

RESEARCH

Open Access



Bayesian fault detection and isolation using Field Kalman Filter

Jerzy Baranowski* , Piotr Bania, Indrajeet Prasad and Tian Cong

Abstract

Fault detection and isolation is crucial for the efficient operation and safety of any industrial process. There is a variety of methods from all areas of data analysis employed to solve this kind of task, such as Bayesian reasoning and Kalman filter. In this paper, the authors use a discrete Field Kalman Filter (FKF) to detect and recognize faulty conditions in a system. The proposed approach, devised for stochastic linear systems, allows for analysis of faults that can be expressed both as parameter and disturbance variations. This approach is formulated for the situations when the fault catalog is known, resulting in the algorithm allowing estimation of probability values. Additionally, a variant of algorithm with greater numerical robustness is presented, based on computation of logarithmic odds. Proposed algorithm operation is illustrated with numerical examples, and both its merits and limitations are critically discussed and compared with traditional EKF.

Keywords: Fault detection, Bayesian reasoning, Data analysis, Field Kalman Filter

1 Introduction

The problem of fault detection and isolation is one of the most important research topics in process control. Efficient methods should quickly detect the occurrence of a fault and determine the fault origin with sufficient probability. However, people usually have to deal with measurement disturbances, process noise, and complicated dynamics when analyzing the actual industrial processes. As the field is very broad for the purpose of this paper, we will focus on faults that can be represented by a change of either parameters in the mathematical model or parameters of stochastic disturbances of the analyzed system. Diagnosis and isolation of such faults relies on two elements—detecting the change in the parameter value and isolating what type of fault causes such change.

A typical approach to such estimation for linear distributed systems is the use of variations of Kalman filters, often used as residue generators, with some kind of classification algorithm. The problem of estimating the state of a linear dynamic system driven by additive Gaussian noise with unknown time varying statistics is considered in [1] through Kalman filtering with Bayesian methods for

estimation of noise variance, while maximum likelihood is used in [2], and correlation methods are used in [3] and [4].

In [5], methodology for the detection and accommodation of actuator faults for a class of multi-input–multi-output (MIMO) stochastic systems is presented. Actuator fault diagnosis is based on the estimation of the state vector, where Kalman filter is used to estimate the state. A Kalman filter approach is proposed for state estimation for stochastic discrete-time linear systems in [6–9] while in [10–12], a Kalman filter is applied for estimation of state vector for non-linear systems. Model-based isolation and estimation of additive faults in discrete-time linear Gaussian systems is presented in [13], where fault estimation is carried out and a Kalman filter is used with Bayesian statistics. A novel sensor fault detection and isolation algorithm based on an extended Kalman filter is presented for noise and efficiency in real-time implementation in [14]. In the paper [15], a diagnostic system based on a uniquely structured Kalman filter is developed for its application to in-flight fault detection of aircraft engine sensors. A bibliographic review on a distributed Kalman filter is provided in [16] with a comparison to the different approaches. Very important group of methods for fault detection and isolation is based on particle

*Correspondence: jb@agh.edu.pl

Department of Automatics and Biomedical Engineering, AGH University of Science and Technology, Aleja Adama Mickiewicza 30, 30-059 Krakow, Poland

filtering. This method, a certain extension of Kalman filtering, allows very good results in a wide spectrum of cases. The disadvantage can be however increased computational complexity. Important works on the subject are among the others [17–22].

It should be also noted that similar approach to system analysis is present in the issues of model selection and working point detection problems. First of this problems is used to verify which model of the possible group fits the best presented data. It is usually connected with some kind of hypothesis testing. The second group has a special place in the process control applications. It is very often encountered in industrial situations that the system is very hard to be modeled fully representing its nonlinear character. In such cases, a set of linearized models in different working points is created. Appropriate algorithm for choice of which linear model should now be used depending on data is very similar to fault detection and isolation one.

In this paper, the authors expand on their earlier results in [23], where the Field Kalman Filter was proposed. The Field Kalman Filter is an advanced signal processing algorithm which in its full form allows full estimation of state and parameters of linear systems with Gaussian disturbances. The main disadvantage of this filter is that in general case, it is infinite dimensional, and for practical applications, moving horizon approximation has to be used.

In this paper, we apply and modify this filter for the problem of fault detection and isolation. In particular, an advantage is taken over the fact that typical faults usually form a limited set of cases. In such situation, problem dimension reduces substantially, and the algorithm requires solution of a parameter-dependent family of systems of equations. Proposed filter is optimal for systems with fixed parameters and highly efficient for those where changes in parameter values occur. We present the algorithm for computing fault probability and also its more numerically robust version for computing logarithmic odds of fault occurrence with respect to each other possibility.

The rest of this paper is organized as follows. First, we present the Field Kalman Filter in the discrete case. Second, we formulate its application for fault detection and isolation problem. We formulate the necessary modification of the filter for parameter-varying systems with formulation of necessary forgetting operator. We discuss algorithms' numerical drawbacks and introduce more robust modification. Then, we illustrate both algorithms' operation with advanced examples of fault detection, where faults are either in parameters or in the disturbances. The paper ends with discussion, comparing our results to those based on EKF and conclusions regarding future work.

