

# An FDR-Oriented Approach to Multiple Sequential Fault Detection and Isolation

Jie Chen

Department of EEIS  
University of Science and  
Technology of China, Hefei, China  
Email: cj12@mail.ustc.edu.cn

Wenyi Zhang

Department of EEIS  
University of Science and  
Technology of China, Hefei, China  
Email: wenyizha@ustc.edu.cn

H. Vincent Poor

Department of Electrical Engineering  
Princeton University  
Princeton, NJ, USA  
Email: poor@princeton.edu

**Abstract**—The problem of sequential fault detection and isolation in multiple data streams is considered. In this work, it is assumed that many independent parallel data streams, each of which has a (possibly infinite) change point, are sequentially observed with a maximum sampling constraint. The pre-change data follow a known distribution, and the post-change data follow one of  $J$  possible distributions. A sequential procedure is proposed to detect the changes for all data streams, and to isolate the types of changes upon their detection. The sequential procedure is shown to control the false discovery rate. An asymptotic upper bound on the average detection delay over the parallel data streams is also derived. A simulation study is presented to illustrate the proposed procedure and to corroborate the analysis.

## I. INTRODUCTION

The problem of sequential change detection is concerned with the situation in which there is an abrupt change at some time in a system that is monitored in real time. After the change, there are  $J \geq 2$  distinct post-change distributions, only one of which is correct, and the goal is to detect the change and identify the correct one as soon as possible after the change occurs, subject to a false alarm and a false isolation constraint. Sequential change detection and isolation has many applications, including fault diagnosis in complex dynamical systems and industrial processes, environment surveillance and monitoring, and radar and sonar signal processing [1], [2], [3]. The general sequential change detection and diagnosis problem was first considered in [4], and subsequently studied in various works [5], [6], [7].

The rapid development of sensing technology allows for the generation of large-scale real-time streaming data [8]. Such data may involve multiple data streams subject to change points, which is the situation considered in this paper. More precisely, suppose that at each time  $t = 1, 2, \dots$ , we have observations of  $K$  parallel data streams,  $X_t^{(k)}$ ,  $k = 1, 2, \dots, K$ . There is a change point  $t^{(k)} \geq 1$  for each of the streams, such that, before the change point, the stream is in a normal condition but after  $t^{(k)}$  it is in an abnormal condition, and it is possible to have  $t^{(k)} = \infty$  (i.e., there is no change in the stream). These  $K$  change points are unknown and possibly different, and the abnormal states are not known exactly. This setting motivates the study of change detection and isolation

in multiple data streams.

Compared with its single-stream counterpart, multi-stream sequential change detection and isolation has been less well studied. To characterize the control of errors in multiple hypothesis testing, there are various error metrics, such as the false discovery rate (FDR) and the familywise error rate (FWER) [8], [9], to evaluate the false-alarm performance. The FDR is the expected value of the proportion of falsely rejected hypotheses among all rejected hypotheses, while the FWER is the probability of rejecting any true null hypotheses. So FDR and FWER represent different methods of quantifying the frequency of false alarms (i.e., Type I errors) among multiple data streams. Procedures based on the control of FDR are more powerful in identifying alternative hypotheses than are FWER-based procedures [8], at the cost of increased error rates.

In this paper, we adapt the definition of FDR to handle both false alarm events and false isolation events. We propose a sequential procedure that achieves FDR control for the Bayesian multiple change detection and isolation problem that is similar to the Bayesian multiple change detection approach of [10]. We impose a constraint on the maximum number of observed samples. With the advantage of FDR control, we can detect and isolate more changes subject to an error constraint than an FWER-based control procedure does. For the  $k$ -th data stream, the procedure updates a statistic  $S_n^{(k)}$  after the  $n$ -th sampling. Then these statistics across the data streams are processed by a variant of Benjamini and Hochberg's procedure [9] to yield detection and isolation results. To characterize the statistical properties of our proposed procedure, we establish a theoretical guarantee for the FDR and derive an asymptotic upper bound of the average detection delay (ADD). We further compare procedures that control the FDR and the FWER via simulation, and illustrate the advantage of the FDR control procedure in terms of smaller ADD.

The rest of this paper is organized as follows. Section II describes the Bayesian multiple change detection and isolation problem and the corresponding definition of FDR. Section III describes the detection and isolation procedure. Section IV develops the performance guarantee of the proposed pro-

cedure. Section V presents the simulation and numerical results. Finally, Section VI concludes this paper and discusses possible extension of this work.

