Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Distributionally Robust Fault Detection by using Kernel Density Estimation ⋆

Ting Xue * **Maiying Zhong** ** **Lijia Luo** *** **Linlin Li** ****
Steven X. Ding *

* *Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Germany ({ting.xue, steven.ding}@uni-due.de)*
** *College of Electrical Engineering and Automation, Shandong University of Science and Technology, China (myzhong@buaa.edu.cn)*
*** *Institute of Process Equipment and Control Engineering, Zhejiang University of Technology, China (lijialuo@zjut.edu.cn)*
**** *School of Automation and Electrical Engineering, University of Science and Technology Beijing, China (linlin.li@ustb.edu.cn)*

Abstract: In this paper, a method of distributionally robust fault detection (FD) is proposed for stochastic linear discrete-time systems by using the kernel density estimation (KDE) technique. For this purpose, an H_2 optimization-based fault detection filter is constructed for residual generation. Towards maximizing the fault detection rate (FDR) for a prescribed false alarm rate (FAR), the residual evaluation issue regarding the design of residual evaluation function and threshold is formulated as a distributionally robust optimization problem, wherein the so-called confidence sets are constituted to model the ambiguity of distribution knowledge of residuals in fault-free and faulty cases. A KDE based solution, robust to the estimation errors in probability distribution of residual caused by the finite number of samples, is further developed to address the targeting problem such that the residual evaluation function, threshold as well as the lower bound of FDR can be achieved simultaneously. A case study on a vehicle lateral control system demonstrates the applicability of the proposed FD method.

Keywords: Fault detection, distributionally robust optimization, kernel density estimation.

1. INTRODUCTION

With the increasing demands for safety and reliability of modern control systems, study on fault detection (FD) has received considerable attention both in theory and practical application fields over the past forty years and a rich body of achievements have been reported, see Li et al. (2018); Ding (2013); Yin et al. (2014); Ding (2014); Odiowei and Cao (2010) and the references therein. Generally speaking, an FD system consists of the residual generation and residual evaluation units. Up to now, plenty of residual generation methods have been studied both in model-based and data-driven frameworks, e.g., the fault detection filter (FDF), parity space-based methods and subspace identification methods, Ding (2013); Huang and Kadali (2008) etc. In residual evaluation phase, the residual evaluation function and an appropriate threshold are determined such that the occurrence of a fault can be detected by comparing them. It is notable that residual evaluation is essential to achieve satisfactory FD performance while few research efforts have been contributed to this topic Jung and Frisk (2018).

Concerning the FD for stochastic processes subject to random noises, hypothesis test schemes are usually adopted for residual evaluation by setting the residual evaluation function as a statistic variable of residual and determining a threshold to maximize the fault detection rate (FDR) for a prescribed false alarm rate (FAR), under the assumption of known precise probability distribution for noises Yin et al. (2014); Ding (2014). In industrial applications, however, we are generally inaccessible to the precise distribution knowledge of noises. To handle this obstacle, an alternative solution is to estimate the distributional information of residual from historical data by using the probability distribution estimation methods such as kernel density estimation (KDE) and histogram estimation Silverman (1986). In this way, the hypothesis test methods or Kullback-Leibler divergence based approaches can be applied for residual evaluation, see, e.g., Zhang et al. (2014). Particularly, KDE as a nonparametric approach can estimate the probability density function (PDF) of a random variable from historical data effectively under no distribution assumption on the samples. Remarkably, though a satisfactory estimate of PDF can be achieved when the data set is sufficiently large, the differences between the true probability distribution and the empirical estimate are naturally inevitable due to the

⋆ This work was supported in part by the National Natural Science Foundation of China under Grants (61873149 and 61733009), the Research Fund for the Taishan Scholar Project of Shandong Province of China, the Beijing Natural Science Foundation under grant 4202045, and the Fundamental Research Funds for the Central Universities under Grant FRF-TP-19-032A2. (Corresponding author: Maiying Zhong).

limited number of samples, which might lead to poor FD performance. For this consideration, study on the residual evaluation robust against the ambiguity of distributional information of noises and faults is of practical importance. This motivates our study.

As for dealing with distributionally robust optimization (DRO) problems, the technique of stochastic programming has attracted considerable attention in recent years Gabrel et al. (2014); Jiang and Guan (2016); Shang and You (2018). In the framework of DRO, instead of making specific distribution assumption on random variables, a so-called confidence set is constituted to model the ambiguity of distribution knowledge in terms of the probability properties, e.g., the moments, PDF, support and structural information or the combination of them, etc. In this manner, the optimization problem can be handled without exact distribution knowledge and the delivered solution would be robust to the distributional uncertainties.

