibilities for approximating this covariance matrix based on $(\gamma\mathbf{R}_{k-1} + I)^{-1}$ and $\gamma H_k^\top H_k$ are summarized in Gunnarsson (1996), in general at the cost of $O(n_{\mathbf{x}}^2)$ flops each time step k. However, such approximation induces the corresponding errors to propagate over time.

Following Lemma 2, the proximal operator in (25) can be approximated by a step in the direction of the negative gradient, for which evaluation of the matrix inversion is not required. Such gradient step involves a computational complexity of $O(n_{\mathbf{x}}^2)$ flops per iteration of the proximal algorithm. This is however at the expense of inducing an upper bound on γ, which in the recursive variant cannot be verified prior to the experiment.

An alternative to the gradient step as introduced in Section 4.1, is to approximate the proximal operator in (25) using the conjugate gradient method provided with a warm start. This requires multiple evaluations of $O(n_{\mathbf{x}}^2)$ flops each iteration of the proximal algorithm and allows the approximation tolerance to be predefined. The warm start could for instance be provided from a previous solution of the proximal algorithm at no computational cost.

5.3. Algorithm overview

The overview of the proximal-based implementation for model-free data driven fault diagnosis is presented in Algorithm 2. It should be noted that this scheme is applied straightforwardly to optimization problems in the form of (24), as well as recursive least-squares problems with convex regularization terms different from the nuclear norm and 1-norm, especially when closed-form expressions of their corresponding proximal operators are available. A list of functions with closed-form expressions for the proximal operators is for instance given in Combettes and Pesquet (2011).

Algorithm 2 Proximal-based implementation for solving (24) recursively.

Initialization
$0 < \rho < \infty$ ▷ Scalar step size
$0 < \omega = [\omega_1, \omega_2, \omega_3] \leq 1$ satisfying $\sum_{i=1}^{3}\omega_i = 1$
$\Gamma = [\gamma_1, \gamma_2, \gamma_3] = \rho\omega$
$\hat{\mathbf{x}}_{s-1}$ ▷ Initial condition
for $k = s, \ldots, N$ **do**
Obtain measurement y_k with regressor H_k
Update $\mathbf{R}_k$ and $\mathbf{r}_k$ in (25) using Table 2
If applicable update $\mathbf{P}_k = (\gamma_1\mathbf{R}_k + I)^{-1}$ using (26) or (28)
$\hat{\mathbf{x}}_k$ = PPXA($\Gamma, \omega, \hat{\mathbf{x}}_{k-1}$) (Algorithm 1) using the updated $\text{prox}_{\gamma_1 f_{r,k}}$ in (25), and from Table 1 $\text{prox}_{\gamma_2 f_2}$ and $\text{prox}_{\gamma_3 f_3}$
end for

In order to solve (24), Algorithm 2 preserves the convergence properties of PPXA (Combettes & Pesquet, 2008) summarized in Algorithm 1. The warm start from $\hat{\mathbf{x}}_{k-1}$ allows to perform only a small number of iterations n_p in PPXA, suppressing the computational effort per time instance k. The experimental demonstration of Algorithm 2 is presented in Section 6.3.

6. Numerical results

Consider the buck converter in Fig. 2 as example for a DC–DC switch mode power supply. In Al-Greer, Armstrong, Ahmeid, and Giaouris (2019) the issue is raised that contemporary system identification procedures for DC–DC switch mode power supplies cannot handle unknown rapidly varying load resistances. However, when a set of possible load profiles is available – for instance when multiple loads are connected but not necessarily activated – the proposed data-driven fault diagnosis procedure can be implemented in order to identify the system characteristics simultaneously to diagnosing the activated loads. The assumptions are as follows:

- The voltage difference over the load resistance $V_R(k)$ is measured as output variable, corrupted with noise $y(k) = V_R(k) + v(k)$.
- The duty ratio of the switch is regarded as the known input variable $u(k) = S_B(k)$. These are typical input/output variables (Al-Greer et al., 2019).
- The load resistance $R_B(k) = R_B + \Delta R_B(k)$ consists of a static part R_B and a time-varying part $\Delta R_B(k)$.
- For the time-varying load resistance $d(k) = \Delta R_B(k) = \theta(k)z$, a dictionary $\theta(k)$ of possible profiles (or: faults) is known, but not necessarily their magnitudes z. Only a small number of possible load resistance profiles from the dictionary is active.
- The variables L_B, C_B, V_B and R_B are unknown.