## II. MODEL AND FORMULATION

Consider a situation in which there are  $K \geq 2$  data streams:

$$X_1^{(k)}, X_2^{(k)}, \dots, X_{t^{(k)}-1}^{(k)}, X_{t^{(k)}}^{(k)}, \dots \text{ for } k = 1, \dots, K. \quad (1)$$

Denote  $[K] = \{1, \dots, K\}$ . For each data stream  $k \in [K]$ , there is a change point  $t^{(k)} \geq 1$ , which is assumed to be a random variable with a known prior distribution  $\pi_m = P\{t^{(k)} = m\}$ . If  $t^{(k)} = \infty$ , there does not exist a change point in the  $k$ th data stream. For a measurable space  $(\Omega, \mathcal{F})$ , consisting of a sample space  $\Omega$  and a  $\sigma$ -field  $\mathcal{F}$  of events, we consider a family  $\{P_{\pi, j}^{(k)} | k \in [K], j \in [J]\}$  of probability measures on  $(\Omega, \mathcal{F})$  that describes both the prior distribution of  $t^{(k)}$  and the distribution of  $\{X_n^{(k)} | n = 1, 2, \dots\}$ , such that, under  $P_{\pi, j}^{(k)}$ ,  $X_1^{(k)}, X_2^{(k)}, \dots, X_{t^{(k)}-1}^{(k)}$  are independent and identically distributed (i.i.d.) with a distribution  $H_0$ , and  $X_{t^{(k)}}^{(k)}, X_{t^{(k)}+1}^{(k)}, \dots$  are i.i.d. with distribution  $H_{j^{(k)}}$ ,  $j^{(k)} \in [J]$  and are independent of  $X_1^{(k)}, X_2^{(k)}, \dots, X_{t^{(k)}-1}^{(k)}$ . For simplicity, we use  $P_{\pi, j}^{(k)}$  to denote  $P_{\pi, j^{(k)}}^{(k)}$ . We assume a prior distribution  $p_j$  on the change-type  $j$ . We further assume mutual independence among the  $K$  data streams, and that  $H_0$  and  $H_j$  have densities  $h_0$  and  $h_j$ ,  $j = 1, \dots, J$ , with respect to some measure  $P$  on  $(\Omega, \mathcal{F})$ .

Now, we consider imposing a constraint on the maximum sample size of each observed data stream; that is, we may observe at most  $\bar{N}$  samples in each data stream, where  $\bar{N}$  may be a very large but finite integer. Then our goal is to define a (multi-)detection-isolation rule  $\{T^{(k)}, \hat{j}^{(k)} | T^{(k)} \leq \bar{N}, \hat{j}^{(k)} \in [J], k \in [K]\}$ , which means that a change is detected at time  $T^{(k)}$  and determined to be of type  $\hat{j}^{(k)}$  for the  $k$ th data stream, such that the FDR is no greater than some given value  $\alpha \in (0, 1)$ . Note that  $T^{(k)} = \bar{N}$  means that no change has been detected. We define the FDR as

$$\text{FDR} = E_{\pi, p} \left\{ \frac{\bar{V}}{\bar{R} \vee 1} \right\}, \quad (2)$$

where  $E_{\pi, p}\{\cdot\}$  denotes expectation under  $\{P_{\pi, j}^{(k)} | k \in [K], j \in [J]\}$  and change-type distribution  $p$ , “ $\bar{R} \vee 1$ ” denotes the maximum between  $\bar{R}$  and 1,  $\bar{R}$  is the number of change points declared, i.e., the size of the subset of  $[K]$  such that  $T^{(k)} < \bar{N}$ , and  $\bar{V}$  is the number of false alarm or false isolation events, i.e. the size of  $\{T^{(k)} < t^{(k)}, T^{(k)} < \bar{N}\} \cup \{t^{(k)} \leq T^{(k)} < \bar{N}, \hat{j}^{(k)} \neq j^{(k)}\}$ . For example, suppose we have ten data streams, six of which possess finite change points, and we have seven stopping times with  $T^{(k)} < \bar{N}$ , among which six of them have  $T^{(k)} \geq t^{(k)}$ , but one of these six results is not isolated correctly, i.e.  $\hat{j}^{(k)} \neq j^{(k)}$ . As a result,  $\bar{V}/\{\bar{R} \vee 1\}$  should be  $2/7$ .