Remark 1 (Notation) We denote the set of symmetrical and positively defined matrices of dimension n by $S^+(n)$. By $x \sim N(m, S)$, we understand a normally distributed random variable with a mean m and a covariance matrix S . Also, we define function $N(x, m, S)$ as

$$N(x, m, S) = \frac{1}{\sqrt{(2\pi)^n |S|}} \exp\left(-\frac{1}{2}(x - m)^T S^{-1}(x - m)\right). \quad (1)$$

2 Field Kalman Filter

Let us consider the following stochastic system

$$x_{k+1} = A(\theta, u_k)x_k + B(\theta, u_k)u_k + G(\theta, u_k)w_k, \quad (2)$$

$$y_k = C(\theta)x_k + v_k. \quad (3)$$

We have the assumptions as following:

1. $x_k \in \mathbb{R}^n, y_k \in \mathbb{R}^m, w_k \in \mathbb{R}^l, v_k \in \mathbb{R}^m, w_k \sim N(0, I), v_k \sim N(0, V(\theta)), u_k \in U, U = \{u \in \mathbb{R}^r : u_{min} \leq u \leq u_{max}\}, k = 0, 1, 2, \dots$
2. Equations (2)–(3) depend on parameter vector $\theta \in \Omega \subset \mathbb{R}^p$. The matrix functions A, B, G, C, V are of C^1 class w.r.t. both arguments and are of appropriate dimensions.
3. The prior distribution of parameter θ will be denoted by p_0 . The set of all initial distributions of θ is defined as $\pi_0 = \{p_0 \in L_1(\Omega; \mathbb{R}) : p_0(\theta) \geq 0 \wedge \|p_0\|_1 = 1\}$.
4. We assume that the initial joint distribution of the variables (x_0, θ) has the following form

$$q_0(x_0, \theta) = p_0(\theta)N(x_0, m_0, S_0). \quad (4)$$

where $m_0 \in \mathbb{R}^n, S_0 \in S^+(n), p_0 \in \pi_0$.

5. The first measurement is performed at $k = 1$, and $Y_k = (y_1, y_2, \dots, y_k) \in \mathbb{R}^{m \times k}$ is a matrix of k first measurements.

Remark 2 We define that observability of the parameters has the following property

$$\theta_1 \neq \theta_2 \Rightarrow p(\theta_1 | Y_k) \neq p(\theta_2 | Y_k) \forall \theta_1, \theta_2, Y_k. \quad (5)$$

In other words, we require that probabilities of different values of parameters conditional on the same set of measurements are different. In this paper, we assume that the parameters of all considered systems are observable.

In the following theorem, we provide a formula for calculating the distribution of the state and parameters of system (2)–(3), which we will call a Field Kalman Filter.

Theorem 1 Let u_k be given by fixed sequence and let $A_i(\theta) = A(\theta, u_i)$, $B_i(\theta) = B(\theta, u_i)$, and $D_i(\theta) = G(\theta, u_i)G(\theta, u_i)^T$. If assumptions 1–5 are fulfilled, then joint density of variables x_k, θ, Y_k is given by

$$p(x_k, \theta, Y_k) = p_0(\theta)N(x_k, m_k(\theta), S_k(\theta)) \times \prod_{i=1}^k N(y_i, c_i(\theta)m_i^-(\theta), W_i(\theta)). \tag{6}$$

where

$$m_0(\theta) = m_0, \tag{7}$$

$$S_0(\theta) = S_0, \tag{8}$$

$$m_i^-(\theta) = A_{i-1}(\theta)m_{i-1}(\theta) + B_{i-1}(\theta)u_{i-1}, \tag{9}$$

$$S_i^-(\theta) = A_{i-1}(\theta)S_{i-1}(\theta)A_{i-1}(\theta)^T + D_{i-1}(\theta), \tag{10}$$

$$W_i(\theta) = V(\theta) + C(\theta)S_i^-(\theta)C(\theta)^T, \tag{11}$$

$$S_{i-1}(\theta) = S_i^-(\theta) - S_i^-(\theta)C(\theta)^TW_i^{-1}(\theta)C(\theta)S_i^-(\theta), \tag{12}$$

$$m_{i-1}(\theta) = m_i^-(\theta) + S_i(\theta)C(\theta)^TV(\theta)^{-1} \times (y_i - C(\theta)m_i^-(\theta)), \tag{13}$$

$$i = 1, 2, \dots, k + 1.$$

Proof The proof of a very similar theorem can be found in [23]. The difference is that here, we consider a discrete-time system instead of a continuous one. Nevertheless, there is a completely analogous distribution construction for these two types of time systems. \square

To create a diagnostic algorithm, we will need the following corollary:

Corollary 1 Under the assumptions of Theorem 1, we have

$$p(x_k | \theta, Y_k) = N(x_k, m_k(\theta), S_k(\theta)). \tag{14}$$

and

$$p(\theta | Y_k) = \frac{p_0(\theta) \prod_{i=1}^k N(y_i, c_i(\theta)m_i^-(\theta), W_i(\theta))}{\int_{\Omega} p_0(\theta) \prod_{i=1}^k N(y_i, c_i(\theta)m_i^-(\theta), W_i(\theta))d\theta}. \tag{15}$$

Proof It is a direct consequence of the Bayes' theorem and Eq. (6). \square

Now, we can formulate the fault detection problem.

3 Fault detection problem

Equations (14) and (15) are the solution to estimating the distributions of state and parameters. This solution (filter), in general, is infinite dimensional. It is caused by the fact that we need to solve Eqs. (7)–(13) for all the possible values of θ , in order to compute the necessary integral in Eq. (15). Because of this sizeable computation, it is greatly difficult to implement in practice.

Situation changes when the number of potential parameters in set Ω is finite. In this case, the integral in (15) becomes a summation operator. Consequently, Eqs. (7)–(13) only need to be calculated for a finite number of θ values.

This problem formulation, i.e., $\theta \in \Omega = \{\theta_0, \theta_1, \theta_2, \dots, \theta_r\}$, is not a common case for parameter estimation; however, it is greatly suited to a fault detection problem, when we already know what type of faulty conditions can occur in the system.