Inspired by these observations, in this paper we endeavor to address the distributionally robust FD problem for stochastic linear discrete-time systems. To this end, an H_2 optimization-based FDF is first designed for residual generation. By modeling the confidence sets with empirical PDFs of residuals in fault-free and faulty cases, we then formulate the residual evaluation as a DRO problem in the context of maximizing the FDR for a prescribed FAR. Furthermore, a KDE based solution is developed to solve the underlying FD problem, achieving the residual evaluation function, a threshold and a quantitative lower bound of FDR without posing any distribution assumption on unknown input and fault. Finally, we show the applicability of the proposed method on a vehicle lateral control system.

Notations: In this paper, $Pr\{\cdot\}$ and $\mathbb{P}_\xi$ denote the probability of $\{\cdot\}$ and the cumulative distribution function (CDF) of random variable ξ, respectively. $\{\xi(i)\}_{i=1}^N$ is a N independent and identically distributed (i.i.d) sample set of ξ. $\mathbb{E}[\cdot]$ is the expectation of $[\cdot]$. $\mathcal{N}(\bar{\xi}, \sigma^2)$ denotes the normal distribution with mean $\bar{\xi}$ and variance σ^2. $\mathcal{U}[a, b]$ represents uniform distribution over the interval $[a,\ b]$.

2. PROBLEM FORMULATION

Consider a stochastic linear discrete-time system as follows

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + B_d d(k) + B_f f(k) \\ y(k) = Cx(k) + Du(k) + D_d d(k) + D_f f(k) \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^q, y \in \mathbb{R}^m, f \in \mathbb{R}^l$ and $d \in \mathbb{R}^p$ are the state, input, output, fault and unknown input vectors, respectively, A, B, C, D, B_d, D_d, B_f, D_f are time invariant matrices with appropriate dimensions. In this paper, we assume that the fault $f(k)$ can be a deterministic or random signal and the unknown input $d(k)$ a random vector without knowing exact distribution knowledge.

For FD purpose, a residual generator should be first constructed. In this paper, we use an FDF to this aim

$$\begin{cases} \hat{x}(k+1) = A\hat{x}(k) + Bu(k) + L(y(k) - \hat{y}(k)) \\ \hat{y}(k) = C\hat{x}(k) + Du(k) \\ r(k) = y(k) - \hat{y}(k) \end{cases} \tag{2}$$

where $\hat{x} \in \mathbb{R}^n$ is the estimate of state, $r(k) \in \mathbb{R}^m$ is the residual signal, L is the observer gain matrix to be designed stabilizing $(A - LC)$.

In residual evaluation stage, the residual evaluation function $J(r)$ and an appropriate threshold J_{th} should be determined so that the occurrence of a fault can be detected according to the following logic

$$\begin{cases} J(r) > J_{th} \ \Rightarrow \ \text{fault alarm}, \\ J(r) \leq J_{th} \ \Rightarrow \ \text{fault-free}. \end{cases} \tag{3}$$

To assess the FD performance, we recall the following definitions of FAR and FDR Ding (2013)

$$FAR = Pr\{J(r) > J_{th} | f(k) = 0\}$$
$$FDR = Pr\{J(r) > J_{th} | f(k) \neq 0\}.$$

In case that the probability distribution knowledge of unknown input $d(k)$ and fault $f(k)$ are known perfectly, precise distributional information of residual $r(k)$ can thus be obtained both in fault-free and faulty cases. In this situation, hypothesis test schemes can be adopted by setting $J(r)$ as a statistic variable of residual and determining a threshold J_{th} to maximize FDR for an acceptable FAR or minimizing the FAR for a satisfactory FDR Ding (2014). Unfortunately, we are generally inaccessible to the exact distribution knowledge of $d(k)$ and $f(k)$ in practice.

To address the residual evaluation issue under mild conditions, we without lose of generality define

$$J(r) = \left| w^T r(k) \right|, \quad J_{th} = b \tag{4}$$

where the weighting vector $w \in \mathbb{R}^m$, $w \neq 0$, scalar $b > 0$. Then the design of residual evaluation function $J(r)$ and threshold J_{th} is converted into the design of w, b. Towards maximizing FDR for a prescribed FAR, we further formulate the design of w, b as the following optimization problem

$$\max_{w \neq 0, b > 0} \beta \tag{5}$$

$$\text{s.t.} \quad \begin{cases} \sup\limits_{f(k)=0} Pr\{\left| w^T r(k) \right| > b\} \leq \rho_0 \\ \inf\limits_{f(k)\neq 0} Pr\{\left| w^T r(k) \right| > b\} \geq \beta \end{cases} \tag{6}$$

where $\rho_0 \in (0, 1)$ is the given upper bound of FAR, $\beta \in (0, 1)$ the lower bound of FDR. Constraints in (6) guarantee that the FDR is achieved no less than β and FAR no greater than ρ_0 in the worst-case scenario with respect to performing (4) and (3) for online FD.

The main objectives of this paper are formulated as: 1) to design the observer gain matrix L for residual generation and 2) to address the problem (5)–(6) for w, b without precise distribution knowledge of unknown input and fault.