## III. DETECTION-ISOLATION PROCEDURE

For each data stream  $k \in [K]$ , we define a statistic according to [5, Thm. 6]

$$S_n^{(k)} = \max_{1 \leq j \leq J} \left\{ \max_{n-m_\alpha \leq l \leq n} \sum_{s=l}^n Z_s^{(k)}(j, 0) - \left( \max_{1 \leq g \leq J, g \neq j} \max_{n-m_\alpha \leq l \leq n} \sum_{s=l}^n Z_s^{(k)}(g, 0) \right)^+ \right\}, \quad (3)$$

in which  $m_\alpha$  is a window size and

$$Z_s^{(k)}(j, 0) = \log \left( h_j(X_s^{(k)}) / h_0(X_s^{(k)}) \right).$$

Let

$$I_{jg} = D(h_j \| h_g) = \int \log \left( \frac{h_j(x)}{h_g(x)} \right) h_j(x) P(dx).$$

As in [5], we assume that  $m_\alpha$  satisfies

$$\liminf m_\alpha / \left| \log \frac{\alpha}{K} \right| > 1 / \min_{1 \leq j \leq J} \min_{0 \leq g \leq J, g \neq j} I_{jg}, \quad (4)$$

$$\log m_\alpha = o \left( \left| \log \frac{\alpha}{K} \right| \right) \text{ as } K \rightarrow \infty.$$

Given a prescribed parameter  $\alpha \in (0, 1)$ , we construct, for each  $k \in [K]$ , threshold values

$$U_K^{(k)} \leq U_{K-1}^{(k)} \leq \dots \leq U_1^{(k)} \quad (5)$$

such that when  $S_{T^{(k)}}^{(k)} \geq U_s^{(k)}$ , in which  $T^{(k)}$  is the first time of  $S_n^{(k)} \geq U_s^{(k)}$ , we have

$$\sum_{j^{(k)}=1}^J \sum_{t^{(k)}=1}^{\infty} p_{j^{(k)}} \pi_{t^{(k)}} P_{t, j^{(k)}}^{(k)} (t^{(k)} \leq T^{(k)} < \infty, \hat{j}^{(k)} \neq j^{(k)}) + \sum_{t^{(k)}=1}^{\infty} \pi_{t^{(k)}} P_t^{(k)} (T^{(k)} < t^{(k)}) \leq \left( \frac{S}{K} \right) \alpha, \forall s \in [K]. \quad (6)$$

To construct the thresholds  $\{U_s^{(k)}\}$  to satisfy (6), according to [5, Thm. 7], we have

$$\sup_{t \geq 1} P_t^{(k)} (t \leq T^{(k)} < t + m_\alpha) \leq \frac{2Jm_\alpha}{e^{U_s^{(k)}}}, \quad (7)$$

$$\sup_{1 \leq t \leq r} P_{t, j^{(k)}}^{(k)} (r \leq T^{(k)} < r + m_\alpha, \hat{j}^{(k)} \neq j^{(k)}) \leq \frac{2(J-1)m_\alpha}{e^{U_s^{(k)}}}. \quad (8)$$

Here  $(T^{(k)}, \hat{j}^{(k)})$  is attained by the procedure DI-FDR described in Algorithm 1. Then we have

$$\sum_{j^{(k)}=1}^J \sum_{t^{(k)}=1}^{\infty} p_{j^{(k)}} \pi_{t^{(k)}} P_{t, j^{(k)}}^{(k)} (t^{(k)} \leq T^{(k)} < \infty, \hat{j}^{(k)} \neq j^{(k)}) + \sum_{t^{(k)}=1}^{\infty} \pi_{t^{(k)}} P_t^{(k)} (T^{(k)} < t^{(k)}) \leq 2(J-1)m_\alpha^2 e^{-U_s^{(k)}} + 2Jm_\alpha e^{-U_s^{(k)}} \left\{ 1 + m_\alpha^{-1} \sum_{t=1}^{\infty} t \pi_t \right\} \leq [2(J-1)m_\alpha + 2J] m_\alpha e^{-U_s^{(k)}}, \quad (9)$$

where the first inequality follows from an argument similar to the proof of [5, Thm. 7], and the second inequality is from the condition we set on the distribution of change points in

Theorem 1 in Section IV-A and the condition we set on the window size (4). Then we set

$$[2(J-1)m_\alpha + 2J]m_\alpha e^{-U_s^{(k)}} = \frac{S}{K}\alpha. \quad (10)$$

Thus we can get the thresholds  $U_s^{(k)}$  as we fix the window size  $m_\alpha$  and  $U_s^{(k)} \sim |\log(\alpha/K)|$  as  $K \rightarrow \infty$ .