Next, we denote the parameter θ_0 as a normal operating condition. The parameters $\theta_1, \theta_2, \dots, \theta_r$ represent system faults. The initial distribution of θ can be rewritten as

$$p_0(\theta) = \sum_{s=0}^r p_{0,s}\delta(\theta - \theta_s). \tag{16}$$

As a consequence, the distribution Eq. (15) becomes a discrete distribution

$$p_0(\theta | Y_k) = \sum_{s=0}^r p_{k,s}\delta(\theta - \theta_s). \tag{17}$$

where,

$$p_{k,s} = p(\theta = \theta_s | Y_k) = \frac{p_{0,s} \prod_{i=1}^k N(y_i, c(\theta_s)m_i^-(\theta_s), W_i(\theta_s))}{\sum_{s=0}^r p_{0,s} \prod_{i=1}^k N(y_i, c(\theta_s)m_i^-(\theta_s), W_i(\theta_s))}, \tag{18}$$

$$s = 1, \dots, r.$$

Equation (18) can be rewritten in a more convenient, recursive form

$$p_{k,s} = p_{k-1,s} \frac{N(y_k, c(\theta_s)m_k^-(\theta_s), W_k(\theta_s))}{\sum_{s=0}^r p_{k-1,s}N(y_k, c(\theta_s)m_k^-(\theta_s), W_k(\theta_s))}, s = 1, \dots, r. \tag{19}$$

Now, the distribution of (x_k, θ) conditioned on Y_k takes the form

$$p(x_k, \theta | Y_k) = N(x_k, m_k(\theta), S_k(\theta)) \sum_{s=0}^r p_{k,s}\delta(\theta - \theta_s). \tag{20}$$

Equations (7)–(13), (19) and (20) describe the finite-dimensional recursive optimal filter for system (2)–(3).

induces an extremely small value of $p_{k,s}$. Consequently, Eq. (29) works out as a quotient of two extremely small numerical values which can cause a numerically intractable result. This numerical limitation was also observed and illustrated in a scalar system example demonstrated in Section 5 of our paper.

Therefore, to eliminate this numerical instability, we propose an alternative approach, using logarithmic odds (see for example [24]). Let us consider the logarithm of probability quotient i.e.

$$e_{k,i,j} = \ln \frac{p_{k,i}}{p_{k,j}}, i, j = 0, 1, \dots, r. \quad (33)$$

By substituting Eqs. (1) and (19) into (33), we get the following recursion

$$e_{k,i,j} = e_{k-1,i,j} + d_{k,i,j}, \quad (34)$$

where

$$\begin{aligned} d_{k,i,j} = & \frac{1}{2} \ln |W_k(\theta_j)| - \frac{1}{2} \ln |W_k(\theta_i)| + \frac{1}{2} |y_k \\ & - C(\theta_j) m_k^-(\theta_j) |_{W_k(\theta_j)^{-1}}^2 \\ & - \frac{1}{2} |y_k - C(\theta_i) m_k^-(\theta_i) |_{W_k(\theta_i)^{-1}}^2. \end{aligned} \quad (35)$$

To use logarithmic odds to determine most probable fault, we need to verify for which fault θ_s all the ratios $e_{k,s,j}$ are positive. Such fault has maximum a posteriori probability.

Whenever we have positive logarithmic odds for one parameter value $\theta_{i_{MAP}}$ against others, i.e., $e_{k,i_{MAP},j} \geq 0$ for all $j = 0, 1, \dots, r$, then it is the most probable value of the parameter, i.e., $p_{k,j} \leq p_{k,i_{MAP}}$.

Furthermore, to introduce the forgetting operation mentioned in Section 3.1, we can implement it as

$$e_{k,i,j} = \beta e_{k-1,i,j} + d_{k,i,j} \quad (36)$$

where $\beta \in [0, 1]$ is a new forgetting factor. The product of $e_{k-1,i,j}$ and $\beta \in [0, 1)$ corresponds to certain new distribution. This new distribution has greater entropy and is closer to a uniform distribution. Hence, if $\beta \in (0, 1)$, the old data need to be gradually forgotten. If $\beta = 0$, then $e_{k,i,j} = 0$ and $p_{k,i} = p_{k,j}$ for all $i, j = 1, \dots, r$. If $\beta = 1$, it means that all the old data are supposed to be taken into account.

Now, we can propose a logarithmic odds algorithm for fault detection and classification:

5 Examples of operation

In this section, we illustrate the algorithm operation using three examples of linear systems. In the first one, we use Algorithm 1 for a scalar discrete system, where all elements of the system depend on the parameters (i.e., state, control and output). In the second example, we use the same algorithm for a second-order, time continuous system, with discrete-time observation, parameter-dependent state, control and noise matrices, and noise

variance. The final example uses the same system to illustrate the log-odds Algorithm 2.

Algorithm 2 Logarithmic odds fault detection

1: **Initialization:**

$$\begin{aligned} \beta & \in [0, 1], k = 0, m_0(\theta_s), S_0(\theta_s), \\ p_0 & = (p_{0,0}, p_{0,1}, \dots, p_{0,r})^T, \\ \Omega & = \{\theta_0, \dots, \theta_r\}. \end{aligned} \quad (37)$$

2: **Prediction step:**

$$m_k^-(\theta_s) = A_{k-1}(\theta_s) m_{k-1}(\theta_s) + B_{k-1}(\theta_s) u_{k-1}, \quad (38)$$

$$S_k^-(\theta_s) = A_{k-1}(\theta_s) S_{k-1}(\theta_s) A_{k-1}(\theta_s)^T + D_{k-1}(\theta_s), \quad (39)$$

$$W_k(\theta_s) = V(\theta_s) + C(\theta_s) S_k^-(\theta_s) C(\theta_s)^T, \quad (40)$$

$$e_{k,s,j}^- = \beta e_{k-1,s,j}, s, j = 1, \dots, r. \quad (\text{Forgetting}) \quad (41)$$