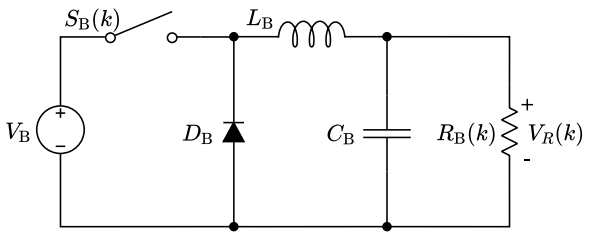

Fig. 2. Buck converter circuit with inductance L_B, diode D_B, capacitance C_B, switch with duty ratio $S_B(k)$, source voltage V_B and time-varying load resistance $R_B(k)$. The voltage difference over the load resistance $V_R(k)$ is measurable.

The simultaneous goals are:

- Identify the system (obtain an equivalent for the knowledge of L_B, C_B, V_B and R_B).
- Diagnose the time-varying load resistance $\Delta R_B(k)$.

In order to show that the dynamics of the buck converter can be captured in (5), an averaged and linearized model of the buck converter is presented in Appendix A.

The nominal numerical values are taken similar to those in Ahmeid, Armstrong, Gadoue, Al-Greer, and Missailidis (2017), namely $V_B = 10$ V, $L_B = 220$ μH, $C_B = 330$ μF and $R_B = 5$ Ω. The switching frequency is 200 kHz and the sampling rate 20 kHz. The initial voltage $V_R(0) = 4.5$ V and initial inductor current $i_L(0) = 0.9$ A. The measurement noise is distributed as $v(k) \sim \mathcal{N}(0, 10^{-3})$. The duty ratio of the switch $S_B(k)$ takes the values 0.4 and 0.5 following a pseudo-random binary signal. The time-varying load resistance $\Delta R_B(k)$ is built up from the dictionary $\theta(k) \in \mathbb{R}^{1\times 50}$ consisting of 50 square waves with linearly increasing frequencies between 1000 and 1900 Hz and amplitude 1 Ω. The fault parameter vector $z \in \mathbb{R}^{50}$ consists of zeros with randomly drawn entries set to one.

6.1. Data-driven fault diagnosis

The simulation experiment tests the proposed approach to model-free data-driven fault diagnosis in (14) with number of data points $N = 1200$, VARX order $s = 3$ and $\tau = \lambda = 10$. It is solved using PPXA in Algorithm 1 with $\gamma = 10^{-3}$ and $n_p = 10^4$.

Fig. 3 shows a typical realization of the simulated buck converter in the case of three active faults. In the bottom figure it can be seen that the entries of z corresponding to the activated square waves are diagnosed correctly up to a multiplicative scalar. After applying the refinement step as elaborated in Section 3.3, the estimates of **B**, **K**, **F** and z are verified with a Variance Accounted For (VAF) (Verhaegen & Verdult, 2007) of 98.8%.

6.2. Comparison to model-based fault diagnosis

The proposed model-free approach with parameters as in Section 6.1 is compared to model-based fault diagnosis using an error-free and erroneous linearized model. The corresponding linear VARX-models required for the model-based approach are constructed as described in Appendix A, after which an error of 0.4 is introduced in the $(2, 1)$th element of the discretized A-matrix. Then, the observer poles for the discretized linear models are placed on 0.2 and 0.21 and a VARX model order of s=8 is chosen. These models are then used in the model-based approach (11) with $\lambda = 0.3$ and solved using CVX (Grant & Boyd, 2014).

Based on 100 realizations, the rates of successful fault isolation against the number of data points are presented in Figs. 4 and 5 for three and four active faults, respectively. The isolation is regarded successful if each nonzero component of $\hat{z}$ satisfying

