Mixed-Data Multi-Modelling for Fault Detection and Isolation

Miroslav Kárný[1], Ivan Nagy[2,1], Jana Novovičová[1,2]

[1]Institute of Information Theory and Automation, AVČR, P.O. Box 18, 182 08 Prague 8, Czech Republic, school@utia.cas.cz

[2]Faculty of Transportation Sciences, Czech Technical University in Prague, Konviktská 20 110 00 Prague 1, Czech Republic

Summary

Early recognition/isolation of a faulty behaviour of a dynamic system is the main task of a fault detection and isolation (FDI). FDI methods based on adaptive probabilistic models with multiple modes represent theoretically well justified way of solution. Their use is severely restricted by inherent computational complexity.

The complexity problem is addressed here by employing an efficient quasi-Bayes estimation algorithm. It is directly applicable to the mixture of components created as products of factors belonging to the exponential family. It opens a novel way to to deal adaptively with mixed continuous-discrete, dynamically-related data. The presented theory and algorithmisation are illustrated by a simple simulation example.

Key words

Fault detection and isolation, mixed mixture models, quasi-Bayes adaptive estimation

1 Introduction

Early recognition/isolation of a faulty behaviour of a dynamic system is the main task of a fault detection and isolation (FDI). Its importance has stimulated a lot of research activities that have provided a wide set of general as well as specific results. For reviews see [16, 7]. The development is definitely not finished yet. Complexity of the inspected systems and complexity of the formally optimal solutions call for a further progress. This paper contributes to a solution of the complexity aspect of FDI. It proposes a quasi-Bayesian estimation algorithm [8] that combats the "curse of dimensionality" connected with use of the multiple-model approach to FDI [4, 3]. It stems from the fact that the likelihood function needed for parameter estimation is a product of sums of functions of parameters, Consequently, it contains exactly unmanageable number of terms.

The proposed approach can be viewed as an extension of existing multiple-model approach to FDI, see for instance [9, 11, 2, 12]. The extensions are made in the following directions:

- mixed mixture models are used for a joint description of multivariate continuous and discrete signals that may also reflect other logical variables than faults,
- models are updated recursively so that detector is able to follow slow changes of the process,
- dynamics of fault occurrence, at least low-dimensional one, is explicitly modelled,
- a way to exploitation of a rich toolset, including structure estimation, quantification of prior knowledge etc., is opened by employing Bayesian problem formulation.

The main contributions of the paper are:

- adaptation of a recent quasi-Bayes algorithm for an efficient treatment of FDI problems,
- novel solution of on-line, possibly adaptive estimation of a dynamic mixture model that is able to cope both with continuous and discrete value data.

The paper is organised as follows. After introducing the necessary mathematical preliminaries the FDI problem is formulated in probabilistic terms, Section 2. Formal tools for description of multiple-model structure are presented in Section 3. Section 4 provides a formal Bayes solution of the addressed problem. The key Section 5 approximates the formal Bayesian estimation needed in FDI. The approximation is based on quasi-Bayes procedure ([15, 13] and the references therein). The result is then elaborated for practically significant and feasible exponential family of models [1], Section 6. A simple example of Section 7 illustrates properties of the proposed algorithm. It deals with two faulty states only but handles both discrete and continuous data, i.e. it handles mixed data mixture model. Finally, conclusions are drawn.

2 System description and problem formulation

In this section, some concepts and definitions concerning the fault behaviour of a dynamic system, the loss function and the type of probabilistic models, used throughout the paper, are introduced.

Notation
The notation generally follows the standard conventions of mathematics and statistics. x^* denotes the *set* with common elements x and $\bar{x}$ means *cardinality* of the set x^*. The term *probability density function* is used for both the continuous and discrete cases, abbreviating it in the usual way to *pdf*. The difference between a random variable and its realisation is omitted. The probability density function of x is denoted by $p(x)$ and the one of y as $p(y)$, even though the functions differ. The symbol $\int f(x)\,dx$ is possibly multivariate integral over all values of x for which the function $f(x)$ is nonzero and well defined. The sign $\propto$ denotes proportionality up to a normalising factor.

FDI scenario
The considered behaviour of the system can be described as follows. At each time $t \in t^* \equiv \{1, 2, \ldots, \bar{t}\}$, the system dynamics may exhibit one fault from a finite set of faults, say $f^* \equiv \{0, 1, \ldots, \bar{f} - 1\}$. As usual, the value $f = 0$ corresponds with a normal (non-faulty) behaviour of the system. The fault model assumes one-to-one mapping between faults and their labels. If multiple faults are expected, their possible configurations are given their special labels. At each time instance $t \in t^*$, some data d_t are observed on the system reflecting the current fault $f_t \in f^*$. The sequences of all vector-valued observations d_τ and all faults f_τ, $\tau = 1, \ldots, t$ are denoted by $d(t) \equiv \{d_1, \ldots, d_t\}$ and $f(t) \equiv \{f_1, \ldots, f_t\}$, respectively.

The main task, the paper deals with, is the fault isolation. It should estimate the true f_t at time t. Thus, a decision $\hat{f}_t \in f^*$ should be taken that classifies the behaviour of the system as exhibiting the fault $f_t = \hat{f}_t$. The early but not pre-mature decision is required. Thus, it can be based on observed data only, i.e. $\hat{f}_t = \hat{f}_t(d(t-1))$. A sequence of rules generating fault estimates $\hat{f}(\bar{t})$, is called the FDI strategy. An optimal FDI strategy is searched for.

Motivation for a model
Intuitively it is clear that a model relating the unknown f_t to the available data $d(t)$ is needed. Its form results from a formalisation of the notion "optimal".

The FDI detection strategy generating the estimates $\hat{f}^0(\bar{t})$, is said to be *optimal FDI strategy* if the expectation $\mathcal{E}$ of the loss function

$$L(f(\bar{t}), \hat{f}(\bar{t})) \equiv \sum_{t \in t^*} l(f_t, \hat{f}_t),$$

determined by a table of particular loss functions $l(f_t, \hat{f}_t) \geq 0,\ t \in t^*$, reaches its minimal value for $\hat{f}(\bar{t}) = \hat{f}^0(\bar{t})$.

It is known that the optimal FDI strategy can be constructed value-wise using dynamic programming technique. It leads to a backward recursive optimisation

$$V(d(t-1)) = \min_{\hat{f}_t \in \hat{f}^*} \mathcal{E}[l(f_t, \hat{f}_t) + V(d(t))|d(t-1), \hat{f}_t], \ t \in t^*, \ \ V(d(\bar{t})) = 0, \tag{1}$$

where the function $V(d(t))$ is known as the optimal cost-to-go.

As shown above, the construction of the optimal FDI strategy relies on the ability to evaluate the conditional expectations in the recursion (1). The probability density function $p(d_t, f_t|d(t-1), \hat{f}_t)$ for all $t \in t^*$ is required. This pdf is the marginal one of the joint pdf

$$p(d_t, F_t|d(t-1), \hat{F}_t). \tag{2}$$

Here, the *fault-configuration* F_t of the length $(n+1)$, $n \geq 0$ is introduced as a finite sequence of faults $F_t = \{f_t, f_{t-1}, \ldots, f_{t-n}\}$ to be able to deal with dynamics of f_t if necessary. Similarly, the fault-configuration estimate $\hat{F}_t = \{\hat{f}_t, \hat{f}_{t-1}, \ldots, \hat{f}_{t-n}\}$ of the same length $(n+1)$ is used.

Remark 1 *It is mostly sufficient to consider the small length n, $n \leq 2$ of fault-configurations F_t and their estimates $\hat{F}_t$ as faults have tendency to persist.*

A design of the algorithm for an efficient (approximate) on-line construction of the pdf (2) is addressed in this paper.

Remark 2 *Other FDI are also possible. For instance, the data item d_{t+i}, $i \geq 0$ can be allowed for estimating $\hat{f}_t$. The model (2) is the basic one even in these cases.*

Types of pdfs used

For constructing the needed pdf (2), relationships of observed data to faults and fault estimates have to be modelled. The desired pdf is then obtained through parameter estimation. It is easy to be lost in the variety of pdfs exploited. Their formal introduction should help to minimise this danger. Adopted assumptions and definitions related to the described system and the solved task are introduced, at the same time.

$\mathcal{A}1$. *The pdf predicting data d_t after observing data $d(t-1)$ and with given faults $f(t)$ as well as their estimates $\hat{f}(t)$ satisfies*

$$p(d_t|d(t-1), f(t), \hat{f}(t)) = p(d_t|d(t-1), F_t, \hat{F}_t).$$

The pdf $p(d_t|d(t-1), f(t), \hat{f}(t))$ satisfying the condition specified in $\mathcal{A}1$ is called *observation model.* The observation model is therefore fully determined by the observed data $d(t-1)$, the newest fault-configuration F_t and its estimate $\hat{F}_t$.