3: **Correction step:**

$$S_k(\theta_s) = S_k^-(\theta_s) - S_k^-(\theta_s) C(\theta_s)^T W_k^{-1}(\theta_s) C(\theta_s) S_k^-(\theta_s), \quad (42)$$

$$\begin{aligned} m_k(\theta_s) & = m_k^-(\theta_s) + S_k(\theta_s) C(\theta_s)^T V(\theta_s)^{-1} \\ & (y_k - C(\theta_s) m_k^-(\theta_s)), \end{aligned} \quad (43)$$

$$e_{k,s,j} = e_{k,s,j}^- + d_{k,s,j}, s, j = 1, \dots, r. \quad (44)$$

4: **goto Prediction step.**

5.1 Example 1. Scalar system with changes in parameters

Let us consider the following scalar system

$$\begin{aligned} x_{k+1} & = a(\theta) x_k + b(\theta) u_k + g w_k, \\ y_k & = c(\theta) x_k + v_k, g = \sqrt{0.1}, E(v_k^2) = 0.1. \end{aligned} \quad (45)$$

In this system, we consider a normal operation state and four fault cases. The settings of these cases are shown in Table 1. The first fault corresponds to increase in system dynamics, combined with substantial increase of system gain. The second fault introduces system oscillation but a lesser gain increase. The third fault slightly increases systems' dynamics, increases input gain, and decreases output gain. The fourth fault introduces a larger increase in the dynamics than the third, but with smaller increment in input gain and lower reduction in output gain.

The initial conditions of the filter were set as $m_0(\theta) = 0$ and $S_0(\theta) = 0.01$. The initial probability distribution of

Table 1 Parameters of system (45) dependent on normal and faulty conditions

	Fault 0, normal operating conditions	Fault 1	Fault 2	Fault 3	Fault 4
$a(\theta)$	0.9	0.5	-0.5	0.8	0.7
$b(\theta)$	0.1	1	0.9	1	0.3
$c(\theta)$	1	2	1.2	0.5	0.6

faults was $p_0 = [0.95, 0.0125, 0.0125, 0.0125, 0.0125]$. The simulation of system started with normal operation conditions. Then, after every 200 steps, the operating mode was changed in such a scheme: normal, fault 1, fault 2, fault 3, fault 4, fault 3, fault 2, fault 1, and normal. The system was excited by a square input signal u with zero mean, a period of 158, and the amplitude as 1. We set the forgetting factor as $\alpha = 0.99$, which is experimentally efficient.

We present the measured signal in Fig. 1. We have applied the algorithm to this signal. Using Eq. (29), we computed the probabilities of particular faults and presented them in the Fig. 2. We can observe consistency between maximal probability and the planned test regime of faults occurring. It is evident in Fig. 3, where the results of fault detection and isolation are presented based on maximum a posteriori estimate.

The results show that estimator reacts very quickly, usually within five samples detecting a change in the parameter value. The numerical instability, described in Section 4, occurs and is visible in the Fig. 3.

5.2 Example 2. Oscillator with unknown noise variances

Let us consider a second-order continuous time system (oscillator) given by

$$dx = (A_c(\theta)x + B_c u)dt + G_c(\theta)dw,$$

$$A_c = \begin{bmatrix} 0 & 1 \\ -\theta_1^2 & -2\xi\theta_1 \end{bmatrix}, B_c = \begin{bmatrix} 0 \\ b_c\theta_1^2 \end{bmatrix}, G_c = \begin{bmatrix} 0 \\ \theta_2 \end{bmatrix}. \tag{46}$$

with discrete-time observation

$$y_k = x_1(t_k) + v_k,$$

$$v_k \sim N(0, \theta_3), \tag{47}$$

$$t_k = kT_0, T_0 > 0, k = 1, 2, \dots$$

where $\xi = 0.1$ and $b_c = 1$. The initial condition is Gaussian, i.e., $x(0) \sim N(m_0, S_0)$. We can see that this is a system with varying natural frequency, varying process, and measurement noise variances.

Assuming that control u is a piecewise constant, i.e., $u(t) = u_k, t \in [t_{k-1}, t_k)$, we can discretize system (46)–(47). The discrete-time system corresponding to Eqs. (46)–(47) has the form

$$x_{k+1} = A(\theta)x_k + B(\theta)u_k + G(\theta)w_k,$$

$$y_k = Cx_k + v_k, \tag{48}$$

$$x_0 \sim N(m_0, S_0), C = [1, 0].$$

where matrices $A, B,$ and D can be calculated from the formulas

$$A(\theta) = e^{A_c(\theta)T_0},$$

$$B(\theta) = \int_0^{T_0} e^{A_c(\theta)\tau} B_c d\tau, \tag{49}$$

$$D(\theta) = \int_0^{T_0} e^{A_c(\theta)\tau} G_c(\theta)G_c(\theta)^T e^{A_c(\theta)^T\tau} d\tau.$$

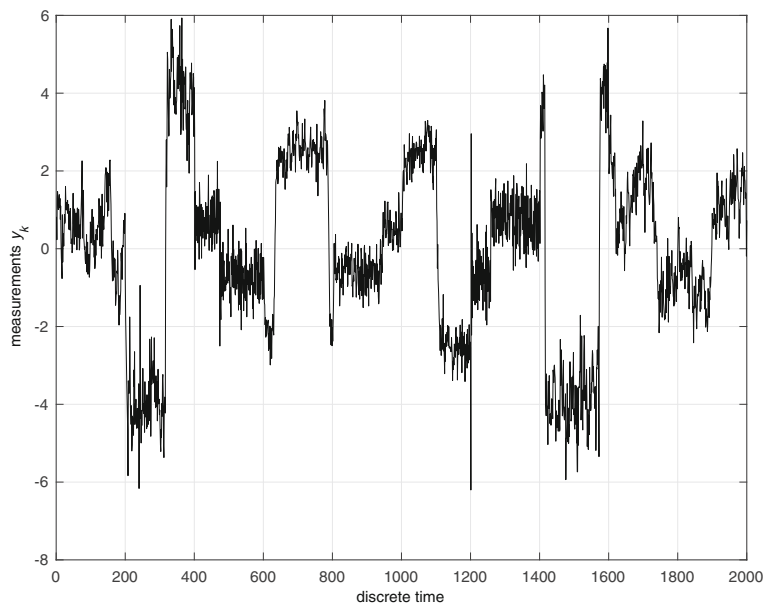


Fig. 1 Measured output signal for the system (45) with operating conditions changing every 200 samples