3. KDE BASED DISTRIBUTIONALLY ROBUST FAULT DETECTION

In this section, we focus on addressing the FD problem (5)–(6) by means of the KDE technique. To this end, the observer gain matrix L is first designed for residual generation in the context of H_2 optimization. Then the problem (5)–(6) is handled for the distributionally robust residual evaluation, followed by a KDE based solution. Finally, the confidentiality of the achieved solution depending on the number of residual samples is briefly discussed.

3.1 H_2 Optimization-based Residual Generation

Based on (1) and (2), we can obtain the dynamics of the residual generator (2) as follows

$$r(k) = \hat{N}_d(z)d(k) + \hat{N}_f(z)f(k)$$

where $\hat{N}_d(z) = C(zI - A + LC)^{-1}(B_d - LD_d) + D_d$ and $\hat{N}_f(z) = C(zI - A + LC)^{-1}(B_f - LD_f) + D_f$.

Under the assumptions that (C, A) is detectable, $D_d D_d^T = I$ and $\forall \theta \in [0, 2\pi]$, $rank\left(\begin{bmatrix} A - e^{j\theta}I & B_d \\ C & D_d \end{bmatrix}\right) = n + m$, we can design the observer gain matrix L by addressing the following problem Ding (2013)

$$\min_L \left\| C(zI - A + LC)^{-1}(B_d - LD_d) \right\|_2 \tag{7}$$

which is an H_2 optimization problem that determines L in the context of minimizing the influence of the norm-bounded unknown input $d(k)$ to residual $r(k)$.

According to Ding (2013), the optimal solution of (7) is obtained as follows

$$L = (AYC^T + B_d D_d^T)(I + CYC^T)^{-1} \tag{8}$$

with $Y \geq 0$ solving the following Riccati equation

$$AYA^T - \bar{L}(I + CYC^T)^{-1}\bar{L}^T - D_d D_d^T = Y$$

where $\bar{L} = AYC^T + B_d D_d^T$.

It is remarkable that various other classes of methods can be used for an optimal design of observer gain matrix L in the context of, e.g., H_∞, H_2/H_2 and H_i/H_∞ optimization, etc. Especially, Ding (2013) presents a unified solution to these problems. Due to the main attention of this paper focuses on the residual evaluation issue with respect to solving (5)–(6), the residual generation will not be discussed here detailedly for brevity. We refer the interested reader to Ding (2013) for more information.

3.2 Distributionally Robust Residual Evaluation

As for solving the problem (5)–(6), the key point lies in coping with the constraints in (6). Despite of the unknown exact distribution knowledge of unknown input and fault, we can usually extract the distributional information of residual from historical data by means of the density estimation methods such as KDE. Meanwhile, it is reasonable to believe that the empirical estimates are nearby the true ones with high probability, especially when the number of samples is large enough. In this context, let

$$r_0(k) = r(k)|_{f(k)=0}, \; r_f(k) = r(k)|_{f(k)\neq 0}$$
$$z(k) = w^T r(k), \; z_0(k) = w^T r_0(k), \; z_f(k) = w^T r_f(k).$$

We constitute the so-called confidence sets Ω_0, Ω_f for fault-free and faulty cases in the following form

$$\Omega_0 = \{\mathbb{P}_z \in \mathcal{M} : D_{KL}(p||\hat{p}_0) \leq \delta_0, p = d\mathbb{P}_z/dz\} \tag{9}$$
$$\Omega_f = \{\mathbb{P}_z \in \mathcal{M} : D_{KL}(p||\hat{p}_f) \leq \delta_f, p = d\mathbb{P}_z/dz\} \tag{10}$$

where $\mathbb{P}_z$ represents the true CDF of $z(k)$, $\mathcal{M}$ denotes the set of all valid CDFs, p is the true PDF of $z(k)$, $\hat{p}_0$, $\hat{p}_f$ are the empirical estimates of p in fault-free and faulty cases, respectively, and

$$D_{KL}(p||\hat{p}_0) = \int p(r) ln\left(\frac{p(r)}{\hat{p}_0(r)}\right) dr$$

is the Kullback–Leibler divergence of p and $\hat{p}_0$, and ditto for $D_{KL}(p||\hat{p}_f)$, δ_0, δ_f are chosen to represent the divergence tolerances measuring the deviations of $\hat{p}_0$, $\hat{p}_f$ from p in fault-free and faulty cases, respectively.

On this basis, the problem (5)–(6) can be reformulated as the following distributionally robust optimization problem

$$\max_{w\neq 0, b>0} \beta \tag{11}$$

$$\text{s.t.} \quad \begin{cases} \sup\limits_{\mathbb{P}_z \in \Omega_0} \mathbb{P}_z\left\{\left|w^T r(k)\right| > b\right\} \leq \rho_0 \\ \inf\limits_{\mathbb{P}_z \in \Omega_f} \mathbb{P}_z\left\{\left|w^T r(k)\right| > b\right\} \geq \beta \end{cases} \tag{12}$$

In (11)–(12), the ambiguity of probability distribution $\mathbb{P}_z$ in fault-free and faulty cases is taken into consideration by modeling the confidence sets Ω_0, Ω_f in terms of the empirical PDFs, which implies the robustness of the solution to problem (11)–(12) against the estimation uncertainties of PDFs $\hat{p}_0$, $\hat{p}_f$ caused by the limited numbers of residual samples in fault-free and faulty cases.