$\mathcal{A}2$. *The pdf predicting data d_t after observing the data $d(t-1)$ with given fault-configuration F_t, its estimate $\hat{F}_t$ and unknown parameters $\theta \in \theta^*$, $\alpha \in \alpha^*$ satisfies*

$$p(d_t|d(t-1), F_t, \hat{F}_t, \alpha, \theta) = p(d_t|\psi_t, F_t, \hat{F}_t, \theta),$$

where ψ_t denotes a finite fixed dimensional regression vector determined by the past data items $d(t-1)$. This regression vector depends possibly on the current system input u_t that is generated on the basis of the information contained in $d(t-1)$.

The pdf $p(d_t|d(t-1), F_t, \hat{F}_t, \theta)$ satisfying the condition $\mathcal{A}2$ is called *parameterised observation model.*

$\mathcal{A}3$. *The pdf predicting the fault-configuration F_t after observing data $d(t-1)$ and selecting the system inputs u_t and fault estimates $\hat{f}(t)$ satisfies*

$$p(F_t|d(t-1), u_t, \hat{f}(t)) = p(F_t|d(t-1)).$$

The pdf $p(F_t|d(t-1), \hat{f}(t))$ satisfying the condition specified in $\mathcal{A}3$ is called *fault-configuration model.*

This assumption that values of fault estimates $\hat{f}(t)$ add no information to $d(t-1)$ are naturally fulfilled by FDI strategies among which the optimal FDI strategy is searched for. The system inputs are just assumed to fulfill the same condition known in control community as natural conditions of control [14].

$\mathcal{A}4$. *The pdf predicting the fault-configuration F_t after observing data $d(t-1)$, fault estimates $\hat{f}(t)$ and having given by the unknown parameters $\theta \in \theta^*$, $\alpha \in \alpha^*$ satisfies*

$$p(F_t|d(t-1), \hat{f}(t), \alpha, \theta) = p(F_t|\hat{F}_t, \alpha).$$

The pdf $p(F_t|d(t-1), \hat{f}(t), \alpha, \theta)$ satisfying the condition specified in $\mathcal{A}4$ is called *parameterised fault-configuration model.*

Throughout the paper, the parameterised fault-configuration model of the time-invariant Markov type

$$p(F_t|d(t-1), \hat{f}(t), \alpha) = \alpha_{F|\hat{F}} \tag{3}$$

is used. The time-invariant array $[\alpha_{F|\hat{F}}]$ of transition probabilities α belongs to the probabilistic simplex

$$\alpha^* = \left\{ \alpha : \alpha_{F|\hat{F}} \geq 0, \sum_{F \in F^*} \alpha_{F|\hat{F}} = 1, \ \forall \hat{F} \in \hat{F}^* \right\}.$$

The joint pdf

$$p(\alpha, \theta, d_t, F_t|d(t-1), \hat{F}_t) \tag{4}$$

of all unobservable quantities, $\alpha \in \alpha^*$, $\theta = \{\theta\}_{F \in F^*}$, $F_t \in F^*$ and the unobserved d_t, given by observed data $d(t-1)$ and by the fault-configuration estimate $\hat{F}_t$ is called *joint model.*

Estimation of unknown parameters deals with the pdf

$$p(d_t|d(t-1), \hat{f}(t), \alpha, \theta) = p(d_t|d(t-1), \hat{F}_t, \alpha, \theta)$$

predicting data d_t after observing data $d(t-1)$ and fault-configuration estimates $\hat{F}_t$, given parameters θ, α but without a knowledge about F_t. It is called *parameterised data model.*

It follows from $\mathcal{A}2$, $\mathcal{A}4$ and from the computation formula for the marginal pdf that the parameterised data model has the mixture form

$$p(d_t|d(t-1), \hat{F}_t, \alpha, \theta) = \sum_{F_t \in F^*} p(d_t|d(t-1), F_t, \hat{F}_t, \theta)\alpha_{F_t|\hat{F}_t}. \tag{5}$$

This is the basic model used when solving the addressed problem, i.e. when evaluating the target pdf (2). The individual terms $p(d_t|d(t-1), F_t, \hat{F}_t, \theta)$ are parameterised observation models and they are called *(parameterised) components.*

3 Tools for description of multiple-model structure

Feasibility of the formulated task is significantly influenced by the form of (parameterised) components used. Here, they are supposed to belong to exponential family [1]. This family is known to possess finite-dimensional sufficient statistics. The existence of such a statistics is the necessary pre-requisite for on-line estimation on potentially infinite data sequences. This situation is typical for FDI. This makes us to restrict to this family that is rich enough so that almost any parameterised data model can be well approximated by it. The exponential family has the form $\exp[b'^F(\theta)c^F(\Psi)]$, where

$'$ denotes transposition,

b^F is a $\bar{b}^F$-dimensional column-vector function of parameters connected with the faulty configuration F,

c^F is a $\bar{c}^F = \bar{b}^F$-dimensional column-vector function of the data vector $\Psi_t = [d'_t, \psi'_t]$, see $\mathcal{A}2$.

Dimensions of vectors b^F, c^F depend generally on labels of components F and the components can further depend on some vector valued discrete data denoted at time t by l_t^F. This dependence can be expressed by decomposing these components to sub-models valid for particular configurations of the discrete data values. With respect to them, the statistics can be further structured. Here, formal tools for describing these structural properties are prepared.

Considered dimensions

$\bar{c}^F$ is a dimension parametric function belonging to the F-th component. This statistics can be structured according to the discrete data configurations.

$\bar{l}^F$ is a number of all possible configurations of the discrete data in the F-th component. Each of the data configurations has its own sub-statistics c_i^F, $i = 1, 2, ..., \bar{l}^F$.

Often, the number of various configurations of the discrete data is the same for all components. Then the superscript F is omitted and the number is denoted by $\bar{l}$.

$\bar{c}_i^F$ is a dimension of the sub-statistic c_i^F corresponding to the i-th configuration of the discrete data within the F-th component.

Obviously, it must hold: $\bar{b}^F = \sum_{i=1}^{\bar{l}^F} \bar{c}_i^F$ and $\bar{b} = \sum_{F=1}^{\bar{F}} \bar{b}^F = \sum_{F=1}^{\bar{F}} \sum_{i=1}^{\bar{l}^F} \bar{c}_i^F = \bar{c}$, where $b = [b'^1, \ldots, d'^{\bar{F}}]'$ is the vector function of parameters θ and $c = [c'^1, \ldots, c'^{\bar{F}}]'$ the vector function of data vector Ψ for the whole mixture model.

Structure vector
The decomposition of the vector b into the sub-vectors b^F is conveniently described by the *structure vector* ν_b

$$\nu_b = [\bar{b}^1, \bar{b}^2, ..., \bar{b}^{\bar{F}}]'. \tag{6}$$

It is a $\bar{F}$-dimensional vector consisting of positive integers – dimensions $\bar{b}^F,\ F \in F^*$ of particular vectors $b^F(\theta)$.

Structure multi-vector
For each component $F \in F^*$, the decomposition of the statistics $c^F = [c_1'^F, \ldots, c_{\bar{l}^F}'^F]$ into sub-vectors c_i^F corresponding to configurations $i = \{1, \ldots, \bar{l}^F\}$ of the discrete data occurring in Ψ is conveniently described by structure vectors ν^F

$$\nu^F = [\bar{c}_1^F, \bar{c}_2^F, ..., \bar{c}_{\bar{l}^F}^F]'. \tag{7}$$

It is a $\bar{\nu}^F$-dimensional vector of positive integers – dimensions $\bar{c}_i^F$ of the sub-vectors c_i^F corresponding to the i-th configuration of the discrete data within the F-th model component.

The overall structure describing all components for possible configurations of faults and discrete data is described by a *structure multi-vector* $\mu_c = [\nu'^1|\nu'^2|...|\nu'^{\bar{F}}]'$. Its expanded form is

$$\mu_c = \left[\bar{c}_1^1, \bar{c}_2^1, \ldots, \bar{c}_{\bar{l}^1}^1 \,|\, \bar{c}_1^2, \bar{c}_2^2, \ldots, \bar{c}_{\bar{l}^2}^2 \,|\, \ldots \,|\, \bar{c}_1^{\bar{F}}, \bar{c}_2^{\bar{F}}, \ldots, \bar{c}_{\bar{l}^{\bar{F}}}^{\bar{F}}\right]' \tag{8}$$

δ-vector
For selecting j-th item of an k-vector $q = [q_1, q_2, ..., q_k]'$ the *δ-vector* $\underline{\delta}(j)$

$$\underline{\delta}(j) = [\delta_{1,j}, \delta_{2,j}, ..., \delta_{k,j}]'\ ;\quad j = 1, 2, ..., k \tag{9}$$

is introduced. Here Kronecker delta $\delta_{i,j}$ is used that equals 1 if $i = j$ and zero otherwise. The j-th item of q can be then expressed as $q_j = q'\,\underline{\delta}(j)$.