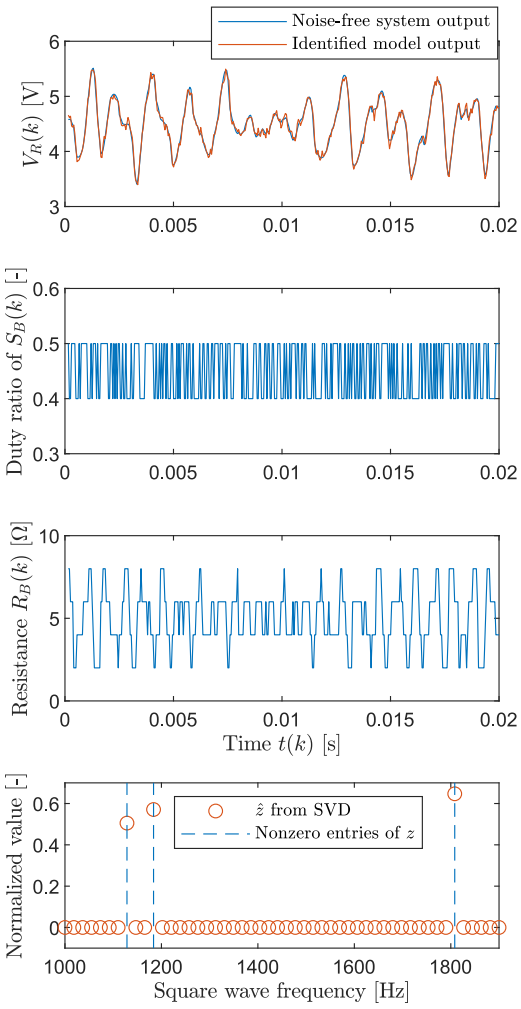

Fig. 3. Part of a realization of the noise-free system output of the buck-converter with corresponding duty ratio as input and load resistance $R_B(k)$. A solution of the optimization problem (14) provides diagnosis of z (bottom, red circles) and a refinement step using (14) with $\lambda = 0$ provides system identification with VAF 98.8% (top, red) with respect to the noise-free system output. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$|\hat{z}| > \epsilon \|\hat{z}\|_\infty$ with $\epsilon = 0.1$ is diagnosed while the other components are not, *i.e.* no misdetection nor overdetection.

For three active faults, Fig. 4 shows that the model-free approach requires more data than the model-based approach in order to reach a high isolation rate. However, if the employed model is erroneous, the isolation rate of the model-based approach remains at a value around 80%, regardless of the increasing data size. For four active faults, Fig. 5 presents a different phenomenon, namely a decreasing isolation rate from a certain data size. Since the net load resistance $R_B(k)$ approaches zero at certain time instances, the buck converter starts showing non-negligible nonlinear behavior. Since the performance of the model-free approach decreases slower than the model-based approach, it appears that the model-free approach can better accommodate this nonlinearity.

6.3. Proximal-based recursive implementation

The second simulation experiment demonstrates the recursive implementation of the proximal algorithm as described in Algorithm 2. Here we consider optimization problem (23) with a finite

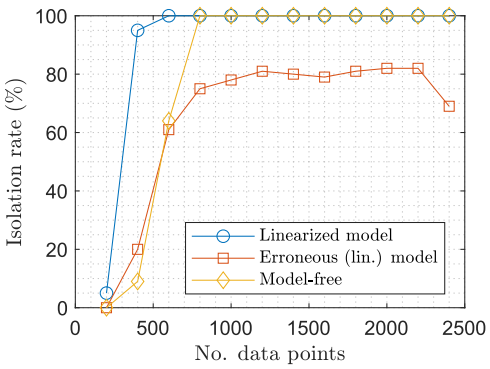

Fig. 4. Successful isolation rate in case of three active faults using a linearized model (blue, ○), an erroneous linearized model (red, □) and the proposed model-free approach (yellow, ◇). More data leads to an increased isolation rate and the model-free approach outperforming the model-based approach with an erroneous model.

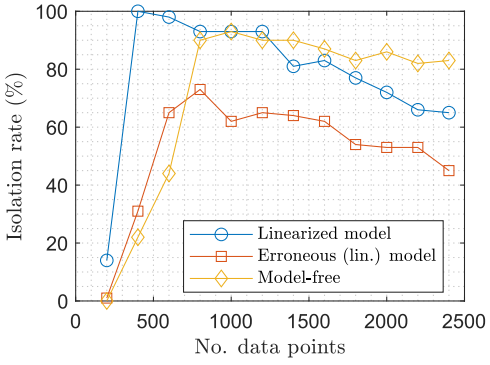

Fig. 5. Successful isolation rate in case of four active faults using a linearized model (blue, ○), an erroneous linearized model (red, □) and the proposed model-free approach (yellow, ◇). More data reveals the nonlinear nature of the buck converter in the condition of four active faults. The model-based approach decreases in performance more than the model-free approach.

window using $\mathcal{L} = 1000$ samples. Equivalently to Section 6.1, $s = 3$, $\tau = \lambda = 10$ and $\gamma = 10^{-3}$. Each time step, the PPXA algorithm performs $n_p = 10$ iterations initialized with a warm start from previous time step. The results are compared to those without regularization using recursive least squares (Sayed, 2003).