We are now in the position of solving the problem (11)–(12). To this aim, the following theorem is recalled. Please refer to Jiang and Guan (2016) for the proof.

Theorem 1. Given random variable $\xi \in \mathbb{R}^{k_\xi}$ and a confidence set $\Omega = \{\mathbb{P} \in \mathcal{M} : D_{KL}(p||\hat{p}) \leq \delta, p = d\mathbb{P}/d\xi\}$, let $\gamma \in (0, 1)$, $\mathcal{F}(w, \xi)$ be a feasible region described in terms of ξ and the decision variable w. The condition

$$\inf_{\mathbb{P} \in \Omega} \mathbb{P}\{\mathcal{F}(w, \xi)\} \geq \gamma \tag{13}$$

can be equally formulated as follows

$$\hat{\mathbb{P}}\{\mathcal{F}(w, \xi)\} \geq \gamma'$$

where $\gamma' = \max\left\{\inf\limits_{t\in(0,1)}\left\{\frac{e^{-\delta}t^\gamma - 1}{t-1}\right\}, 0\right\} \in [0, 1)$, $\hat{\mathbb{P}}$ is the empirical estimate of $\mathbb{P}$.

Let $\mathcal{F}(w, r) := \left\{\left|w^T r(k)\right| > b\right\}$. According to Theorem 1, it yields from (6) that

$$\sup_{\mathbb{P}_z \in \Omega_0} \mathbb{P}_z\left\{\left|w^T r(k)\right| > b\right\} \leq \rho_0 \Leftrightarrow \inf_{\mathbb{P}_z \in \Omega_0} \mathbb{P}_z\left\{\left|w^T r(k)\right| \leq b\right\} \geq 1 - \rho_0$$
$$\Rightarrow \hat{\mathbb{P}}_{z_0}\left\{\left|w^T r_0(k)\right| \leq b\right\} \geq 1 - \rho_0'$$
$$\inf_{\mathbb{P}_z \in \Omega_f} \mathbb{P}_z\left\{\left|w^T r(k)\right| > b\right\} \geq \beta \Rightarrow \hat{\mathbb{P}}_{z_f}\left\{\left|w^T r_f(k)\right| > b\right\} \geq \beta'$$

where $\hat{\mathbb{P}}_{z_0}$, $\hat{\mathbb{P}}_{z_f}$ denote the estimates of $\mathbb{P}_z$ in fault-free and faulty cases, respectively, and

$$\rho_0' = 1 - \max\left\{\inf_{t\in(0,1)}\left\{\frac{e^{-\delta_0}t^{1-\rho_0} - 1}{t-1}\right\}, 0\right\} \in (0, 1] \tag{14}$$

$$\beta' = \max\left\{\inf_{t\in(0,1)}\left\{\frac{e^{-\delta_f}t^{\beta} - 1}{t-1}\right\}, 0\right\} \in [0, 1). \tag{15}$$

Then the problem (11)–(12) can be rewritten as follows

$$\max_{w\neq 0, b>0} \beta' \tag{16}$$

$$\text{s.t.} \quad \begin{cases} \hat{\mathbb{P}}_{z_0}\left\{\left|w^T r_0(k)\right| \leq b\right\} \geq 1 - \rho_0' \\ \hat{\mathbb{P}}_{z_f}\left\{\left|w^T r_f(k)\right| > b\right\} \geq \beta'. \end{cases} \tag{17}$$

It is remarkable that (16)–(17) is an approximation of the problem (11)–(12), the solution of which highly relies on the quality of empirical estimates $\hat{\mathbb{P}}_{z_0}$, $\hat{\mathbb{P}}_{z_f}$. Moreover, the theoretical upper bound of FAR ρ_0 and the lower bound of

FDR β are achievable based on (14), (15) with probability one on condition that the parameters δ_0, δ_f are chosen appropriately to depict the differences between $\hat{p}_0$, $\hat{p}_f$ and their true values, respectively.

Regarding estimating $\hat{\mathbb{P}}_{z_0}$, $\hat{\mathbb{P}}_{z_f}$ based on historical residual data, KDE as an efficient nonparametric technique has been adopted in a rich body of literature thanks to its capability of estimating PDF of random variables directly from data without any distribution assumption on samples as required in the parametric schemes Zhang et al. (2014). For this merit, in the following subsection we propose a KDE based solution to the problem (16)–(17).