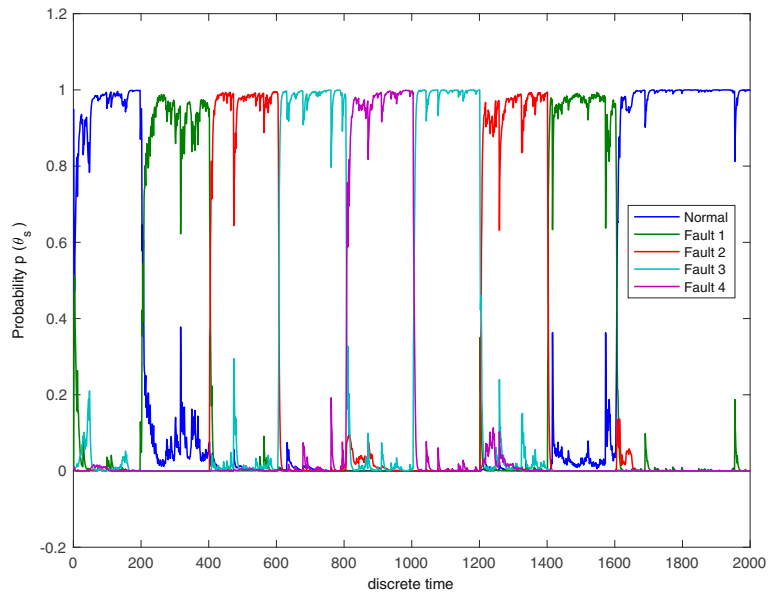


Fig. 2 Probability of which operation state system (45) is in conditional on the measurements. Probabilities are computed using formula (29)

$G(\theta)$ is non-uniquely defined as a factor of $D(\theta) = G(\theta)G(\theta)^T$. However, it is not needed for filter computation. In addition, it should be noted that discretization operation (49) increases the nonlinearity of system dependence on parameters. The discretization step is $T_0 = 0.05$.

In this example, noise variances θ_2 and θ_3 are unknown and must be estimated, which is not a typical task. We consider four cases: normal operating conditions, change

in the oscillator natural frequency (fault 1), increase in process noise variance by an order of magnitude (fault 2), and increase in measurement noise by an order of magnitude (fault 3). Values of parameters corresponding to normal operating conditions and faults are shown in Table 2.

The simulation was started with normal conditions. Next, after every 2500 steps (125 s), the operating mode

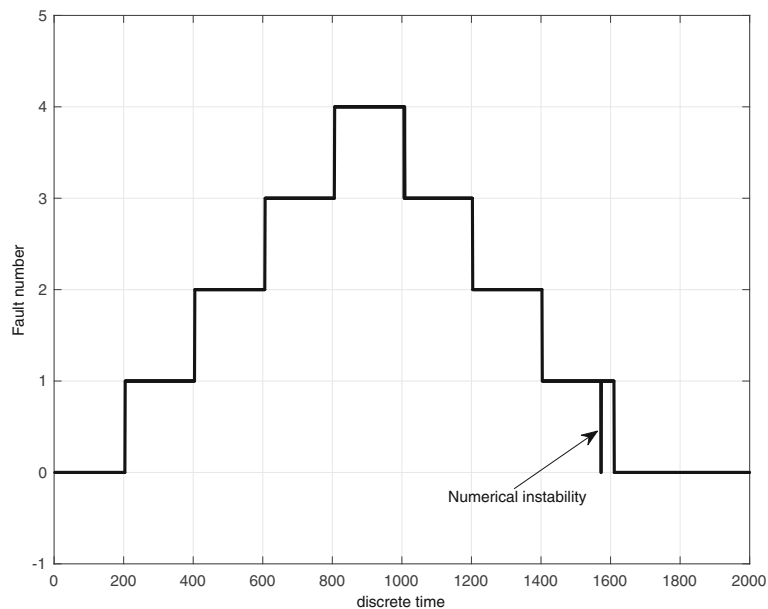


Fig. 3 Maximum a posteriori estimate of faulty condition. Vertical axis determines faulty condition. In sample 1572, there occurs a numerical instability, causing mislabeling the condition

Table 2 Parameters corresponding to normal operation and faulty conditions of oscillator (46)

	Fault 0, normal operating conditions	Fault 1	Fault 2	Fault 3
θ_1	3.0	5.0	3.0	3.0
θ_2	10^{-1}	10^{-1}	1.0	10^{-1}
θ_3	10^{-2}	10^{-2}	10^{-2}	10^{-1}

Parameter θ_1 is the oscillator natural frequency. Parameters θ_2 and θ_3 are process and measurement noise covariances, respectively

was changed according to the scheme as normal, fault 1, fault 2, and fault 3. The input u was a zero mean, square signal. The forgetting factor α was 0.99. Measurements are presented in Fig. 4. Probabilities obtained with Eq. (32) are presented in Fig. 5. Again, in this case, faults with maximum probability correspond to the fault schedule.