The results are shown in Fig. 6. In the initial configuration ($0 \leq t < 0.06$ s) the VAF for system identification converges around $t = 0.01$ s and the faults can be diagnosed correctly from circa $t = 0.04$ s. In the second phase ($0.06 \leq t < 0.12$ s) the active faults have changed to different square wave frequencies. The VAF decreases slightly after which it slowly increases again. Simultaneously, the new faults can be diagnosed from circa $t = 0.11$ s. In the third phase ($0.12 \leq t < 0.18$ s) the system characteristics have changed by adjusting the nominal load resistance R_B from 5 Ω to 3.5 Ω. This results in a reduction in the VAF, after which it increases again. The diagnosed faults remain roughly constant. In the fourth phase ($0.18 \leq t < 0.3$ s) both the active faults and model change to their original configuration. Now the VAF shows a sharp drop, but again it recovers after some time and also the original faults are diagnosed again around $t = 0.23$ s. By the regularization on the nuclear norm and the 1-norm, the recursive proximal algorithm outperforms unregularized recursive least squares in identifying the system in a small number of measurements, and in highlighting the active faults while suppressing the entries of the non-active faults.

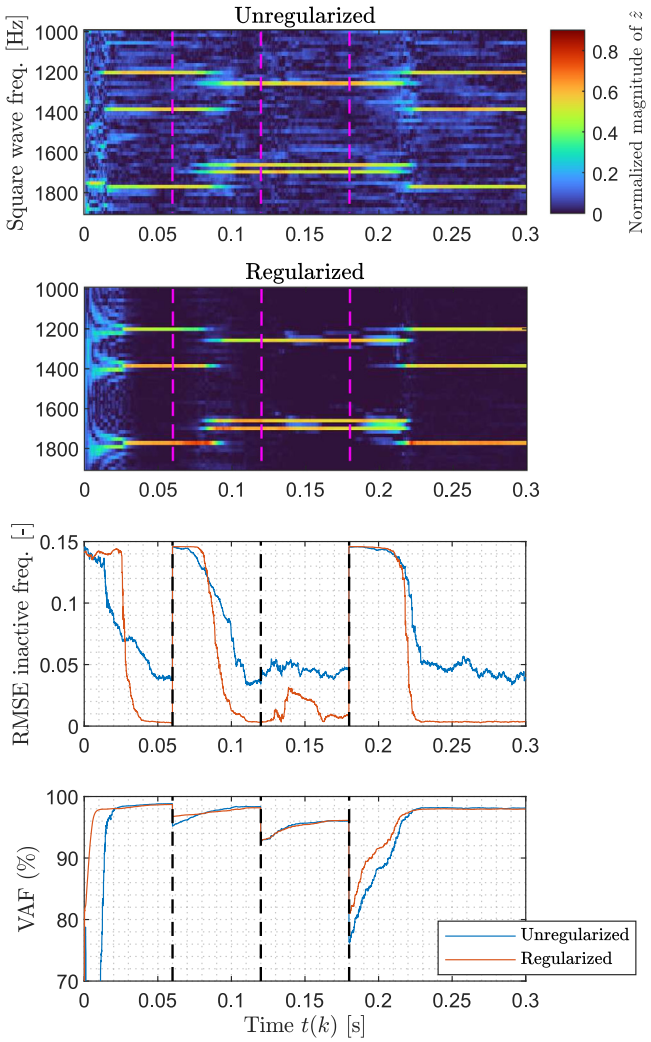

Fig. 6. The diagnosed faults over time without regularization using recursive least squares ((Sayed, 2003), FW with $\mathbf{R}_0 = I$, top) and with regularization using the recursive proximal algorithm ((23), FW with $\mathbf{R}_0 = 0$ solved using Algorithm 2, second), together with their corresponding Root Mean Square Error (RMSE) with respect to the inactive faults (third) and the VAF of the simultaneously identified model (bottom). The vertical dashed lines at $t = 0.06$, $t = 0.12$ and $t = 0.18$ s indicate a change in active faults, system characteristics and both, respectively. The average computational time per time instance k on an Intel i7-9750H CPU was 929 μs for Algorithm 2 versus 148 μs for unregularized recursive least squares. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

7. Conclusions

Model-free data-driven fault diagnosis aims at identifying the system and diagnosing the faults *simultaneously*, eliminating the necessity of an extensive identification phase prior to diagnosing faults. A proposed approach reformulates it as a convex optimization problem in order to make it computationally attractive. This approach is implemented online using a recursive implementation of a proximal algorithm.