3.3 A KDE based Solution

We start with the basic idea of KDE method. Denote by $\{\xi(i)\}_{i=1}^{N}$ a N i.i.d sample set of ξ being with PDF p. The number of samples, i.e., N, is assumed to be sufficiently large enough. A KDE based estimate of p is obtained as

$$\hat{p}(\xi) = \frac{1}{Nh}\sum_{i=1}^{N}\mathcal{K}\left(\frac{\xi-\xi_i}{h}\right) \tag{18}$$

where $\mathcal{K}(\cdot)$ is a kernel function satisfying $\mathcal{K}(\cdot) > 0$, $\int \mathcal{K}(\xi)d\xi = 1$, $\int \xi\mathcal{K}(\xi)d\xi = 0$, $\int \xi^2\mathcal{K}(\xi)d\xi > 0$, $h > 0$ is the bandwidth of kernel function. Furthermore, the estimate of the CDF of ξ is given by

$$\hat{\mathbb{P}}_{\xi}\left\{\xi \le l\right\} = \frac{1}{Nh}\sum_{i=1}^{N}\mathcal{I}\left(\frac{l-\xi(i)}{h}\right) \tag{19}$$

where $\mathcal{I}(l) = \int_{-\infty}^{l}\mathcal{K}(\xi)d\xi$.

Given $\{r_0(i)\}_{i=1}^{N_0}$, $\{r_f(i)\}_{i=1}^{N_f}$ the i.i.d sample sets of residuals in fault-free and faulty cases, respectively, the empirical estimate of the CDF $\mathbb{P}_{z_0}$ is obtained as follows

$$\begin{aligned} &P_0(w,b) = \hat{\mathbb{P}}_{z_0}\{\left|w^T r_0(k)\right| \le b\} \\ &= \hat{\mathbb{P}}_{z_0}\{-b \le w^T r_0(k) \le b\} \\ &= \frac{1}{N_0 h_0}\sum_{i=1}^{N_0}\left[\mathcal{I}\left(\frac{b-w^T r_0(i)}{h_0}\right) - \mathcal{I}\left(\frac{-b-w^T r_0(i)}{h_0}\right)\right]. \end{aligned} \tag{20}$$

For the faulty case, we have

$$\hat{\mathbb{P}}_{z_f}\{\left|w^T r_f(k)\right| > b\} = 1 - \hat{\mathbb{P}}_{z_f}\{\left|w^T r_f(k)\right| \le b\}$$

with

$$\begin{aligned} &P_f(w,b) = \hat{\mathbb{P}}_{z_f}\{\left|w^T r_f\right| \le b\} \\ &= \frac{1}{N_f h_f}\sum_{i=1}^{N_f}\left[\mathcal{I}\left(\frac{b-w^T r_f(i)}{h_f}\right) - \mathcal{I}\left(\frac{-b-w^T r_f(i)}{h_f}\right)\right] \end{aligned} \tag{21}$$

Based on (20), (21), the problem (16)–(17) can be intuitively converted into the following form

$$\max_{w\neq 0, b>0} \beta' \tag{22}$$

$$\text{s.t.} \quad \begin{cases} P_0(w,b) \ge 1-\rho_0' \\ P_f(w,b) \le 1-\beta'. \end{cases} \tag{23}$$

To achieve a feasible solution to (22)–(23), a bisection line search method Jiang and Guan (2016) is adopted in this paper, as summarized in **Algorithm 1**. More specifically, for a given acceptable ρ_0', we first search for w, b satisfying

Algorithm 1 A KDE based solution to (22)–(23)

Construct residual sample sets $\{r_0(i)\}_{i=1}^{N_0}, \{r_f(i)\}_{i=1}^{N_f}$;
Select kernel function $\mathcal{K}(\cdot)$ and bandwidths h_0, h_f. Initialize $\rho_0' \in (0,1)$, δ_0, $\delta_f > 0$, $\beta_{low}' = 0$, $\beta_{up}' = 1$, a sufficient small parameter $\tau > 0$;
while $\beta_{up}' - \beta_{low}' > \tau$ **do**
  Set $\beta_n' = \frac{\beta_{up}'+\beta_{low}'}{2}$, $\beta' = \beta_n'$;
  if feasible w, b satisfying (23) can be found **then**
    $\beta_{up}' = \beta_n'$;
  else
    $\beta_{low}' = \beta_n'$.
  end if
end while
Set $\beta' = \beta_n'$ and compute β according to (15). Compute ρ_0 based on (14) for given ρ_0';
Return w, b, β, β' and ρ.

Algorithm 2 Online FD

Compute residual $r(k)$ with residual generator (2);
Compute $J(r)$ and threshold J_{th} with (4);
Detect the occurrence of a fault with decision logic (6).

the conditions (23) for a fixed β', then update β' within (0, 1) and repeat this procedure until a maximum value of β' is found. Correspondingly, the solutions of ρ_0, β to (11)–(12) can be achieved based on the relationships between ρ_0', β' and ρ_0, β as specified in (14), (15), respectively.