Let  $I_q$  denote the indices of active data streams (i.e., data streams that have not been stopped) at the beginning of the  $q$ th stage of sampling, and let  $n_q$  denote the accumulated sample size of active data streams at the end of the  $q$ th stage of sampling. Clearly  $I_1 = [K]$  and  $n_0 = 0$ . Algorithm 1 describes the  $q$ th stage of sampling ( $q = 1, 2, \dots$ ).

---

**Algorithm 1** DI-FDR

---

- 1) Sample the active data streams  $\{X_n^{(k)}\}_{k \in I_q, n > n_{q-1}}$ , and update the statistics  $\{S_n^{(k)}\}_{k \in I_q, n > n_{q-1}}$  using (3).
- 2) Sort the statistics in ascending order as  $\{S_n^{(i(n,l))}\}$ , where  $i(n,l)$  denotes the index of the  $l$ th ordered active detection statistic at sample size  $n$ .
- 3) Continue until  $n$  equals

$$n_q = \bar{N} \wedge \inf \left\{ n > n_{q-1} \mid S_n^{(i(n,l))} \geq U_{K-l+1}^{(i(n,l))}, \exists l \in [I_q] \right\}. \quad (11)$$

- a) If  $n_q < \bar{N}$ , then stop sampling and declare change points at  $n_q$  for the following data streams:

$$X_{n_q}^{(i(n_q, l_q))}, X_{n_q}^{(i(n_q, l_q+1))}, \dots, X_{n_q}^{(i(n_q, |I_q|))},$$

where

$$l_q = \min \left\{ l \in [I_q] \mid S_n^{(i(n,l))} \geq U_{K-l+1}^{(i(n,l))} \right\}. \quad (12)$$

Declare the corresponding change types as

$$\hat{j}^{(k)} = \arg \max_{1 \leq j \leq J} \left\{ \max_{n_q - m_\alpha \leq l \leq n_q} \sum_{s=l}^{n_q} Z_s^{(k)}(j, 0) \right\}, \quad (13)$$

respectively. Update  $I_{q+1}$  to include the indices of the remaining active data streams and proceed to the  $(q+1)$ th stage; if  $|I_{q+1}| = 0$ , then stop.

- b) Otherwise,  $n_q = \bar{N}$  so declare all the remaining active data streams to have no change points, and stop.
- 

#### IV. PROPERTIES OF DI-FDR

In this section we analyze statistical properties of the proposed parallel detection-isolation procedure DI-FDR in Algorithm 1. We consider two performance metrics: (i) the FDR; and (ii) the ADD.

##### A. False Discovery Rate

The following result, Theorem 1, states that DI-FDR controls FDR.

**Theorem 1.** Let  $I^* = \min_{1 \leq j \leq J} \min_{0 \leq g \leq J, g \neq j} I_{jg} > 0$  and let the change-point distribution be such that

$$\log \left( \sum_{t=1}^{\infty} t \pi_t \right) = o \left( \left| \log \frac{\alpha}{K} \right| \right), \quad (14)$$

as  $K \rightarrow \infty$ . Then for a given  $0 < \alpha < 1$ , as  $K \rightarrow \infty$ , the procedure DI-FDR satisfies

$$FDR \leq \alpha + \frac{KE \left\{ t^{(k)} \mid t^{(k)} < \infty \right\} + (K \log K - K \log \alpha) / I^*}{\bar{N}}. \quad (15)$$

So if we choose  $\bar{N}$  to satisfy  $(K \log K) / \bar{N} \rightarrow 0$  as  $K \rightarrow \infty$ , we have

$$FDR \leq \alpha.$$

*Proof.* The proof is similar to that of [10, Thm. 1]. Here we provide only the difference between the proofs. We use the following result about the expected delay in the fault detection-isolation problem from [5].