$\underline{1}$-vector
For summing items of an k-vector $q = [q_1, q_2, ..., q_k]'$ the *vector of ones* $\underline{1}_k$ is introduced. It is a k-dimensional column vector with unit entries. With it, $\sum_{i=1}^{k} q_i = q'\,\underline{1}_k$.

A generalisation of the previously introduced symbol is a *block-vector of ones* $\underline{1}_\kappa$ with $\bar{\kappa}$-vector index $\kappa = [\kappa_1, \kappa_2, \ldots, \kappa_{\bar{\kappa}}]'$ of nonnegative integers. It is defined

$$\underline{1}_\kappa = [\underline{1}'_{\kappa_1}, \underline{1}'_{\kappa_2}, ..., \underline{1}'_{\kappa_{\bar{\kappa}}}]'. \tag{10}$$

This vector consists of blocks of ones with dimensions $\kappa_1, \kappa_2, \ldots, \kappa_{\bar{\kappa}}$.

I-matrix
The current values of faults and discrete data l are characterised by the structure multi-

vector μ_c. For selecting the corresponding model components, the *I-matrix*

$$I(\mu_c, r) = I([\nu'^1|\nu'^2|\cdot|\nu'^{\bar{F}}], r) = \mathrm{diag}\{\,\underline{1}'_{\nu^1} \odot r,\, \underline{1}'_{\nu^2} \odot r, ..., \,\underline{1}'_{\nu^{\bar{F}}} \odot r\} \tag{11}$$

is introduced, where

$\mu_c = [\nu'^1|\nu'^2|\cdot|\nu'^{\bar{F}}]'$ is a structure multi-vector,

$\underline{1}'_{\nu^F},\ F = 1, 2, \cdots, \bar{F}$ is a block-vector of ones with the block dimensions $\bar{\nu}_1^F, \bar{\nu}_2^F, \ldots, \bar{\nu}_{\bar{l}}^F$,

r is a real vector of dimension $\bar{r} = \bar{l}$, (the case $\bar{l}^1 = \bar{l}^2 = ... = \bar{l}^{\bar{F}} = \bar{l}$ is considered),

$\odot$ is so called Khatri-Rao product defined for block vectors q, r by means of Kronecker product $\otimes$ as follows

$$q \odot r = [q_1 \otimes r_1, q_2 \otimes r_2, ..., q_{\bar{q}} \otimes r_{\bar{q}}]. \tag{12}$$

Particular items in the diagonal of (11) are Khatri-Rao products of unit block-vectors $\underline{1}'_{\nu^F}$ with $\bar{l}$ blocks by (scalar) items of $\bar{l}$-dimensional vector r. Here the Kronecker product means multiplication of all items of a sub-block by a corresponding item of the vector r. The dimensions of the sub-blocks are $\bar{\nu}_j^F,\ j = 1, 2, \ldots, \bar{l}$, for $F = 1, 2, \ldots, \bar{F}$.

When dealing only with the structure of the parametric function b, which is induced by the faulty configurations F, the *I-matrix* is

$$I(\nu, r) = \mathrm{diag}\{\nu \odot r\} = \mathrm{diag}\{\,\underline{1}_{\nu_1} r_1,\, \underline{1}_{\nu_2} r_2, ..., \,\underline{1}_{\nu_{\bar{l}}} r_{\bar{l}}, \}. \tag{13}$$

where particular items represent multiplications of vectors $\underline{1}_{\nu_F} = \underline{1}_{\bar{b}^F}$, see (6), by scalars $r_F,\ F = 1, 2, \ldots, \bar{F}$.

Remark 3 *The δ-vector $\underline{\delta}$ of Kronecker delta functions is mainly substituted for the vector r. It selects a component of the mixture model. Substitution of $\underline{1}$ for r chooses all the components.*

With the introduced notation, the exponent of the F-th model component pdf, corresponding to i-th configuration of observed discrete data can be written

$$b_i'^F c_i^F = b'\ I(\nu_b,\, \underline{\delta}(F)) I(\mu_c,\, \underline{\delta}(i))\ c, \qquad \text{where} \tag{14}$$

ν_b and μ_c are structure vector of the parametric function b and the structure multi-vector of the data function c, respectively,

b and c are vector functions of parameters θ and vector functions of data statistics composed of data from Ψ ordered according to their structure vectors,

$\underline{\delta}(F)$ and $\underline{\delta}(i)$ are δ-vectors choosing proper b^F and c_i^F for constructing the model of the active component labeled by ^{F} for actually observed discrete data configuration labeled by $_i$.

Remark 4 *The product of I-matrices $I(\nu_b,\, \underline{1}) I(\mu_c,\, \underline{\delta}(i))$ chooses all sub-models corresponding to discrete data configuration i and all components $F \in F^*$.*

4 Bayesian estimation

This section presents formal relationships of the target pdf (2) $p(d_t, F_t|d(t-1), \hat{F}_t)$ to the underlying pdfs and observed data. The resulting formulas are converted into a practicable algorithm in Sections 5 and 6.

Constructing $p(d_t, F_t|d(t-1), \hat{F}_t)$
The investigated pdf (2) can be expressed by using the joint model (4) as

$$\begin{aligned}
p(d_t, F_t|d(t-1), \hat{F}_t) &= \int\int p(\alpha, \theta, d_t, F_t|d(t-1), \hat{F}_t)\, d\alpha\, d\theta = \qquad (15)\\
&= \int\int p(d_t, F_t|d(t-1), \hat{F}_t, \alpha, \theta) f(\alpha, \theta|d(t-1), \hat{F}_t)\, d\alpha\, d\theta = \\
&= \int\int p(d_t|d(t-1), F_t, \hat{F}_t, \alpha, \theta) p(F_t|d(t-1), \hat{F}_t, \alpha, \theta) f(\alpha, \theta|d(t-1), \hat{F}_t)\, d\alpha\, d\theta \underbrace{=}_{\mathcal{A}2, \mathcal{A}3, \mathcal{A}4} \\
&= \int p(d_t|d(t-1), F_t, \hat{F}_t, \theta) p(\theta|d(t-1))\, d\theta \int p(F_t|\hat{F}_t, \alpha) p(\alpha|d(t-1))\, d\alpha = \\
&= p(d_t|d(t-1), F_t, \hat{F}_t) p(F_t|d(t-1), \hat{F}_t).
\end{aligned}$$

In the formula (15), assumptions $\mathcal{A}2$, $\mathcal{A}3$, $\mathcal{A}4$ and Fubini theorem are exploited together with identities $\int f(\alpha, \theta|d(t-1))\, d\alpha = f(\theta|d(t-1))$ and $\int f(\alpha, \theta|d(t-1))\, d\theta = f(\alpha|d(t-1))$.

Moreover, the integral

$$\int p(d_t|d(t-1), F_t, \hat{F}_t, \theta)\, p(\theta|d(t-1), F_t, \hat{F}_t) d\theta = p(d_t|d(t-1), F_t, \hat{F}_t) \qquad (16)$$

represents the observation model for the given fault-configuration F_t, cf. assumption $\mathcal{A}1$, and

$$\int p(F_t|\hat{F}_t, \alpha)\, p(\alpha|d(t-1), \hat{F}_t) d\alpha = p(F_t|d(t-1), \hat{F}_t) \qquad (17)$$

is the fault-configuration model, cf. assumption $\mathcal{A}3$.

Parameter estimation
Marginal pdfs $p(\theta|d(t-1))$, $p(\alpha|d(t-1))$ of the posterior pdf of unknown parameters $p(\alpha, \theta|d(t-1))$ are only used in (15), see (16 and (17). They are obtained through the Bayesian parameter estimation, i.e. by updating the chosen prior pdf $p(\alpha, \theta)$ according to the Bayes formula

$$p(\alpha, \theta|d(t)) = \frac{p(d_t|d(t-1), \hat{F}_t, \alpha, \theta)\, p(\alpha, \theta|d(t-1))}{p(d_t|d(t-1), \hat{F}_t)}. \qquad (18)$$

Remark 5 *The form of the mixture model (5) implies that the joint pdf $p(\alpha, \theta|d(t))$ has to be updated in spite of the fact that its marginal pdfs are needed only in (16), (17) determining the needed pdf (2).*

The mixture form of the parameterised observation model (5) in the Bayes formula (18) makes on-line exact use of (18) impossible. In other words, the computational difficulties, which are faced and solved approximately in the subsequent sections, stem from the *fact that the parameterised data model does not possess natural conjugate. Consequently, its functional form does not reproduce within a feasible set while estimating unknown parameters α, θ.*