After achieving w, b with Algorithm 1, we can perform online FD by computing the residual evaluation function $J(r)$ and threshold J_{th} with (4). The algorithm is summarized in **Algorithm 2**. In this manner, the achieved FAR would be no greater than ρ_0 and FDR no less than β.

It should be pointed out that the selection of the kernel function $\mathcal{K}(\cdot)$ and bandwidths h_0, h_f in KDE is important to achieve satisfactory FD performance with respect to solving the problem (22)–(23). Roughly speaking, there are several types of kernel function available such as the Gaussian, Epanechnikov, rectangular and triangular, Barbé et al. (2014); Odiowei and Cao (2010); Silverman (1986) to name just a few. Concerning a trade-off between the bias and variance of estimation error, Epanechnikov kernel is one of the most popular choice in KDE, the kernel function is defined as

$$\mathcal{K}(\xi) = \frac{3}{4}(1-\xi^2), \quad |\xi| \le 1 \tag{24}$$

and the function $\mathcal{I}(\cdot)$ can then be easily derived as

$$\mathcal{I}(l) = \begin{cases} 0 & l < -1 \\ 0.25(-l^3+3l+2), & |l| \le 1 \\ 1 & l > 1. \end{cases}$$

With respect to the bandwidths h_0, h_f, it is notable that a larger bandwidth is prone to be used for simplicity but obscuring bimodal nature of the true distribution and a smaller bandwidth delivers good local density information but possible spurious structure of density by comparison Silverman (1986). To choose an optimal bandwidth, the methods of plug-in, cross-validation and rules-of-thumb can be utilized Gramacki (2018) while at the cost of high computational complexity. Though there is no unified rule for the choice of bandwidth yet, we can usually choose

$$h_0 = 1.06\hat{\sigma}_0 N_0^{-\frac{1}{5}},\ h_f = 1.06\hat{\sigma}_f N_f^{-\frac{1}{5}} \quad (25)$$

for brevity with $\hat{\sigma}_0$, $\hat{\sigma}_f$ being the estimates of variances of $z_0(k)$, $z_f(k)$, respectively. Particularly, in this manner, the bandwidths are optimal in the context of minimizing the mean squared estimation error when the underlying distribution for unknown input and fault being Gaussian Barbé et al. (2014).

3.4 Discussion

As mentioned previously, parameters δ_0, δ_f introduced in the confidence sets (9), (10) are chosen to represent the divergence tolerances of empirical estimates $\hat{p}_0$, $\hat{p}_f$ from the true PDF p in fault-free and faulty cases, respectively, which, according to (14), (15), further determine the theoretical values of the upper bound of FAR ρ_0 and the lower bound of FDR β with respect to solving (22)–(23) for ρ_0', β'. Since the choice of δ_0, δ_f depends on the sample numbers of historical residuals in fault-free and faulty cases, the values of δ_0, δ_f can be further regarded as the robustness indices of the solution to (11)–(12) against the ambiguity of distribution knowledge of residuals in fault-free and faulty cases for the limited number of samples.

So far, few achievements about the determination of δ_0, δ_f are available with regard to estimating $\hat{p}_0$, $\hat{p}_f$ using KDE. As one possible solution suggested by Jiang and Guan (2016), δ_0, δ_f can be chosen as decreasing functions of sample numbers N_0, N_f, respectively, i.e., $\delta_0 = \mathcal{L}(N_0)$, $\delta_f = \mathcal{L}(N_f)$. Then it holds that $\delta_0 \to 0$, $\delta_f \to 0$ and $\rho_0' \to \rho_0$, $\beta' \to \beta$ as $N_0 \to \infty$, $N_f \to \infty$. Viewing in this point, we can without lose of generality choose

$$\delta_0 = \mathcal{L}(N_0^{-\frac{2}{5}}),\ \delta_f = \mathcal{L}(N_f^{-\frac{2}{5}}) \quad (26)$$

with $|\delta_0/N_0^{-\frac{2}{5}}|$, $|\delta_f/N_f^{-\frac{2}{5}}|$ being bounded. It follows that the confidence sets Ω_0, Ω_f are believable with levels no less than ϵ_0, ϵ_f, respectively, where ϵ_0 satisfies

$$Pr\left\{\int (p(r) - \hat{p}_0(r))^2 dr \geq \delta_0\right\} \leq \frac{\mathbb{E}\left[\int (p(r) - \hat{p}_0(r))^2 dr\right]}{\delta_0} = 1 - \epsilon_0$$

and ditto for ϵ_f. We can use ϵ_f, ϵ_f to represent the lower bounds of confidence levels of ρ_0, β, respectively. Unfortunately, it remains difficult to achieve analytical results of ϵ_0, ϵ_f in the probabilistic context. Quantitative analysis of the relationship between δ_0, δ_f in terms of N_0, N_f and ϵ_0, ϵ_f needs to be addressed in the future.