**Lemma 2.** [5, Thm. 7] If  $(T^{(k)}, \hat{j}^{(k)})$  satisfies (6), then for every  $1 \leq j \leq J$

$$\begin{aligned} & \sum_{t^{(k)}=1}^{\infty} \pi_{t^{(k)}} E_{t,j}^{(k)} \left\{ T^{(k)} - t^{(k)} \mid T^{(k)} \geq t^{(k)} \right\} \\ & \sim (1 + o(1)) \log \alpha \Big| \min_{0 \leq g \leq J, g \neq j} I_{jg} \text{ as } \alpha \rightarrow 0. \end{aligned} \quad (16)$$

Hence, we have

$$\begin{aligned} & E_{\pi,p}^{(k)} \left\{ T^{(k)} - t^{(k)} \mid T^{(k)} \geq t^{(k)} \right\} \\ & = \sum_{j^{(k)}=1}^J \sum_{t^{(k)}=1}^{\infty} p_{j^{(k)}} \pi_{t^{(k)}} E_{t,j}^{(k)} \left\{ T^{(k)} - t^{(k)} \mid T^{(k)} \geq t^{(k)} \right\} \\ & \leq \frac{1}{I^*} \log \frac{K}{\alpha} (1 + o(1)), \end{aligned} \quad (17)$$

where  $I^* = \min_{1 \leq j \leq J} \min_{0 \leq g \leq J, g \neq j} I_{jg}$ , and  $o(1)$  tends to 0 as  $K \rightarrow \infty$  in our situation. Then we can take this result into the proof of [10, Thm. 1] to establish Theorem 1.  $\square$

##### B. Average Detection Delay

We are interested in the expected number of additional observations after the change point required for detection and isolation. We consider the following ADD function as the average of ADDs for those data streams detected:

$$\begin{aligned} & \text{ADD} \\ & = \frac{1}{K} \sum_{k=1}^K \sum_{t^{(k)}} \pi_{t^{(k)}} \sum_{j^{(k)}=1}^J p_{j^{(k)}} E_{t,j}^{(k)} \left\{ T^{(k)} - t^{(k)} \mid T^{(k)} \geq t^{(k)} \right\} \\ & = \frac{1}{K} \sum_{k=1}^K E_{\pi,p}^{(k)} \left\{ T^{(k)} - t^{(k)} \mid T^{(k)} \geq t^{(k)} \right\}. \end{aligned} \quad (18)$$

According to (17), we have the following result providing an asymptotic upper bound on the ADD.

**Theorem 3.**

$$\text{ADD}(FDR \leq \alpha) \leq 1/I^* \left( \left| \log \alpha (1 - \pi_\infty) \right| + 1 \right) (1 + o(1)), \text{ as } K \rightarrow \infty. \quad (19)$$

The proof is similar to that of [10, Thm. 3].

## V. NUMERICAL EXPERIMENTS

In this section we perform a Monte Carlo experiment to illustrate our procedure proposed in Section III, and compare the performance of the proposed procedure with another procedure, which will be introduced shortly below.

Before presenting the setup of the experiment, we first introduce FWER control procedures. The FWER is defined as the probability of rejecting any true hypothesis in multiple hypothesis testing problems. In our problem, this means the probability of false alarm or false isolation on any data streams. To achieve an FWER no greater than  $\alpha$  within  $K$  data streams, Bonferroni's procedure [8] requires the error probability in (6) on each data stream to be no greater than  $\alpha/K$ .

We describe the  $q$ th stage of this FWER-control procedure in Algorithm 2.

---

### Algorithm 2 DI-FWER

---

- 1) Sample the active data streams  $\{X_n^{(k)}\}_{k \in I_q, n > n_{q-1}}$ , and update the statistics  $\{S_n^{(k)}\}_{k \in I_q, n > n_{q-1}}$  using (3).
- 2) Continue until  $n$  equals

$$n_q = \bar{N} \wedge \inf \{n > n_{q-1} \mid S_n^{(k)} \geq U_1^{(k)}, \exists k \in I_q\}. \quad (20)$$

- a) If  $n_q < \bar{N}$ , then stop sampling and declare change points at  $n_q$  for those data streams that satisfy (20). Declare the corresponding change types as

$$\hat{j}^{(k)} = \arg \max_{1 \leq j \leq J} \left\{ \max_{n_q - m_n \leq l \leq n_q} \sum_{s=l}^{n_q} Z_s^{(k)}(j, 0) \right\}, \quad (21)$$

respectively. Update  $I_{q+1}$  to include the indices of the remaining active data streams and proceed to the  $(q+1)$ th stage; if  $|I_{q+1}| = 0$ , then stop.