5.3 Example 3. Oscillator with unknown noise variances—logarithmic odds approach

In this example, we analyze the same second-order continuous time system as in the example 2. The settings of parameters were as $\xi = 0.1$, $b_c = 1$, $T_0 = 0.05$, $m_0 = (0, 0)^T$, and $S_0 = 10^{-2}I$. Initial distribution was uniform. The fault case 3 is now different than in the previous example, as the natural frequency also changes.

The simulation started with the normal conditions. Next, after every 2500 steps, the operating mode was changed in the scheme as normal, fault 1, fault 2, fault 3, and normal (see Table 3). The input u was a zero mean, square signal with a period of 630. The forgetting factor β was 0.995. The estimation results are shown in Fig. 6. We can see that there are no numerical errors;

however, there is a case of wrongly identified fault, when fault 2 is identified instead of normal state for a short time (comparing with a schedule).

We can reconstruct the probabilities from log-odds using formula

$$p_{k,j} = \frac{e^{-e_{k,i_{MAP},j}}}{\sum_{s=0}^r e^{-e_{k,i_{MAP},s}}} \tag{50}$$

and they are presented in Fig. 7. As one can see, the probability estimate is much less varying than in the case of using Algorithm 1.

6 Discussion

While proposed fault detection and isolation algorithm gives promising results, it is more than natural to compare it with the industry standard which is Extended Kalman Filter. We will now explain the differences between our approach and EKF.

The most important fact is that the Extended Kalman Filter (EKF) forms Gaussian approximation of true probability density of the state variables and parameters. The original state equations are extended by adding new equations which represent parameters as constant state variables. Prediction of their mean values is based on the extended system of state equations while the covariance matrix is calculated on the basis of linearized model. The correction step is the same as in the classical Kalman filter. Input noise and measurement noise covariances have to be known a priori and cannot be estimated by EKF. Due to the nonlinearity of the extended system, the EKF estimate can often diverge. In the considered fault isolation

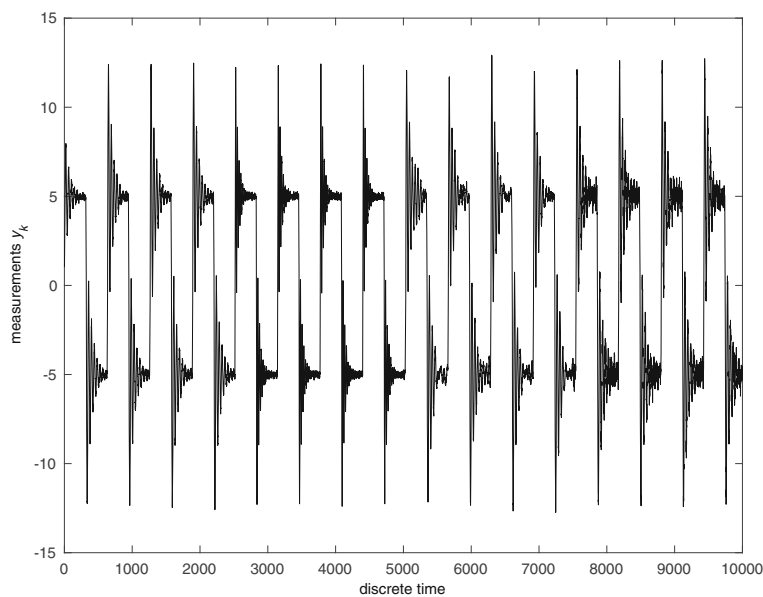


Fig. 4 Measured output signal for the oscillator (46) with operating conditions changing every 2500 samples

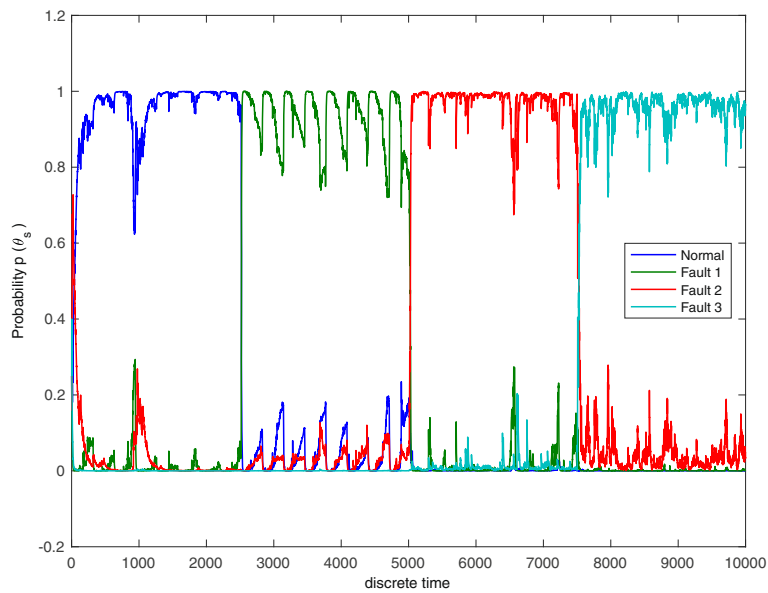


Fig. 5 Probability of which operation state the oscillator (46) is in conditional on the measurements. Probabilities are computed using formula (29)

problem, the probability density of the parameter values is discrete and its Gaussian approximation formed by EKF is continuous, so it cannot represent their character properly.

All these issues suggest that EKF is not an optimal (or sufficiently sub-optimal) tool for the considered fault detection and isolation problem. Nevertheless, it is possible to apply EKF for this task, using it to estimate the parameter values and providing appropriate thresholding to isolate appropriate faults.