The numerical results on a buck converter show how the faults are diagnosed simultaneous to system identification. A second simulation shows how the recursive implementation handles varying system parameters and (dis-)appearance of faults during operation.

The newly introduced methodology provides ample room for future research. A few examples are outlined as follows. First, the diagnosis performance trade-off with respect to the data

size could be analyzed theoretically, including a comparison with model-based and state-of-the-art data-driven techniques. This is presumably dependent on the chosen input signal and the fault dictionary. The second is the evaluation of the uncertainty of the estimated quantities with respect to the number of active faults, the properties of the I/O-data and those of the dictionary. On the algorithmic side we mention a third area of future research related to competitive optimization algorithms able to handle more than two possibly non-smooth objective term. In line of that research objective the computational efficiency of the recursive implementation could be further improved exploiting possible structure in the problem.

Appendix A. Buck converter system equations

Regarding the continuous-time variables

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} i_L(t) \\ V_R(t) \end{bmatrix} \qquad u(t) = S_B(t) \tag{A.1}$$

where $i_L(t)$ is the current through the inductor, the averaged (Tan & Hoo, 2015) continuous-time model for the buck converter is

$$\dot{x}(t) = \begin{bmatrix} 0 & \frac{-1}{L_B} \\ \frac{1}{C_B} & \frac{-1}{R_B C_B} \end{bmatrix} x(t) + \begin{bmatrix} \frac{V_B}{L_B} \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ \frac{\Delta R_B(t) V_R(t)}{(R_B + \Delta R_B(t)) R_B C_B} \end{bmatrix}$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t) + v(t) \tag{A.2}$$

Linearization around $\overline{\Delta R_B} = 0$ and $\overline{V}_R$ yields

$$\dot{x}(t) = \begin{bmatrix} 0 & \frac{-1}{L_B} \\ \frac{1}{C_B} & \frac{-1}{R_B C_B} \end{bmatrix} x(t) + \begin{bmatrix} \frac{V_B}{L_B} \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ \frac{\overline{V}_R}{R_B^2 C_B} \end{bmatrix} d(t)$$
$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(t) + v(t) \tag{A.3}$$

where $d(t) = \Delta R_B(t)$. Assuming zero-order hold for $u(t)$ and $d(t)$, the discretized state-space representation is

$$x(k+1) = Ax(k) + Bu(k) + Fd(k)$$
$$y(k) = x(k) + v(k), \tag{A.4}$$

which can be approximated by the VARX model (5).

References

Ahmeid, Mohamed, Armstrong, Matthew, Gadoue, Shady, Al-Greer, Maher, & Missailidis, Petros (2017). Real-time parameter estimation of DC–DC converters using a self-tuned Kalman filter. *IEEE Transactions on Power Electronics*, *32*(7), 5666–5674.

Ajalloeian, Amirhossein, Simonetto, Andrea, & Dall'Anese, Emiliano (2020). Inexact online proximal-gradient method for time-varying convex optimization. In *2020 American control conference* (pp. 2850–2857).

Al-Greer, Maher, Armstrong, Matthew, Ahmeid, Mohamed, & Giaouris, Damian (2019). Advances on system identification techniques for DC–DC switch mode power converter applications. *IEEE Transactions on Power Electronics*, *34*(7), 6973–6990.

Ali, Alnur, & Tibshirani, Ryan J. (2019). The generalized lasso problem and uniqueness. *Electronic Journal of Statistics*, *13*(2), 2307–2347.

Angelosante, Daniele, & Giannakis, Georgios B. (2009). RLS-weighted lasso for adaptive estimation of sparse signals. In *2009 IEEE international conference on acoustics, speech and signal processing* (pp. 3245–3248).

Basseville, Michèle, & Nikiforov, Igor V. (1993). *Detection of abrupt changes: theory and application*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc..

Beck, Amir, & Teboulle, Marc (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, *2*(1), 183–202.

Blanke, Mogens, Kinnaert, Michel, Lunze, Jan, & Staroswieck, Marcel (2006). *Diagnosis and fault-tolerant control* (2nd). Springer-Verlag Berlin Heidelberg.