- b) Otherwise,  $n_q = \bar{N}$  so declare all the remaining active data streams to have no change points, and stop.
- 

To make our comparison fair, we set the same upper bound for FDR and FWER as  $\alpha = 0.1$ . As for the prior distributions of the change points, we set the probability of infinite change points as  $\pi_\infty$ , and assume that the finite change points each follows a (conditional) geometric distribution with  $p = 0.05$  (that is, the mean change point on each stream is 20). We choose the distribution of the post-change type as a uniform distribution on  $\{1, \dots, J\}$ . The maximum sample size is set as 2000. We consider i.i.d  $N(\mu, 1)$  data streams with pre/post-change distributions:

$$H_0 : \mu = 0 \quad \text{vs.} \quad H_j : \mu = \mu_j.$$

We consider four configurations for comparison as follows:

- $P1$  :  $J = 2, \pi_\infty = 0.1, [\mu_1, \mu_2] = [2, 6]$ ,
- $P2$  :  $J = 2, \pi_\infty = 0.1, [\mu_1, \mu_2] = [2, -2]$ ,
- $P3$  :  $J = 2, \pi_\infty = 0.6, [\mu_1, \mu_2] = [2, 6]$ ,
- $P4$  :  $J = 3, \pi_\infty = 0.1, [\mu_1, \mu_2, \mu_3] = [2, 6, 4]$ .

TABLE I  
DI-FDR AND DI-FWER FOR MULTIPLE CHANGE DETECTION/ISOLATION OF I.I.D. GAUSSIAN DATA STREAMS.

$P1$				
$K$	Procedure	Estimated FDR/FWER	ADD	$\bar{R}$
10	DI-FDR	0.0143	3.49	9.13
	DI-FWER	0.0154	3.98	9.02
100	DI-FDR	0.0137	3.50	90.87
	DI-FWER	0.0137	4.68	89.65
500	DI-FDR	0.0136	3.54	455.53
	DI-FWER	0.0151	5.29	449.50
1000	DI-FDR	0.0138	3.54	912.00
	DI-FWER	0.0146	5.53	900.05

TABLE II  
DI-FDR AND DI-FWER FOR MULTIPLE CHANGE DETECTION/ISOLATION OF I.I.D. GAUSSIAN DATA STREAMS.

$P2$				
$K$	Procedure	Estimated FDR/FWER	ADD	$\bar{R}$
10	DI-FDR	0.0306	4.96	9.28
	DI-FWER	0.0379	5.69	9.04
100	DI-FDR	0.0319	5.06	92.38
	DI-FWER	0.0392	6.89	89.68
500	DI-FDR	0.0297	5.07	462.91
	DI-FWER	0.0343	7.77	449.52
1000	DI-FDR	0.0300	5.07	926.56
	DI-FWER	0.0320	8.17	900.07

Tables I-IV show the detection and isolation results with different configurations.

DI-FDR and DI-FWER satisfy the upper bound  $\alpha$ , notably much less than the prescribed value. We can see that a decrease in  $\{I_{jg} \mid 1 \leq j \leq J, 0 \leq g \leq J, g = j\}$  will increase the FDR/FWER and ADD, which is consistent with intuition. From Table III, we see that an increase in the probability of infinite change points can increase the FDR/FWER significantly. Because of the assumption on the distribution of change points in (14), a larger probability of infinite change points can worsen the FDR control. From Table IV, we note that an increase in the number of post-change types  $J$  decreases the FDR/FWER. Because the threshold values increase with  $J$ , and according to (7), the error probability is the worst case probability, we can have a smaller error probability in practice. This means that when  $J$  is large, we may set smaller threshold values in order to control the error at the same level with a better detection delay.

As for ADD, the performance of the FDR control pro-

TABLE III  
DI-FDR AND DI-FWER FOR MULTIPLE CHANGE DETECTION/ISOLATION OF I.I.D. GAUSSIAN DATA STREAMS.

$P3$				
$K$	Procedure	Estimated FDR/FWER	ADD	$\bar{R}$
10	DI-FDR	0.0835	3.59	4.39
	DI-FWER	0.0917	3.88	4.09
100	DI-FDR	0.0790	3.80	43.81
	DI-FWER	0.0787	4.77	40.43
500	DI-FDR	0.0788	3.79	217.31
	DI-FWER	0.0760	5.27	200.27
1000	DI-FDR	0.0790	3.80	435.56
	DI-FWER	0.0800	5.54	401.40

TABLE IV  
DI-FDR AND DI-FWER FOR MULTIPLE CHANGE DETECTION/ISOLATION OF I.I.D.  
GAUSSIAN DATA STREAMS.