Data prediction
The data predictor $p(d_t|d(t-1), \hat{F}_t)$ exploiting only observed data can be obtained by summation the pdf (15) over all fault configurations from the set F^* of the faulty configurations F

$$p(d_t|d(t-1), \hat{F}_t) = \sum_{F_t \in F^*} p(d_t, F_t|d(t-1), \hat{F}_t).$$

Fault-configuration prediction and estimation
The model (17) provides prediction of fault configuration that is needed for considered design of "predictive" FDI strategies. As mentioned in Remark 2, other strategies differing in speed of the fault-isolation response can be considered. For instance, estimation of the current fault-configuration F_t can be based on data $d(t)$, i.e. with the current data d_t included. The pdf needed for an optimal Bayesian estimation can also be constructed using the Bayes formula. It holds

$$p(F_t|d(t), \hat{F}_t) = p(F_t|d_t, d(t-1), \hat{F}_t) = \frac{p(d_t|d(t-1), F_t, \hat{F}_t)p(F_t|d(t-1), \hat{F}_t)}{p(d_t|d(t-1), \hat{F}_t)}. \tag{19}$$

The denominator in (19) is independent of F_t, thus (19) can be expressed in the alternative instructive way

$$p(F_t|d(t), \hat{F}_t) \propto p(d_t|d(t-1), F_t, \hat{F}_t)p(F_t|d(t-1), \hat{F}_t) = p(d_t, F_t|d(t-1), \hat{F}_t). \tag{20}$$

It says that $p(F_t|d(t), \hat{F}_t)$ is proportional to the pdf (2).

Remark 6 *The pdf (16) and/or (17) are available if parameters θ and/or α are known, i.e. if pdfs $p(\theta|d(t-1), \hat{F}_t)$ and $p(\alpha|d(t-1), \hat{F}_t)$ concentrate of known values θ, α.*

Recall, that the aim is to *compute the pdf (2). The computation is unfeasible. Thus, it will be done approximately in the way applicable in on-line mode.* An algorithm for solving this problem is described in the next section.

5 Quasi-Bayes estimation for known observation model

First, the simpler case when the parameterised observation model contains no unknown parameter θ is considered. Various formulae below are simplified by using the δ-vector $\underline{\delta}$ defined by (9) and vector of ones $\underline{1}$ of an appropriate dimension.

According to (17) and (15), the needed pdf (2) is

$$p(d_t, F_t|d(t-1), \hat{F}_t) = p(d_t|d(t-1), F_t, \hat{F}_t) \int p(F_t|\hat{F}_t, \alpha)p(\alpha|d(t-1))\, d\alpha. \tag{21}$$

The pdf $p(\alpha|d(t-1))$ of the unknown α have to be updated in order to get the desired (approximately) Bayesian solution (18) of the considered problem.

α-prior pdf
Let the prior pdf for α has the form of Dirichlet density [6]

$$p(\alpha|d(t-1),\hat{F}_t) \propto \prod_{F\in F^*} \prod_{\hat{F}\in\hat{F}^*} \alpha_{F|\hat{F}_t}^{n_{F|\hat{F}_t}(t-1)-1}, \tag{22}$$

determined by the $(\bar{F})^2 \equiv (\bar{f}+1)^{2(n+1)}$-way table of positive statistics $n_{F|\hat{F}}(t-1)$. It is known to be natural conjugate models describing discrete-valued random quantities when fault-configuration model is parameterised by the values of transition probabilities [6]. This pdf can be written in the exponential form

$$p(\alpha|d(t-1),\hat{F}_t) \propto \exp\{\text{tr } \lambda'(\alpha)[n(t-1) - \underline{1}\,\underline{1}']\}, \qquad \text{where}$$

$\lambda(\alpha)$ is the $(\bar{F},\bar{F})$-dimensional matrix function of α with entries

$$\lambda_{F|\hat{F}} = \ln(\alpha_{F|\hat{F}}),\ F\in F^*,\ \hat{F}\in F^* \tag{23}$$

$n(t)$ is $(\bar{F},\bar{F})$-dimensional matrix of statistics with entries $n_{F|\hat{F}}(t)$ "counting" the number of occurrences of the configurations $F_\tau|\hat{F}_\tau$ for $\tau \le t$.

Parameterised fault-configuration model
The parameterised fault-configuration model can be given the same multiplicative form as the α-prior pdf

$$p(F_t|\hat{F}_t,\alpha) = \prod_{F\in F^*} \prod_{\hat{F}\in\hat{F}^*} \alpha_{F|\hat{F}}^{\delta_{F,F_t}\delta_{\hat{F},\hat{F}_t}}. \tag{24}$$

Using (9), the model (24) can be rewritten in the form

$$p(F_t|\hat{F}_t,\alpha) = \exp\{\,\underline{\delta}'(F_t)\lambda(\alpha)\,\underline{\delta}(\hat{F}_t)\} = \exp\{\text{tr } \lambda'(\alpha)\,\underline{\delta}(F_t)\,\underline{\delta}'(\hat{F}_t)\}. \tag{25}$$

Computing the fault-configuration prediction
After evaluating the integral in (21) by using relations (22) and (24), the target pdf (2) is obtained in the following form

$$p(d_t,F_t|d(t-1),\hat{F}_t) = p(d_t|d(t-1),F_t,\hat{F}_t)\frac{\underline{\delta}'(F_t)n(t-1)\,\underline{\delta}(\hat{F}_t)}{\underline{1}'n(t-1)\,\underline{\delta}(\hat{F}_t)}. \tag{26}$$

Updating the α-prior pdf
By the Bayes chain formula [14] it holds

$$p(\alpha,F_t|d(t),\hat{F}_t) \propto p(d_t|d(t-1),F_t,\hat{F}_t)\exp\{\text{tr } \lambda'(\alpha)[n(t-1) - \underline{1}\,\underline{1}' + \underline{\delta}(F_t)\,\underline{\delta}'(\hat{F}_t)]\}, \tag{27}$$

where the δ-vector $\underline{\delta}(F_t)$ on the right-hand side of (27) is the only object that prevents updating because the true value of F_t is not observed. Its formally correct removal by computing marginal pdf provides the constructed posterior pdf of a mixture form. Its further updating by new data would increases the number of terms to be stored too quickly. Thus, some approximation is needed.

Quasi-Bayes approximation
The adopted quasi-Bayes estimation [8] simply replaces δ_{F,F_t} by its expectation

$$\delta_{F,F_t} \approx \mathcal{E}[\delta_{F,F_t}|d(t), \hat{F}_t] = p(F_t = F|d(t), \hat{F}_t) \equiv w_F(t), \tag{28}$$

i.e. by the pdf $p(F_t|d(t), \hat{F}_t)$ given in (17) and proportional to pdf (26) with the measured $d(t)$ and chosen $\hat{F}_t$.

Probability weights
Using the approximation (28), the posterior pdf of α is again the Dirichlet density with the matrix $n(\cdot)$ evolving according to the simple formula

$$n(t) = n(t-1) + w(t)\,\underline{\delta}'(\hat{F}_t). \tag{29}$$

The F-th entry of the $\bar{F}$-vector $w(t)$ contains probability of the fault configuration $F_t = F$ conditioned on the data $d(t)$ and the fault-configuration estimate $\hat{F}_t$. Its entries are given by the formula

$$w_F(t) = p(F_t = F|d(t), \hat{F}_t) \propto p(d_t|d(t-1), F_t = F, \hat{F}_t)\,\underline{\delta}'(F_t)n(t-1)\,\underline{\delta}(\hat{F}_t). \tag{30}$$

It is implied by equations (21) and (26).

Note that

$$p(F_t = F|d(t-1), \hat{F}_t) = \frac{\underline{\delta}'(F_t)n(t-1)\,\underline{\delta}(\hat{F}_t)}{\underline{1}'n(t-1)\,\underline{\delta}(\hat{F}_t)} \tag{31}$$

is the prediction of the fault-configuration $F_t = F$. It can be used for constructing FDI strategy when the dynamic nature of (1) can be neglected and the loss function cannot be specified in detail.

The vector $w(t)$ consists of probabilities of fault configurations F_t when even d_t is included in the condition. These probabilities are generally better indicators of the occurrence of fault configurations. If the selection of the fault-configuration estimate can be postponed after observing d_t then these weights should be used for the construction of $\hat{F}_t(d(t))$.

Practical usefulness of the final formulas (21), (29) and (30) depends heavily on our ability to provide the observation model $p(d_t|d(t-1), F_t = F, \hat{F}_t)$ indexed by the arbitrary fault-configuration $F_t = F$. This problem is addressed in the subsequent section.

6 Quasi-Bayes estimation for exponential family

The observation model $p(d_t|d(t-1), F_t, \hat{F}_t)$ is rarely available from a direct modeling. Thus, they must be learnt using observed data. Input-output models are considered here so that learning reduces to parameter estimation. The presented results can be extended to state space models belonging to Daume's class that guarantees feasibility of the state estimation problem [5]. Discussion of this case is, however, out of scope of this paper.