In contrast to EKF, the proposed Field Kalman Filter forms true density of the state and parameter if the forgetting factor is 1. Noise covariances need not to be known a priori and they can be estimated on the basis of data. Since FKF allows calculation of probabilities of particular faults it is possible to easily construct statistical tests for fault detection and isolation.

The only actual advantage of EKF over FKF is a substantially reduced computational complexity. FKF needs to solve a system of equations for every considered fault type, while EKF requires solution of only one system of equations (but with extended state). This reduction

comes at a price; as for system with complicated parameter dependence (like the one in examples 2 and 3), EKF might not converge at all. Lack of convergence for parameter estimation (natural frequency of the oscillator) is presented in the Fig. 8.

7 Conclusions

We have presented two algorithms for fault estimation. Both of them are based on the proposed by the first two authors the Field Kalman Filter. Algorithm 1, i.e., direct application of FKF is attractive because it can be easily used for determination of fault probabilities, providing level of confidence in fault detection. There is however some numerical sensitivity connected to computations close to 0/0 operation.

At the cost of dropping, the direct probability computation one can use log-odds-based approach of the Algorithm 2. In this case, the problem of division by zero is reduced to addition and subtraction of numbers in the same order of magnitude. The main disadvantage is requirement of additional memory storage—with Algorithm 1, only the vector of r probabilities is needed, while Algorithm 2 requires a $r \times r$ matrix of logarithmic odds. Additionally, the result in form of odds can be complicated to interpret if the resulting probability distribution of faults is multi-modal. Original probabilities can be reconstructed but also with additional computational cost.

Additionally, as it was mentioned before, the forgetting operator is not unique. Proposed forgetting for Algorithms 1 and 2 are working properly for presented

Table 3 Parameters corresponding to normal operation and faulty conditions of oscillator (46) in the example 3

	Fault 0, normal operating conditions	Fault 1	Fault 2	Fault 3
θ_1	3.0	5.0	3.0	1.0
θ_2	10^{-1}	10^{-1}	1.0	1.0
θ_3	10^{-2}	10^{-2}	10^{-2}	10^{-1}

Parameter θ_1 is the oscillator natural frequency. Parameters θ_2 and θ_3 are process and measurement noise covariances, respectively

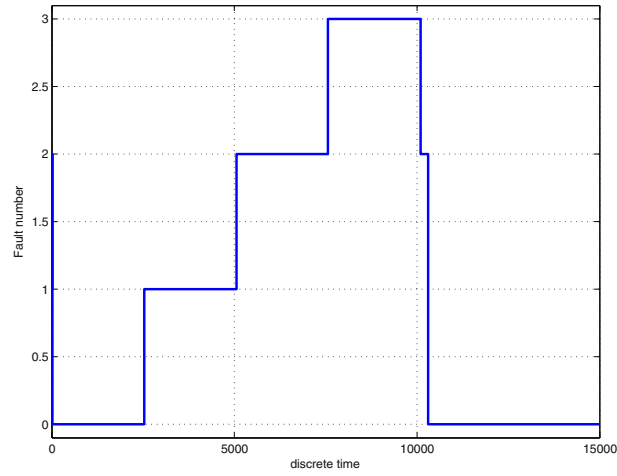


Fig. 6 Estimated fault number in the example 3, using Algorithm 2 as a function of time

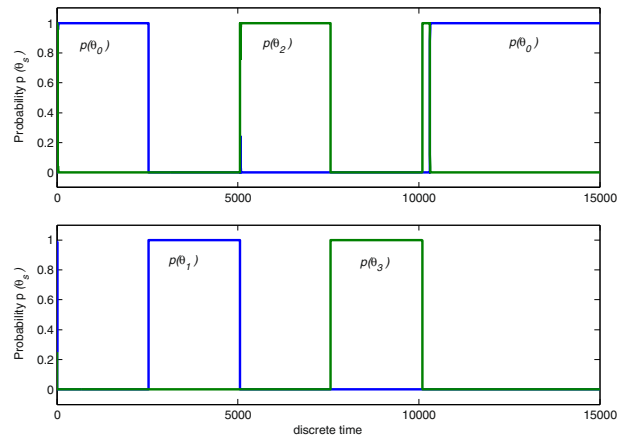


Fig. 7 Reconstructed probabilities of faults in example 3 estimated by the Algorithm 2

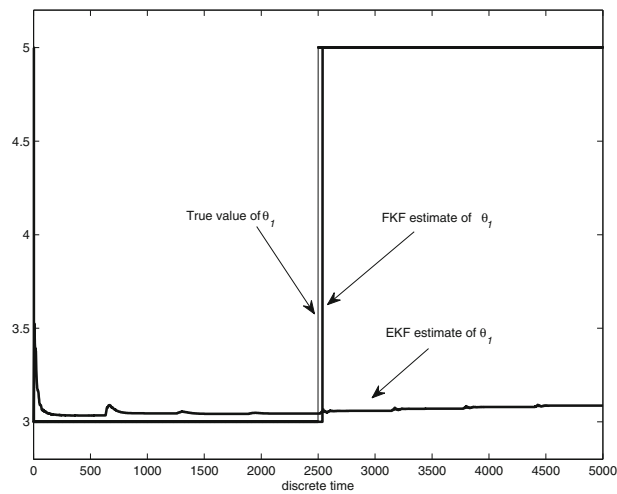


Fig. 8 Estimation of the oscillator's natural frequency with Extended Kalman Filter (EKF) and Field Kalman Filter (FKF). EKF does not converge

and other considered test systems. It should be however noted that they are not equivalent and their selection may require future work.

Future directions of work will include addressing the abovementioned issue, as well as implementation for more advanced systems. In particular, we want to extend the fault detection algorithm to the problems continuous in time. Additionally, the real-time performance of proposed algorithms will be tested in the laboratory environment systems.