4. SIMULATION RESULTS

In this section, a case study on a vehicle lateral control system Ding (2013) is demonstrated to show the applicability of the proposed method. Denote by α, θ, μ_L, ω the vehicle side slip angle, yaw rate, steering angle and lateral accelerated velocity, respectively. We define the state, system input and output respectively by

$$x = [\alpha \quad \theta]^T, \quad u = \mu_L, \quad y = [\omega \quad \theta]^T.$$

In this case study, we assume the vehicle works steadily with velocity $V_{ref} = 50m/s$. By setting the sampling time with $T_s = 0.1s$, a linear discrete-time model of the system in form of (1) can be established with

$$A = \begin{bmatrix} 0.6333 & -0.0672 \\ 2.0570 & 0.6082 \end{bmatrix},\ B = \begin{bmatrix} -0.0653 \\ 3.4462 \end{bmatrix},\ B_d = [I \quad 0]$$

$$C = \begin{bmatrix} -152.7568 & 1.2493 \\ 0 & 1 \end{bmatrix},\ D = \begin{bmatrix} 56 \\ 0 \end{bmatrix},\ B_f = [B \quad 0]$$

$$D_d = [0 \quad I],\ D_f = [D \quad I]$$

The unknown input is assumed to be $d(k) = [v_1(k)\ v_2(k)]^T$ with $v_1(k) \in \mathbb{R}^2$ and $v_2(k) \in \mathbb{R}^2$ being the process and measurement noise vectors, respectively. For simulation purpose, we generate $v_1(k)$ and $v_2(k)$ with zero-mean white noise sequences. Let $f(k) = [f_1(k)\ f_2(k)\ f_3(k)]^T$, where $f_1(k)$ is the fault in steering angle measurement and $f_2(k)$, $f_3(k)$ the faults in lateral acceleration and yaw rate sensors. We consider the following two faulty cases:

Case 1: $f_1(k) = 0.04 + n_1(k), n_1(k) \sim \mathcal{N}(0, 0.1^2)$
$f_2(k) = f_3(k) = 0;$

Case 2: $f_3(k) = -0.2 + n_2(k), n_2(k) \sim \mathcal{U}[-0.01, 0.01]$
$f_1(k) = f_2(k) = 0.$

For residual generation purpose, an optimal observer gain matrix L is obtained based on (8) as

$$L = \begin{bmatrix} -0.0042 & -0.0340 \\ -0.0134 & 0.3427 \end{bmatrix}.$$

Then we can construct the residual generator (2) with $r(k) = [r_1(k) \quad r_2(k)]^T$ and establish the sample sets $\{r_0(i)\}_{i=1}^{N_0}$ with $N_0 = 5000$ for fault-free case and $\{r_f(i)\}_{i=1}^{N_f}$ with $N_f = 2000$ for faulty case 1 and $N_f = 5000$ for faulty case 2, respectively. The profiles of the residuals in normal operation and each faulty case are illustrated in Fig. 1.

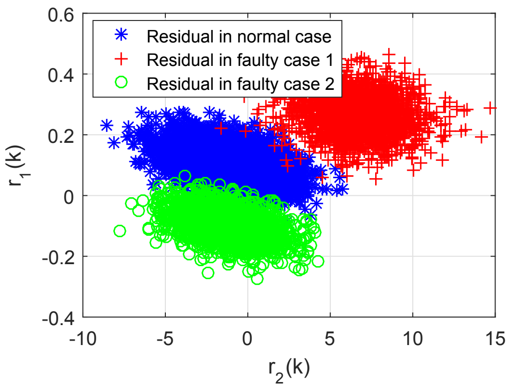

Fig. 1. Profiles of residuals in fault-free and faulty cases.

Now we confine ourselves to solve the problem (22)–(23) for w, b without using distribution knowledge of $d(k)$, $f(k)$. By choosing Epanechnikov kernel (24) as the kernel function, the bandwidths h_0, h_f are set as (25) with $\hat{\sigma}_0$, $\hat{\sigma}_f$ being determined by performing post-optimization analysis. Let $\delta_0 = J_0/log(N_0^{2/5}) = 2.9352 \times 10^{-4}$, $\delta_f = J_f/log(N_f^{2/5})$ with $J_0 = J_f = 0.001$. The confidence sets Ω_0 and Ω_f for each faulty case can thus be determined. Given $\rho_0' = 0.05$, an acceptable theoretical FAR $\rho_0 = 0.0554$ can then be obtained based on (14). A KDE based solution of w, b, β' are achieved by solving the problem (22)–(23) with Algorithm 1. The theoretical value of β are then computed according to (15). The results of δ_f, β, β' for the two faulty cases are summarized in Tab. 1. The

Table 1. Feasible solutions of β, β' with $\rho_0 = 0.0554$, $\rho'_0 = 0.05$

Faults	N_f	$\delta_f(\times 10^{-4})$	β	β'
Faulty case 1	2000	3.2891	0.9642	0.9688
Faulty case 2	5000	2.9352	0.9117	0.9184

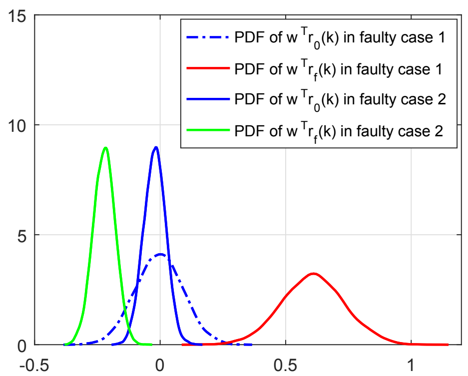

Fig. 2. The profiles of $\hat{p}_0$ and $\hat{p}_f$ for faulty cases 1 and 2.