Bliek, Laurens, Verstraete, Hans R. G. W., Verhaegen, Michel, & Wahls, Sander (2018). Online optimization with costly and noisy measurements using random Fourier expansions. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(1), 167–182.

Cai, Jian-Feng, Candès, Emmanuel J., & Shen, Zuowei (2010). A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, *20*(4), 1956–1982.

Chen, Zhiwen (2017). *Data-driven fault detection for industrial processes*. Springer Fachmedien Wiesbaden.

Chiuso, Alessandro (2007). The role of vector autoregressive modeling in predictor-based subspace identification. *Automatica*, *43*(6), 1034–1048.

Combettes, Patrick L., & Pesquet, Jean-Christophe (2007). A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, *1*(4), 564–574.

Combettes, Patrick L., & Pesquet, Jean-Christophe (2008). A proximal decomposition method for solving convex variational inverse problems*. *Inverse Problems*, *24*(6), Article 065014.

Combettes, Patrick L., & Pesquet, Jean-Christophe (2011). Proximal splitting methods in signal processing. In Heinz H. Bauschke, Regina S. Burachik, Patrick L. Combettes, Veit Elser, D. Russell Luke, & Henry Wolkowicz (Eds.), *Fixed-point algorithms for inverse problems in science and engineering*: *vol. 49*, (pp. 185–212). Springer New York.

Combettes, Patrick L., & Wajs, Valérie R. (2005). Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, *4*(4), 1168–1200.

Dai, Xuewu, & Gao, Zhiwei (2013). From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Transactions on Industrial Informatics*, *9*(4), 2226–2238.

Davis, Damek, & Yin, Wotao (2017). A three-operator splitting scheme and its optimization applications. *Set-Valued and Variational Analysis*, *25*(4), 829–858.

Ding, Steven X. (2013). *Model-based fault diagnosis techniques* (2nd). London: Springer-Verlag.

Ding, Steven X. (2014). *Data-driven design of fault diagnosis and fault-tolerant control systems*. London: Springer-Verlag.

Dixit, Rishabh, Bedi, Amrit Singh, Tripathi, Ruchi, & Rajawat, Ketan (2019). Online learning with inexact proximal online gradient descent algorithms. *IEEE Transactions on Signal Processing*, *67*(5), 1338–1352.

Freeman, Paul, Seiler, Peter, & Balas, Gary J. (2013). Air data system fault modeling and detection. *Control Engineering Practice*, *21*(10), 1290–1301.

Gao, Zhiwei, Cecati, Carlo, & Ding, Steven X. (2015a). A survey of fault diagnosis and fault-tolerant techniques – Part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Transactions on Industrial Electronics*, *62*(6), 3757–3767.

Gao, Zhiwei, Cecati, Carlo, & Ding, Steven X. (2015b). A survey of fault diagnosis and fault-tolerant techniques – Part II: Fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Transactions on Industrial Electronics*, *62*(6), 3768–3774.

Grant, Michael, & Boyd, Stephen (2014). CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx.

Gunnarsson, S. (1996). Combining tracking and regularization in recursive least squares identification. In *Proceedings of 35th IEEE conference on decision and control*: *vol. 3*, (pp. 2551–2552).

Ljung, L. (1999). System identification: Theory for the user. In *Prentice hall information and system sciences series*, Prentice Hall PTR.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin, Heidelberg: Springer-Verlag.

Noom, Jacques, Soloviev, Oleg, & Verhaegen, Michel (2023). Data-driven fault diagnosis under sparseness assumption for LTI systems. *IFAC-PapersOnLine*, *56*(2), 7722–7727.

Parikh, Neal, & Boyd, Stephen (2014). Proximal algorithms. *Foundations and Trends® in Optimization*, *1*(3), 127–239.

Phan, M. Q., & Longman, R. W. (1996). Relationship between state-space and input-output models via observer Markov parameters. *WIT Transactions on the Built Environment*, *22*, 185–200.

Raguet, Hugo, Fadili, Jalal, & Peyré, Gabriel (2013). A generalized forward-backward splitting. *SIAM Journal on Imaging Sciences*, *6*(3), 1199–1226.

Samad, Tariq, Bauer, Margret, Bortoff, Scott, Di Cairano, Stefano, Fagiano, Lorenzo, Odgaard, Peter Fogh, et al. (2020). Industry engagement with control research: Perspective and messages. *Annual Reviews in Control*, *49*, 1–14.