Observation model
Now, the unknown observation model (16) is assumed. It is computed by integration of the product of parameterised observation model and the Bayesian parameter estimation that provides the pdf $f(\theta|d(t-1))$.

For control application, it is reasonable to split data items d_t inputs u_t (a directly manipulable part) and outputs y_t, i.e. $d_t = [y_t, u_t]$. Using chain rule for pdfs, the parameterised observation model can be expressed as the following product

$$\begin{aligned} p(d_t|d(t-1), F_t, \hat{F}_t, \theta) &= p(u_t|d(t-1), \hat{F}_t, \theta) p(y_t|u_t, d(t-1), F_t, \hat{F}_t, \theta) = \\ &= p(u_t|d(t-1), \hat{F}_t) p(y_t|\psi_t, F_t, \hat{F}_t, \theta). \end{aligned} \tag{32}$$

The second equality is implied by the adopted natural conditions of control, see $\mathcal{A}3$, [14]. They are met whenever the input u_t is generated without additional knowledge of θ and F_t in addition to that knowledge contained in $d(t-1)$. Consequently, the sole information about input u_t adds nothing to information carried by data $d(t-1)$ about θ. It is reflected in omission of θ in the first factor of the product (32). Thus, the parameterised model $p(y_t|\psi_t, F_t, \hat{F}_t, \theta)$ has to be chosen only. It is taken from *exponential family* because of reasons outlined in Section 3. Thus, the following exponential form is considered for each $F \in F^*$

$$p(y_t|\psi_t, F_t, \hat{F}_t, \theta) = \exp\{b'^F(\theta) c^F(\Psi_t)\}, \tag{33}$$

where

$b^F(\theta)$ is a vector function of parameters relevant to the F-th fault configuration,

$c^F(\Psi_t)$ is a vector function of data vector $\Psi_t = [y_t, \psi_t']'$ relevant to F-th fault configuration.

For a compact description, I-matrices introduced in (13) are exploited. The parameterised observation model with $\bar{F}$ components from the exponential family is characterised by the vector function $b(\theta) = b = [b'^1, b'^2, ..., b'^{\bar{F}}]'$. It is partitioned into $\bar{F}$ blocks. Each block is allowed to have different dimension and nature of the involved parameters. Its data dependent counterpart is $c(\Psi) = c = [c'^1, c'^2, ..., c'^{\bar{F}}]'$. Generally, arguments Ψ of individual $c^F(\Psi)$ differ in dimension and nature, too.

With the introduced vector functions $b(\theta), c(\Psi)$, the considered parameterised observation model becomes

$$p(y_t|\psi_t, F_t, \hat{F}_t, \theta) = \exp\{b'(\theta) I(\nu_b, \underline{\delta}(F)) c(\Psi_t)\}, \tag{34}$$

where

$b(\theta)$ is a $\bar{b}$-dimensional known vector function of θ composed of particular functions of parameter b^F for $F \in F^*$,

$I(\cdot,\cdot)$ is I-matrix defined by (13),

ν_b is the structure vector introduced in (6),

$\underline{\delta}(F)$ is the δ-vector defined by (9),

$c(\Psi_t)$ is a known $\bar{c}$-dimensional vector function of the data vector Ψ_t composed of particular functions c^F for $F \in F^*$,

Ψ_t is data vector that can be recursively evaluated using the observed data $d(t)$ and the fault-configuration estimate $\hat{F}_t$.

(α, θ)-prior pdf
Dirichlet pdf of α reproduces with adopted quasi-Bayes approximation. It indicates that proper candidates for "self-reproducing" class of pdfs of the unknown parameters (α, θ) should have the form

$$p(\alpha, \theta | d(t-1), \hat{F}_t) \propto \exp\{\mathrm{tr}\ \lambda'(\alpha)(n(t-1) - \underline{1}\,\underline{1}')\} \exp\{b'(\theta) C(t-1)\}. \qquad (35)$$

The factor related to parameters θ is the natural conjugate pdf related to the model (34). It is determined by a $(\bar{b}, \bar{c})$-dimensional statistics $C(t-1)$. It contains the prior information and further one brought by the past observed data $d(t-1)$. With it, the prediction and the data update of statistics can be performed similarly as in the previous section.

Computing the fault-configuration prediction

$$\begin{aligned} &p(d_t, F_t | d(t-1), \hat{F}_t) \propto \\ &\quad \propto \int \exp\left\{b'(\theta)\left[C(t-1) + I(\nu_b, \underline{\delta}(F)) c(\Psi_t)\right]\right\} d\theta\, \frac{\underline{\delta}'(F_t) n(t-1)\, \underline{\delta}(\hat{F}_t)}{\underline{1}' n(t-1)\, \underline{\delta}'(\hat{F}_t)}. \end{aligned} \qquad (36)$$

Updating Bayesian estimates of (α, θ)
The Bayes rule and $\mathcal{A}3$ imply

$$\begin{aligned} p(\alpha, \theta, F_t | d(t), \hat{F}_t) \ \propto\ & \exp\{\mathrm{tr}\ \lambda'(\alpha)[n(t-1) - \underline{1}\,\underline{1}' + \underline{\delta}(F_t)\, \underline{\delta}'(\hat{F}_t)]\} \\ & \exp\{b'(\theta)[C(t-1) + I(\nu_b, \underline{\delta}(F)_t) c(\Psi_t)]\}. \end{aligned} \qquad (37)$$

Quasi-Bayes approximation

Again, the correct removal of the unknown values of $\underline{\delta}(F_t)$ would lead to computational burst. For this reason, the approximation

$$\delta_{F,F_t} \approx p(F_t = F | d(t), \hat{F}_t)$$

for the known $d(t)$ and $\hat{F}_t$ is used. It can be computed as the marginal pdf of (37).

Probability weights

$$w_F(t) = p(F_t = F|d(t), \hat{F}_t) \propto \mathcal{I}(I(\nu_b,\, \underline{\delta}(F))[C(t-1) + c(\Psi_t)])\, \underline{\delta}'(F)(n(t-1) + \underline{1}\,\underline{1}')\, \underline{\delta}(\hat{F}_t) \tag{38}$$

with

$$\mathcal{I}(I(\nu_b,\, \underline{\delta}(F))C) \equiv \int \exp[b'(\theta)I(\nu_b,\, \underline{\delta}(F))C]\, d\theta^F.$$

Substituting the approximation of the weight (38) into the I-matrix $I(\nu_b,\, \underline{\delta}(F))$ in (37), the form of the approximate parameter distribution is reproduced. The key influence of the approximation adopted is that the increments $c(\Psi_t)$ of the statistics $C(t-1)$ are added with the weight (38). Having the (approximate) pdf $p(\theta|d(t), \hat{F}(t))$ the needed observation model can be simply computed. The results found can be summarized. In the final algorithm, preparatory and on-line stages are distinguished.

Quasi-Bayes estimation in multiple-model exponential family

Preparatory stage

- Choose the parameterised models, i.e. specify the form of the functions $b(\theta)$, $c(\Psi)$ determining the members of the exponential family (33) and the form of the data vector Ψ.
- Specify the I-matrix $I(\nu_b,\, \underline{\delta}(F_t))$ determining structure of the mixture model defined in Section 3.
- Specify initial values of the statistics $C(0)$, see (35).
- Specify the "counters" $n_{F|\hat{F}}(0) > 0$ for all $F \in F^*$ and $\hat{F} \in \hat{F}^*$, see (35).
- Set time $t = 1$ and select the first input u_1.

On-line stage

1. Compute, if need be, the characteristics of the pdf (38),

$$p(y_t, F_t|\psi_t, \hat{F}_t) \propto \mathcal{I}\left(I(\nu_b,\, \underline{\delta}(F))(C(t-1) + c(\Psi_t))\right)\;\; \underline{\delta}'(F_t)n(t-1)\, \underline{\delta}(\hat{F}_t).$$

2. Select the point prediction $\hat{F}_t(d(t-1))$ of the fault configuration F_t.
3. Observe the output y_t and evaluate the current value of the vector Ψ_t, see (33).
4. Make "trial" updating of the following statistics

$$\begin{aligned} \tilde{C}(t) &= C(t-1) + I(\nu_b,\, \underline{1})c(\Psi_t) \\ \tilde{n}(t) &= n(t-1) + \underline{1}\, \underline{\delta}'(\hat{F}_t). \end{aligned}$$

5. Compute the weights $w_F(t)$ for all $F \in F^*$ by the formula

$$w_F(t) = \frac{\mathcal{I}(I(\nu_b, \underline{\delta}(F_t))\tilde{C}(t))\ \underline{\delta}'(F_t)\tilde{n}(t)\,\underline{\delta}(\hat{F}_t)}{\sum_{F \in F^*} \mathcal{I}(I(\nu_b, \underline{\delta}(F))\tilde{C}(t))\ \underline{\delta}'(F)\tilde{n}(t)\,\underline{\delta}(\hat{F}_t)}, \quad (38).$$