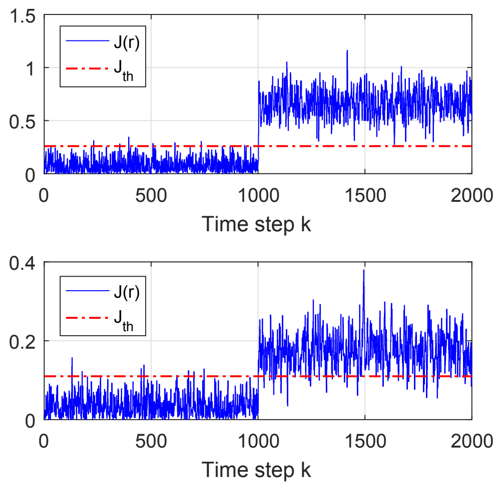

Fig. 3. FD results for the faulty case 1 and 2.

estimates of PDFs of $w^T r_0(k)$ and $w^T r_f(k)$ are shown in Fig. 2. It is seen from the Tab. 1 and Fig. 2 that the obtained solution of w can deliver good fault detectability for the two faulty cases.

To show the effectiveness of the proposed FD method, we inject each fault at $k = 1000$. By setting the residual evaluation function and threshold with (4), we perform online FD with Algorithm 2. The FD results are given in Fig. 3. It is seen from Tab. 1, Fig. 2 and Fig. 3 that satisfactory FD results by virtual of acceptable FAR and higher FDR can be achieved without knowing precise distribution knowledge of unknown input and fault.

5. CONCLUSIONS

In this paper, a distributionally robust FD approach has been proposed for stochastic linear discrete-time processes by using KDE without precise distribution knowledge of unknown input and fault. An H_2 optimization-based FDF was first constructed for residual generation. By introducing the confidence sets to model the ambiguity of distribution knowledge of residuals in fault-free and faulty cases, a distributionally robust optimization problem was formulated for residual evaluation in the context of maximizing the FDR for a given acceptable FAR. A KDE based solution was developed to address the underlying optimization problem. Its robustness to the estimation errors of PDF of residual caused by the limited number of samples was discussed. Simulation results on a vehicle lateral control system demonstrated the applicability of the proposed approach.

REFERENCES

Barbé, K., Fuentes, L., Barford, L., and Lauwers, L. (2014). A guaranteed blind and automatic probability density estimation of raw measurements. *IEEE Trans. Instrum. Meas.*, 63(9), 2120–2128.

Ding, S. (2013). *Model-based fault diagnosis techniques: design schemes, algorithms, and tools.* Springer-Verlag, London, 2nd edition.

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

Gabrel, V., Murat, C., and Thiele, A. (2014). Recent advances in robust optimization: An overview. *European Journal of operational research*, 235(3), 471–483.

Gramacki, A. (2018). *Nonparametric kernel density estimation and its computational aspects.* Springer, Switzerland.

Huang, B. and Kadali, R. (2008). *Dynamic modeling, predictive control and performance monitoring: a data-driven subspace approach.* Springer-Verlag, London.

Jiang, R. and Guan, Y. (2016). Data-driven chance constrained stochastic program. *Mathematical Programming*, 158(1–2), 291–327.

Jung, D. and Frisk, E. (2018). Residual selection for fault detection and isolation using convex optimization. *Automatica*, 97, 143–149.

Li, M., Chen, W., and Zhang, T. (2018). A novel seizure diagnostic model based on kernel density estimation and least squares support vector machine. *Biomed Signal Proces. Contr.*, 41, 233–241.

Odiowei, P.E.P. and Cao, Y. (2010). Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. *IEEE Trans. Ind. Inf.*, 6(1), 36–45.

Shang, C. and You, F. (2018). Distributionally robust optimization for planning and scheduling under uncertainty. *Comput. & Chem. Eng.*, 110, 53–68.

Silverman, B. (1986). *Density estimation for statistics and data analysis.* Springer, London.

Yin, S., Ding, S., Xie, X., and Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. *IEEE Trans. Ind. Electron.*, 61(11), 6418–6428.

Zhang, F., Liu, Y., Chen, C., Li, Y., and Huang, H. (2014). Fault diagnosis of rotating machinery based on kernel density estimation and kullback-leibler divergence. *J. Mech. Sci. Technol.*, 28(11), 4441–4454.