Sayed, A. H. (2003). Fundamentals of adaptive filtering. Wiley.

Scobee, Dexter, Ratliff, Lillian, Dong, Roy, Ohlsson, Henrik, Verhaegen, Michel, & Sastry, S. Shankar (2015). Nuclear norm minimization for blind subspace identification (N2BSID). In *54th IEEE conference on decision and control* (pp. 2127–2132).

Simani, Silvio (2021). Data-driven methods for fault diagnosis. In Vicenç Puig, & Silvio Simani (Eds.), *Diagnosis and fault-tolerant control 1: data-driven and model-based fault diagnosis techniques* (pp. 131–195). ISTE Ltd and John Wiley & Sons, Inc..

Tan, Rodney H. G., & Hoo, Landon Y. H. (2015). DC–DC converter modeling and simulation using state space approach. In *2015 IEEE conference on energy conversion* (pp. 42–47).

Venkatasubramanian, Venkat, Rengaswamy, Raghunathan, Kavuri, Surya N., & Yin, Kewen (2003). A review of process fault detection and diagnosis – Part III: Process history based methods. *Computers & Chemical Engineering*, *27*(3), 327–346.

Verhaegen, Michel, & Verdult, Vincent (2007). *Filtering and system identification*. Cambridge University Press.

Yin, Shen, Ding, Steven X., Xie, Xiaochen, & Luo, Hao (2014). A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, *61*(11), 6418–6428.

Zhang, Qinghua (2021). Dynamic system fault diagnosis under sparseness assumption. *IEEE Transactions on Signal Processing*, *69*, 2499–2508.

Jacques Noom received a B.Sc. degree in Mechanical Engineering in 2017 and a M.Sc. degree in Systems & Control in 2019, both cum laude at Delft University of Technology, The Netherlands. From 2019 he pursues a Ph.D. degree in Systems & Control at Delft University of Technology, The Netherlands. His current research interests include fault diagnosis, blind system identification and the application of convex optimization to control problems.

Oleg Soloviev received his M.Sc. degree in mathematics and applied mathematics in 1994 from Moscow State University (Russia) and his Ph.D. degree in 2006 from Delft University of Technology, the Netherlands, with a thesis titled "Methods and Sensors for Accurate Wavefront Measurements". His expertise lies in the design and development of deformable mirrors and adaptive optics systems, as well as in developing algorithms for phase retrieval, wavefront sensing, imaging, and machine vision. Since 2006, he has been employed as a Senior Associate by Flexible Optical BV. Additionally, he holds a part-time position as a Senior Research Fellow at DCSC, TU Delft, since 2015.

Michel Verhaegen received an engineering degree in aeronautics from the Delft University of Technology, Delft, The Netherlands, in 1982 and the doctoral degree in applied sciences from the Catholic University Leuven, Belgium, in 1985. During his graduate study, he held a research assistantship sponsored by the Flemish Institute for scientific research (IWT). From 1985 to 1994, he was a Research Fellow of the U.S. National Research Council (NRC), affiliated with the NASA Ames Research Center in California, and of the Dutch Academy of Arts and Sciences, affiliated with the Network Theory Group of the Delft University of Technology. From 1994 to 1999, he was an Associate Professor of the Control Laboratory, Delft University of Technology and was appointed as Full Professor with the Faculty of Applied Physics, University of Twente, The Netherlands, in 1999. In 2001, he moved back to the University of Delft and is now a member of the Delft Center for Systems and Control. He has held short sabbatical leaves at the University of Uppsala, McGill, Lund, and the German Aerospace Research Center (DLR) in Munich and is participating in several European Research Networks. His main research interest is the interdisciplinary domain of numerical algebra and system theory. In this field he has published over 150 peer reviewed papers in top Journals in the field of Systems and Control and optical systems. This includes 38 papers in IEEE Transactions and 17 papers in AUTOMATICA journal. He received the prestigious IFAC [International Federation for Automatic Control] Fellow Award in 2014. This award is given to persons who have made outstanding and extraordinary contributions in the field of interest of IFAC, in the role as an Engineer/Scientist, Technical Leader, or Educator. He also was a recipient of an Advanced Personal Grant from the European Research Council in 2017. Currently he has a renewed interest in broadening the field of system identification in areas such as imaging systems, data driven fault detection and control.