6. Select, if need be, the filtered estimate $\hat{F}(d(t))$ of F_t.
7. Compute, for the chosen fault prediction $\hat{F}_t$, the statistics

$$\begin{aligned} C(t) &= C(t-1) + I(\nu_b, w(t))c(\Psi_t) \\ n(t) &= n(t-1) + w(t)\,\underline{\delta}'(\hat{F}_t). \end{aligned}$$

8. Compute, if need be, the characteristics of the pdf (18)

$$p(\alpha, \theta | d(t), \hat{f}(t)) \propto \exp\{\mathrm{tr}\ \lambda'(\alpha)(n(t) - \underline{1}\,\underline{1}')\}\exp\{b'(\theta)C(t)\}.$$

9. Increase the time counter, choose the new system input u_t and return to the step 1.

It is worth noticing that the presented algorithm for the solution of the stated problem can be easily extended to the case of slowly varying parameters by adopting stabilised forgetting [10]. Then the recursions for the statistics $n(t)$, $C(t)$ are modified using forgetting factors $\varphi \in [0, 1]$ as follows

$$\begin{aligned} C(t) &= \varphi(C(t-1) + I(\nu_b, w(t))c(\Psi_t)) + (1-\varphi)C(0) \\ n(t) &= \varphi(n(t-1) + w(t)\,\underline{\delta}'(\hat{F}_t)) + (1-\varphi)n(0). \end{aligned}$$

7 Example – multiple fault model with mixed data

In this section, a particular case of the parameterised observation model is elaborated in detail in order to provide insight into the presented algorithm. The example can be interpreted as an (over)simplified version of fault detection on a controlled rolling mill. The controlled output y_t is the thickness of the produced metal sheet. At the time t, the controlled system can exhibit one of the two possible faults

$$f_t = 0 \text{ (no fault)}, \ \ f_t = 1 \text{ (fault)}, \ \ f^* = \{0, 1\}, \ \ \bar{f} = 2.$$

An operator observes the system and estimates its behaviour. He can choose his estimate from the set $\hat{f} \in f^* = \{0, 1\}$, $\bar{f} = 2$, of all possible faults

$$\hat{f}_t = 0 \text{ when no fault is expected}, \ \hat{f}_t = 1 \text{ when fault is expected}, \ \bar{\hat{f}} = 2.$$

The operator switches between two types of control actions $u \in u^* = \{0, 1\}$, $\bar{u} = 2$ in dependence on his estimate $\hat{f}_t$ of f_t.

$u_t = 0$ means that feedback is switched on or $u_t = 1$ means that feedback is switched off.

The continuous output y_t and discrete input u_t form the mixed data pair $d_t = [y_t, u_t]$. For sake of simplicity, the following assumptions are adopted.

$\mathcal{A}5$ *Fault configuration has length 1, $n = 0$, $F_t = f_t$, $\hat{F}_t = \hat{f}_t$, and system output y_t is normally distributed with moments depending just on fault f_t and the chosen action u_t.*

Parameterised observation model

$$p(d_t|d(t-1), F_t, \hat{F}_t, \alpha, \theta) \underbrace{=}_{\mathcal{A}2,\mathcal{A}5} p(d_t|f_t, \theta) \underbrace{=}_{\mathcal{A}3} p(y_t|u_t, f_t, \theta)p(u_t|\hat{f}_t) = \mathcal{N}_{y_t}(m_{u_t|f_t}, r_{u_t|f_t})p(u_t|\hat{f}_t)$$

where $\mathcal{N}_y(m, r)$ stands for normal pdf with mean m and variance r, i.e.

$$\begin{aligned} \mathcal{N}_y(m,r) &= (2\pi r)^{-0.5} \exp\left\{-0.5r^{-1}(y-m)^2\right\} = \qquad (39)\\ &= \exp\left\{-0.5[r^{-1}, 2r^{-1}m, (m^2 + \ln(2\pi r))][y^2, y, 1]'\right\},\\ \Psi &= [y, \psi],\ \psi = 1,\\ \theta &= [m_{u|f}, r_{u|f}],\ m_{u|f} \text{ is real}, r_{u|f} > 0,\ u \in u^*,\ f \in f^* \qquad (40) \end{aligned}$$

The two-way table $p(u_t|\hat{f}_t)$ describes operator strategy. It contains probabilities with which the operator switches on or off the feedback. These actions are driven by the current fault estimate. The time-invariant strategy is simulated with $p(u_t|\hat{f}_t)$ given by the following table

$p(u_t\|\hat{f}_t)$	no fault $\hat{f}_t = 0$	fault $\hat{f}_t = 1$
feedback $u_t = 0$	0.6	0.2
no feedback $u_t = 1$	0.4	0.8

Such a table may reflect recommended operator actions.

Parameterised fault-configuration model
In this example, the parameterised fault-configuration model (3) reduces to $p(f_t|\hat{f}_t, \alpha) = \alpha_{f_t|\hat{f}_t}$. It is again a two-way table of probabilities $\alpha_{f_t|\hat{f}_t}$

	no fault $\hat{f}_t = 0$	fault $\hat{f}_t = 1$
no fault $f_t = 0$	$\alpha_{0\|0}$	$\alpha_{0\|1}$
fault $f_t = 1$	$\alpha_{1\|0}$	$\alpha_{1\|1}$

These probabilities are unknown and have to be estimated. The table represents (fictitious) fault occurrences on the mill. This interpetation is qualitatively reflected in specific values used in simulation $\alpha_{0|0} = 0.95$, $\alpha_{1|1} = 0.8$, $\alpha_{1|0} = 0.05$, $\alpha_{0,1} = 0.2$.

The parameterised fault-configuration model (25) can be given the form

$$p(f_t|\hat{f}_t, \alpha) = \exp\{\text{tr }\underline{\delta}'(f_t)\lambda(\alpha)\,\underline{\delta}(\hat{f}_t)\}, \quad \text{with}$$

$$\lambda(\alpha) = \begin{bmatrix} \ln(\alpha_{1|1}) & \ln(\alpha_{1|2}) \\ \ln(1-\alpha_{1|1}) & \ln(1-\alpha_{1|2}) \end{bmatrix}, \quad \underline{\delta}(f_t) = \begin{bmatrix} \delta_{0,f_t} \\ \delta_{1,f_t} \end{bmatrix}, \qquad \underline{\delta}(\hat{f}_t) = \begin{bmatrix} \delta_{0,\hat{f}_t} \\ \delta_{1,\hat{f}_t} \end{bmatrix}.$$

Estimate of α

The pdf describing knowledge on parameter α (22) is assumed to be Dirichlet pdf. It is presented here in the form stressing that it belongs to exponential family

$$p(\alpha|d(t-1),\hat{f}_t) \propto \prod_{f\in f^*}\prod_{\hat{f}\in\hat{f}^*} \alpha_{f|\hat{f}_t}^{n_{f|\hat{f}_t}(t-1)-1} = \exp\{\mathrm{tr}\ \lambda'(\alpha)(n(t-1)-\underline{1}\,\underline{1}')\}.$$

The statistics $n(t-1)$ determining it form the matrix $n(t-1) = \begin{bmatrix} n_{1,1}(t-1) & n_{1,2}(t-1) \\ n_{2,1}(t-1) & n_{2,2}(t-1) \end{bmatrix}$.

In simulations, $n(0) = \begin{bmatrix} 500 & 50 \\ 100 & 150 \end{bmatrix}$ was used.

Structure of parameterised observation model

The parameterised observation model belongs to exponential family. Its behaviour depends both on the fault $f \in \{0,1\}$ and the discrete control variable $u \in \{0,1\}$. This dependence induces decomposition of components indexed by f to sub-models each of them valid for a particular value of the discrete input u_t. In the considered case, the function c of the data vector $\Psi = [y,1]'$ is

$$\begin{aligned} c &= [c_0^{'0}, c_1^{'0}, c_0^{'1}, c_1^{'1}]' \quad \text{with} \\ c_u^f(y) &= \left[y^2,\, y,\, 1\right]', \quad u,f \in u^* = f^* = \{0,1\}. \end{aligned} \tag{41}$$

The structure vectors $\nu^f = \nu^F$ (7) serve for describing it, see Section 3. For each fault configuration $F = f \in f^* = \{0,1\}$, ν^f is defined

$$\nu^f = [\bar{c}_0^f, \bar{c}_1^f]' = [3,3]', \quad \text{where}$$

$\bar{c}_u^f = 3,\ u \in \{0,1\}$ are dimensions of statistics corresponding to the u-th configuration of discrete data within the f-th component. The structure multi-vector μ_c defined in (8) is composed of the structure vectors ν^f. In this example, it is

$$\mu_c = [\nu^{'0}, \nu^{'1}]' = \left[[\bar{c}_0^0, \bar{c}_1^0], [\bar{c}_0^1, \bar{c}_1^1]\right]' = [[3,3],[3,3]]'$$

reflecting the considered case with common $c(\Psi) = c(y) = c_u^f$ for all f and u.