$P_4$				
$K$	Procedure	Estimated FDR/FWER	ADD	$R$
10	DI-FDR	0.0031	7.21	9.04
	DI-FWER	0.0028	8.13	9.02
100	DI-FDR	0.0028	7.31	90.44
	DI-FWER	0.0035	9.70	90.20
500	DI-FDR	0.0029	7.32	451.75
	DI-FWER	0.0025	10.97	450.52
1000	DI-FDR	0.0029	7.33	903.22
	DI-FWER	0.0027	11.62	900.79

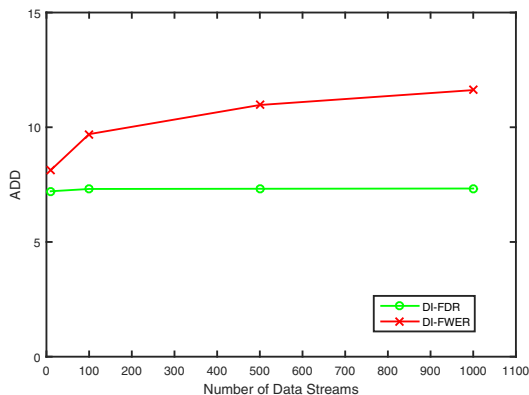


Fig. 1. ADD of DI-FDR and DI-FWER with  $J = 3$ .

cedure is always better than that of the FWER control procedures. With an increase in the number of data streams  $K$ , the ADD is not visibly affected by  $K$  in DI-FDR, but tends to grow logarithmically with  $K$  in DI-FWER (see Fig. 1). In this simulation, when  $J$  changes from 2 to 3, the minimum Kullback-Leibler divergence  $I^*$  does not change, but the ADD still get worse. Recall that the threshold values increase with  $J$ , and for different change types the expected detection delay is different. The ADD is an average expected detection delay over different change types, so it will get worse.

## VI. CONCLUSION

In this paper, we have considered a multiple change detection-isolation problem under a maximum sampling constraint. We have applied the FDR metric to evaluate the performance of the detection-isolation procedure, with a modification of the definition of FDR, and we have proposed a sequential procedure that controls the FDR. For statistical properties, we have established an upper bound on the FDR, and derived an asymptotic upper bound on the ADD. Moreover, we have shown that the proposed procedure is superior to procedures controlling the FWER in terms of ADD. Although our problem has been developed within a Bayesian setting, we have not made use of the prior distribution of changes in the detection-isolation procedure. So there may be other Bayesian statistics for change detection and isolation with improved statistical properties. A further analysis along that direction is work in progress.

## ACKNOWLEDGMENT

The work of J. Chen and W. Zhang was supported in part by the Foundation of China Scholarship Council and the National Natural Science Foundation of China under Grant 61379003, and the work of H. V. Poor was supported in part by the U.S. National Science Foundation under Grants CNS-1456793, CNS-1702808 and ECCS-1343210.

## REFERENCES

- [1] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy—A survey and some new results," *Automatica*, vol. 26, pp. 459-474, 1991.
- [2] I. V. Nikiforov, V. Varavva, and V. Kireichikov, "Application of statistical fault detection algorithms to navigation systems monitoring," *Automatica*, vol. 29, pp. 1275-1290, 1993.
- [3] A. S. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, pp. 601-611, 1976.
- [4] I. V. Nikiforov, "A generalized change detection problem," *IEEE Trans. Inform. Theory*, vol. 41, no. 1, pp. 171-187, 1995.
- [5] T. L. Lai, "Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 595-608, 2000.
- [6] T. Oskiper and H. V. Poor, "Online activity detection in a multiuser environment using the matrix cusum algorithm," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 477-493, 2002.
- [7] I. V. Nikiforov, "A simple recursive algorithm for diagnosis of abrupt changes in random signals," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2740-2746, 2000.
- [8] B. Efron, *Large-scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*, Cambridge, UK: Cambridge University Press, 2010.
- [9] Y. Benjamini, and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Statist. Soc. Ser. B*, vol. 57, pp. 289-300, 1995.
- [10] J. Chen, W. Zhang, and H. V. Poor, "On parallel sequential change detection controlling false discovery rate," in *Proc. 50th Annual Asilomar Conference on Signals, Systems, and Computers*, pp. 107-111, 2016.