Acknowledgements

We acknowledge this work under the scope of project titled "PRONTO: PROcess NeTwork Optimization". This project is financed by European commission's HORIZON 2020 /Marie Skłodowska-Curie/ITN/EID grant agreement no: 675215.

Authors' contributions

All the authors have participated in writing the manuscript and have revised the final version. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 31 January 2017 Accepted: 6 November 2017

Published online: 28 November 2017

References

1. D Alspach, A parallel filtering algorithm for linear systems with unknown time varying noise statistics. *IEEE Trans. Autom. Control.* **19**(5), 552–556 (1974). doi:10.1109/TAC.1974.1100645
2. R Kashyap, Maximum likelihood identification of stochastic linear systems. *IEEE Trans Autom Control.* **15**(1), 25–34 (1970). doi:10.1109/TAC.1970.1099344
3. BM Åkesson, JB Jørgensen, NK Poulsen, SB Jørgensen, in *17th European Symposium on Computer Aided Process Engineering, Computer Aided Chemical Engineering*, ed. by V Pleşu, PŞ Agachi. A tool for kalman filter tuning, vol. 24 (Elsevier, 2007), pp. 859–864. doi:10.1016/S1570-7946(07)80166-0. <http://www.sciencedirect.com/science/article/pii/S1570794607801660>. Accessed 15 Nov 2017
4. B Carew, P Belanger, Identification of optimum filter steady-state gain for systems with unknown noise covariances. *IEEE Trans Autom Control.* **18**(6), 582–587 (1973). doi:10.1109/TAC.1973.1100420
5. B Jiang, FN Chowdhury, Fault estimation and accommodation for linear mimo discrete-time systems. *IEEE trans control syst technol.* **13**(3), 493–499 (2005)
6. J-y Keller, DDJ Sauter, Kalman Filter for discrete-time stochastic linear systems subject to intermittent unknown inputs. *IEEE Trans Autom Control.* **58**(7), 1882–1887 (2013)
7. L Giovanini, R Dondo, A fault detection and isolation filter for discrete linear systems. *ISA Trans.* **42**(4), 643–9 (2003)
8. S Huang, KK Tan, TH Lee, Fault diagnosis and fault-tolerant control in linear drives using the Kalman filter. *IEEE Trans Ind Electron.* **59**(11), 4285–4292 (2012)
9. RH Chen, JL Speyer, A generalized least-squares fault detection filter. *Int. J. Adapt. Control Sig. Process.* **14**(7), 747–757 (2000)
10. GG Rigatos, A derivative-free Kalman filtering approach to state estimation-based control of nonlinear systems. *IEEE Trans. Ind. Electron.* **59**(10), 3987–3997 (2012)
11. W Li, SL Shah, D Xiao, Kalman filters in non-uniformly sampled multirate systems : For FDI and beyond. *Automatica.* **44**, 199–208 (2008). doi:10.1016/j.automatica.2007.05.009
12. P Li, V Kadiramanathan, Particle filtering based likelihood ratio approach to fault diagnosis in nonlinear stochastic systems. *IEEE Trans Syst Man Cybern Part C: Appl Rev.* **31**(3), 337–343 (2001)
13. D Ant6nio, T Yoneyama, Automatica A Bayesian solution to the multiple composite hypothesis testing for fault diagnosis in dynamic systems. *Automatica.* **47**(1), 158–163 (2011). doi:10.1016/j.automatica.2010.10.030
14. GH Foo, X Zhang, S Member, DM Vilathgamuwa, S Member, A sensor fault detection and isolation method in interior permanent-magnet synchronous motor drives based on an extended Kalman filter. *IEEE Trans. Ind. Electron.* **60**(8), 3485–3495 (2013)
15. T Kobayashi, DL Simon, Hybrid Kalman filter approach for aircraft engine in-flight diagnostics: sensor fault detection case. *J. Eng. Gas Turbines Power.* **129**(3), 746–754 (2006). doi:10.1115/1.2718572
16. MS Mahmoud, HM Khalid, Distributed Kalman filtering: a bibliographic review. *IET Control. Theory Appl.* **7**(4), 483–501 (2013). doi:10.1049/iet-cta.2012.0732. <http://ieeexplore.ieee.org/document/6544363/>. Accessed 15 Nov 2017
17. CC Drovandi, J McGree, AN Pettitt, A sequential monte carlo algorithm to incorporate model uncertainty in bayesian sequential design. *J. Comput. Graphical Stat.* **23**(1), 3–24 (2014)
18. L Martino, J Read, V Elvira, F Louzada, Cooperative parallel particle filters for online model selection and applications to urban mobility. *Digital Signal Process.* **60**, 172–185 (2017)
19. CM Carvalho, MS Johannes, HF Lopes, NG Polson, Particle learning and smoothing. *Statist. Sci.* **25**(1), 88–106 (2010)
20. C Andrier, A Doucet, R Holenstein, Particle markov chain monte carlo methods. *JR Statist. Soc. B.* **72**(3), 269–342 (2010)
21. I Urteaga, MF Bugallo, PM Djuric, Sequential Monte Carlo methods under model uncertainty. *IEEE Work. Stat. Sig. Process. Proc.* 1–5 (2016). doi:10.1109/SSP.2016.7551747
22. N Chopin, PE Jacob, O Papaspiliopoulos, Smc2: an efficient algorithm for sequential analysis of state-space models. *JR Statist. Soc. B.* **75**, 397–426 (2013)
23. P Bania, J Baranowski, in *2016 IEEE 55th Conference on Decision and Control (CDC)*. Field kalman filter and its approximation, (2016), pp. 2875–2880. doi:10.1109/CDC.2016.7798697
24. ET Jaynes, *Probability theory: the logic of science.* (Cambridge University Press, Cambridge, 2003)

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com