The composed function $b = b(\theta)$ of parameters is

$$\begin{aligned} b &= [b^{'0}, b^{'1}]' = \left[[b_0^0, b_1^0], [b_0^1, b_1^1]\right]' \quad \text{with} \\ b_u^f &= \left[-\frac{1}{2r_{u|f}}, \frac{m_{u|f}}{r_{u|f}}, -\frac{m_{u|f}^2}{2r_{u|f}} - 0.5\ln(r_{u|f})\right]', \quad u,f \in u^* = f^* = \{0,1\}. \end{aligned} \tag{42}$$

Its structure vector is $\nu_b = [6,6]$ reflecting that $b^f = [b_0^f, b_1^f]$ is a pair of three-dimensional functions (42).

Use of the identity (14) provides the exponential form of the observation model pdf (33)

$$p(y_t|u_t, f_t, \theta) \propto \exp\{b'(\theta) I(\nu_b, \underline{\delta}(f)) I(\mu_c, \underline{\delta}(u)) c(y_t)]\}.$$

The I-matrices, reflecting the discussed structures are constructed by means of the structure vectors (6), (7), (8) and the vectors of ones (10). Here, they are $\nu_b = [6,6]$, $\quad \nu^1 = \nu^2 = [3,3]$, $\quad \mu_c = [[3,3],[3,3]]$.

According to definitions (13) concerning the structure of b and (11) related to the structure of c, the I-matrices look as follows

$$\begin{aligned} I(\nu_b, \underline{\delta}(f)) &= \operatorname{diag}\{[[1\ 1\ 1\ 1\ 1\ 1]\delta_{0,f},\ [1\ 1\ 1\ 1\ 1\ 1]\delta_{1,f}]\} = \\ &= \operatorname{diag}\{\delta_{0,f}, \delta_{0,f}, \delta_{0,f}, \delta_{0,f}, \delta_{0,f}, \delta_{0,f}, \delta_{1,f}, \delta_{1,f}, \delta_{1,f}, \delta_{1,f}, \delta_{1,f}, \delta_{1,f}\} \quad \text{and} \\ I(\mu_c, \underline{\delta}(u)) &= \operatorname{diag}\{[[[1\ 1\ 1],[1\ 1\ 1]] \odot [\delta_{0,u},\ \delta_{1,u}],\ [[1\ 1\ 1],[1\ 1\ 1]] \odot [\delta_{0,u},\ \delta_{1,u}]]\} = \\ &= \operatorname{diag}\{\delta_{0,u}, \delta_{0,u}, \delta_{0,u}, \delta_{1,u}, \delta_{1,u}, \delta_{1,u}, \delta_{0,u}, \delta_{0,u}, \delta_{0,u}, \delta_{1,u}, \delta_{1,u}, \delta_{1,u}\}. \end{aligned}$$

Remark 7 *In this example, the statistics depend only on the observed output. The input u and the fault f determine just their "position" with respect to functions b of parameters.*

Estimate of θ
The pdf describing knowledge on parameter θ (35) is assumed to be in naturally conjugate form

$$p(\theta|d(t-1), \hat{f}_t) \propto \exp[b'(\theta) C(t-1)].$$

The function $b(\theta)$ is defined by (42) and the statistic

$$C = [C'_{0|0}, C'_{1|0}, C'_{0|1}, C'_{1|1}]'$$

consists of four three-dimensional vectors $C_{u|f} = [V_{u|f}, W_{u|f}, \kappa_{u|f}]'$, $u \in u^*, f \in f^*$, $u^* = f^* = \{0,1\}$. The intial values of statistics $V_{u|f}(0)$, $\kappa_{u|f}(0)$ and $R_{u|f}(0) = V_{u|f}(0) - W^2_{u|f}(0)/\kappa_{u|f}(0)$ are assumed to be positive. It guarantees properness of the prior pdf, [14].

Completion of preparatory stage
Setting time $t = 1$ and selection the first input u_t complete specification of the preparatory stage of the quasi-Bayes estimation algorithm, described in Section 5. Thus, it remains to specialise its on-line stage.

On-line quasi-Bayes estimation of a mill with discrete control

1. Make prediction of y_t, f_t, if need be.

2. Select the prediction $\hat{f}_t(d(t-1))$ of the fault f_t. For instance, set

$$\hat{f}_t(d(t-1)) = \arg\max_{f \in f^*} \hat{\alpha}_{f|u_t}(t-1) = \arg\max_{f \in f^*} \frac{n_{f|u_t}(t-1)}{n_{0|u_t}(t-1) + n_{1|u_t}(t-1)}.$$

3. Observe y_t and evaluate $\Psi_t = [y_t, 1]'$.
4. Make trial updating of the statistics $\tilde{C}(t)$ using the following recursions

$$\tilde{C}(t) = C(t-1) + I(\nu_b, \underline{1})c(\Psi_t).$$

In the given case, it becomes

$$\tilde{V}_{u_t|f}(t) = V_{u_t|f}(t-1) + y_t^2,\ \tilde{W}_{u_t|f}(t) = W_{u_t|f}(t-1) + y_t,\ \tilde{\kappa}_{u_t|f}(t) = \kappa_{u_t|f}(t-1) + 1$$

where u_t is applied input and both possible values of f are considered.

Similarly, make trial update of counters

$$\tilde{n}(t) = n(t-1) + \underline{1}\,\underline{\delta}'(\hat{f}_t).$$

In the given case, it becomes

$$\tilde{n}_{f,\hat{f}_t}(t) = n_{f,\hat{f}_t}(t-1) + 1$$

where $\hat{f}_t$ is the chosen prediction of f_t and both possible values of f are considered.

5. Compute the weights $w_f(t)$, i.e. probabilities of $f_t = f$ for $f \in f^*$ and actual u_t and $\hat{f}_t$, by the formula

$$w_f(t) = \frac{\mathcal{I}_f(t)\tilde{n}_{f|\hat{f}}(t)}{\sum_{\tilde{f}\in f^*} \mathcal{I}_{\tilde{f}}(t)\tilde{n}_{\tilde{f}|\hat{f}}(t)}, \quad \text{where}$$

$$\mathcal{I}_f(t) = \frac{1}{\tilde{\kappa}_{u_t|f}(t)}\Gamma\left(\frac{\tilde{\kappa}_{u|f}(t) - 3}{2}\right)\left(\frac{2}{\tilde{R}_{u_t|f}(t)}\right)^{\frac{\tilde{\kappa}_{u_t|f}(t)-3}{2}},\ \tilde{R}_{u_t|f}(t) = \tilde{V}_{u_t|f}(t) - \frac{\tilde{W}^2_{u_t|f}(t)}{\tilde{\kappa}_{u_t|f}(t)}.$$

6. Select, if need be, the filtered estimate $\hat{f}_t(d(t))$ of the fault f_t. For instance, choose $\hat{f}_t(d(t)) \arg\max_{f\in f^*} w_f(t)$.
7. Compute the statistics

$$C(t) = C(t-1) + I(\nu_b, w(t))c(\Psi_t).$$

In the given case, it becomes

$$\tilde{V}_{u_t|f}(t) = V_{u_t|f}(t-1) + w_f(t)y_t^2,\ \tilde{W}_{u_t|f}(t) = W_{u_t|f}(t-1) + w_f(t)y_t,\ \tilde{\kappa}_{u_t|f_t}(t) = \kappa_{u_t|f}(t-1) + w_f(t)$$

where u_t is applied input and both possible values of f are considered.

Similarly, update counters

$$\tilde{n}(t) = n(t-1) + w(t)\,\underline{\delta}'(\hat{f}_t).$$

In the given case, it becomes

$$\tilde{n}_{f,\hat{f}_t}(t) = n_{f,\hat{f}_t}(t-1) + w_f(t)$$

where $\hat{f}_t$ is the chosen prediction of f_t, or its improved filtered value, and both possible values of f are considered.

8. Update, if need be, the parameter estimates

$$\hat{m}_{f|u}(t) = \frac{W_{u|f}(t)}{\kappa_{u|f}(t)}, \qquad \hat{r}_{u|f}(t) = \frac{R_{u|f}(t)}{\kappa_{u|f}(t)}, \qquad \hat{\alpha}_{f|\hat{f}}(t) = \frac{n_{f|\hat{f}}(t)}{\sum_{i \in f^*} n_{i|\hat{f}}(t)}$$

   for the changed statistics, i.e. having indices $u = u_t$ and both f.

9. Increase the time counter, choose the new system input u_t and return to the step 1.

7.1 Results of experiments

The following figures demonstrate behaviour of the presented algorithm on simulated data.

The data sample presented is chosen rather noisy – see the figure a) where the (univariate) data is plotted against time or the figure b) with a histogram of sampled data. The interesting case is when the noise is so big that the particular clusters get hardly discriminable. On the basis of the figure b) one can be suspicious that the cluster centers could be around values 3, 5, 9, 12 and 14.

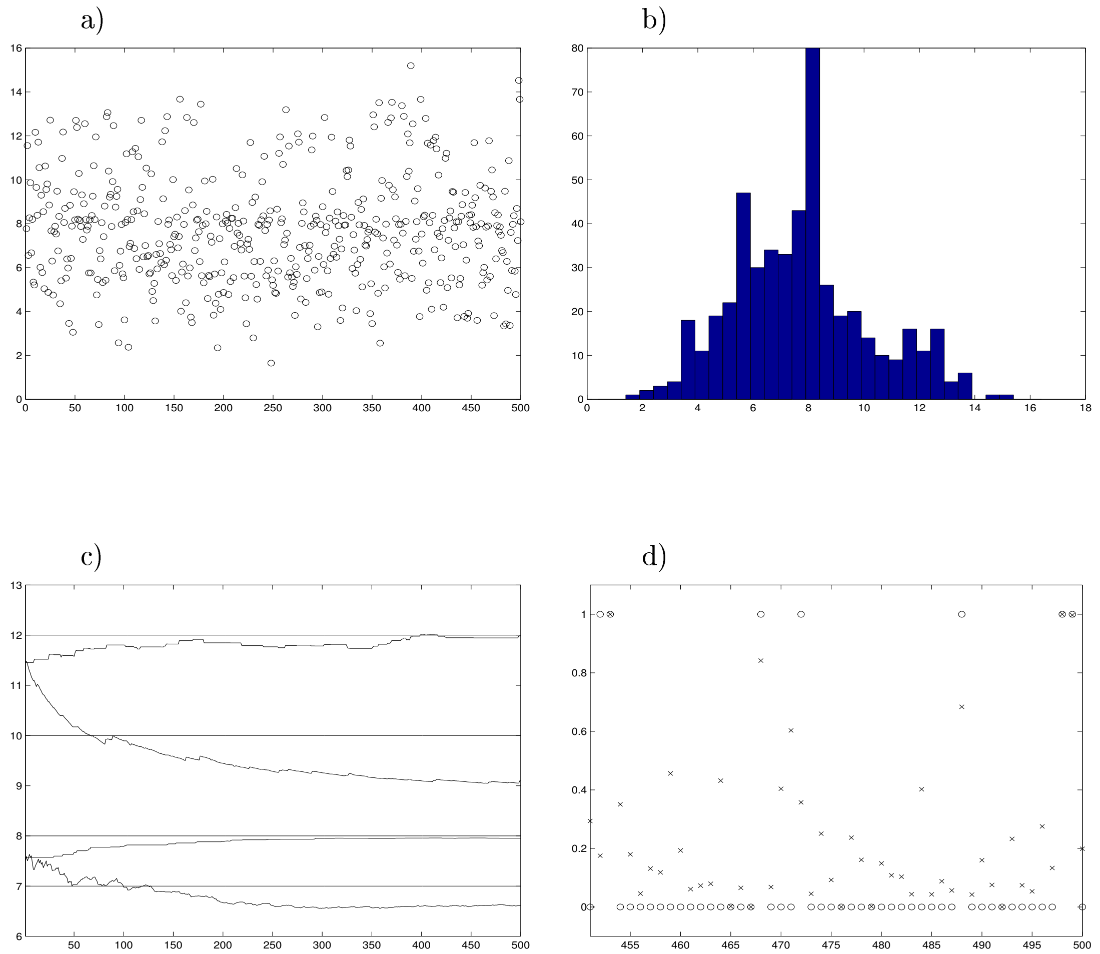
In reality, as described in previous section, two components have been simulated. The center (mean value) of each component depends on the discrete two valued variable u. Thus it is composed of two clusters. So, the simulation generates data with four clusters, whose centers are in positions 7, 8 for the first component and 10, 12 for the second one. The algorithm has not only to find their positions, but also to assort proper couples of clusters to particular components and to decide which component is active in each time.

Figure c) plots courses of parameter estimates during the simulation. It can be seen that some of them are not precise. Intentionally chosen excitation can be blamed for it.

Even under this extreme situation, a rather good performance of the quasi-Bayes-based fault detector has been achieved. It is demonstrated on the figure d). Crosses are probabilities $w_1(t)$ of the activity of the first component that give direct estimate of of the faulty configuration $\hat{f}(t) = 1$ iff $w_1(t) > 0.5$. Circles show the true activity. If the first component is used for data simulation, the circle is put at position 1 otherwise it is drawn at zero level.

Good performance in estimating of the first component means also good performance of the other one due to the dichotomy considered.

Just 7% of false decisions were recorded on the example. By construction, the quasi-Bayesian estimator is adaptive. Its learning properties are reflect to the decrease of this ratio to 6% when initial phase of the run is omitted.



a) simulated data in time, b) histogram of data,
c) courses of parameter estimates during the simulation,
d) correct (o) and estimated (x) probabilities of components

The conclusion visible from the pictures is that the algorithm is able to classify correctly even in the case when the data are so noisy that the components can hardly be recognised visually. Even a poor excitation of one component resulting into imprecise estimation of its parameters does not prevent correct classification. But, it is necessary to stress again that the main purpose of this simple example is to illustrate the theory and its algorithmic features. Assessment of its real value will need extensive simulation, pilot plant and real-life studies.

8 Conclusions

Mixture proved to be powerful modelling tool able to describe systems exhibiting multiple modes of behaviour. Fault detection and isolation (FDI) deal exactly with such a case. Its modes depend, however, in a dynamic way on the system history that often contains discrete-valued (often logical) quantities. Thus, there is an urgent need for efficient algorithms that are able to

- adapt model in on-line mode,
- describe dynamic relationships,
- deal with mixed continuous and discrete valued quantities.

In this paper, a novel quasi-Bayes estimation algorithm has been adapted to FDI problems and extended to the case with mixed quantities involved. Altogether, the above requirements are fully met by the algorithm proposed. Its potential applicability is far behind the addressed FDI area.

At the end, it is fair to list open problems that remained to be addressed

- extensive real-life tests,
- feasible estimation of the model structure,
- initiation of the estimation.

In spite of these open problems, the proposed algorithm is believed worth of extensive tests. From theoretical and algorithmic view-point it has reached the applicability stage.

Acknowledgement

This work has been partially supported by grants of EC, contract No. IST-99-12058, from the Grant Agency of the Czech Republic, contract No GAČR 102/99/1564, and from the Czech Ministry of Education, contract numbers No VS96063 and No ME245.

References

[1] O. Barndorff-Nielsen. *Information and exponential families in statistical theory.* Wiley, New York, 1978.

[2] M. Basseville and I. Nikiforov. *Detection of abrupt changes: theory & applications.* Prentice Hall, Englewood Cliffs, New Jersey, 1993. ISBN 0 13 126780 9.

[3] L. Berec. *Model Structure Identification: Global and Local Views. Bayesian Solution.* Ph.D. Thesis, Czech Technical University, Faculty of Nuclear Sciences and Physical Engineering, Prague, Czech Republic, 1998.

[4] L. Berec. A multi-model method to fault detection and diagnosis: Bayesian solution. An introductory treatise. *International Journal of Adaptive Control and Signal Processing*, 12(1):81–92, 1998.

[5] E. Daum. New exact nonlinear filters. In J. Spall, editor, *Bayesian Analysis of Time Series and Dynamic Models.* Marcel Dekker, New York, 1988.

[6] T. Fergusson. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1:209–230, 1973.

[7] R. Isermann. Process fault detection based on modeling and estimation – a survey. *Automatica*, 20:387–404, 1984.

[8] M. Kárný, J. Kadlec, and E. L. Sutanto. Quasi-Bayes estimation applied to normal mixture. In J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, editors, *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, pages 77–82, Praha, September 1998. ÚTIA AV ČR.

[9] K.Loparo and et al. Leak detection in experimental heat exchanges process. a multiple model approach. *IEEE Trans. AC*, 36(2):167–177, 1991.

[10] R. Kulhavý and M. B. Zarrop. On general concept of forgetting. *International Journal of Control*, 58(4):905–924, 1993.

[11] M.Basseville. Detecting changes in signals and systems - a survey. *Automatica*, 24(3):309–326, 1988.

[12] M.Basseville. Information criteria for residual generation and fault detection and isolation. *Automatica*, 33(5):783–803, 1997.

[13] G. J. McLachlan. *Finite Mixture Models.* Wiley, New York, 1999.

[14] V. Peterka. Bayesian system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.

[15] D. Titterington, A. Smith, and U. Makov. *Statistical Analysis of Finite Mixtures.* John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1985. ISBN 0 471 90763 4.

[16] A. S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, 12:610–611, 1976.