Online Quickest Change Detection for Multiple Gaussian Sequences Using Stochastic Bandits

Joshua S. Kartzman *, Benjamin D. Robinson †, Matthew T. Hale ‡

ABSTRACT

This paper considers the online multi-stream quickest change-point detection problem. An agent is faced with a set of independent data streams, one of which contains a change-point at an unknown time step which shifts the mean of its distribution by an unknown amount. The goal of the agent is to minimize its detection delay while controlling for false alarms. Uninterrupted monitoring of every stream can be costly due to resource limitations, so the agent only observes one stream at each point in time. We propose an adaptive algorithm which combines an ϵ-greedy selection rule with a change-point detection algorithm for unknown post-change means. Our main contributions are performance bounds of our algorithm which show that it matches the asymptotic detection delay (to within a constant factor) of single-stream CUSUM. Compared with previous work, our algorithm relies on considerably fewer assumptions.

1. INTRODUCTION

We propose an algorithm for online multi-stream quickest change detection in the case of an unknown post-change mean. All streams are initially identically distributed according to a distribution known to the observer. At an unknown change-point, the mean of an unknown stream's generating distribution shifts by an unknown amount. The agent that is searching for the change-point is constrained to sample only one stream at each time step, which introduces an exploration-exploitation tradeoff in our problem. We combine an ϵ-greedy approach, which induces a small amount of forced exploration, with a change-point detection procedure known as the Generalized Likelihood Ratio (GLR) statistic, which we use to drive exploitation.

Multi-stream quickest change detection with unknown post-change means has been addressed in some prior works, and the most relevant such works are [11] and [3], from which our approach differs in significant ways. Developments in [11] also consider a multi-stream formulation in which only one stream is sampled at each time step, and that work proposes a round-robin algorithm for stream selection. However, a drawback of their approach is that it requires the post-change parameter to have a known lower bound greater than the pre-change parameter, which limits its effectiveness in adversarial settings in which an adversary may attempt to evade detection by deliberately minimizing the changes that they cause, e.g., by minimizing their radar cross section. Work in [3] also employs an ϵ-greedy based strategy in their multi-dimensional change detection algorithm. However, their approach requires the set of possible post-change parameters to be finite, which limits the usability of their algorithm in continuous signal settings.

In this paper, we eliminate these assumptions for the case of univariate Gaussian sequences with an unknown post-change mean shift, and this setup is ubiquitous in radar detection and other remote-sensing problems involving matched filters and outlier detectors. To analyze our approach, we derive worst-case minimax detection delay results for our approach based on a surrogate used in [9]. Then we show that, despite considering a multi-stream setting and relying on weaker assumptions, our approach has an asymptotic detection delay that matches (to within a constant factor) the asymptotic detection delay of a single-stream approach.

2. PROBLEM FORMULATION

We first provide background on online quickest change detection, then state the problem that we solve.

2.1 Background

The cumulative sum (CUSUM) algorithm, introduced in [6], monitors the maximum of a set of partial sums of log-likelihood ratios to detect a change in the distribution of samples from a known distribution f_0 to a known distribution f_1. The statistic (for one stream) at time t is defined as

$$T_t^{\mathrm{CUSUM}} = \max_{0 \le k < t} \sum_{i=k+1}^{t} \log\left(\frac{f_1(X_i)}{f_0(X_i)}\right), \tag{1}$$

where X_i denotes the observation at time i. For a threshold λ, a detection is flagged at the earliest time step t at which T_t^{CUSUM} exceeds λ. At each time t, (1) can be computed in $O(1)$ time with the recursive formula

$$T_t^{\mathrm{CUSUM}} = \max\left\{\log\left(\frac{f_1(X_t)}{f_0(X_t)}\right) + T_{t-1}^{\mathrm{CUSUM}}, 0\right\}.$$

CUSUM was shown to be minimax optimal in [5]. However

*Joshua Kartzman is with the Department of Applied Mathematics and Statistics at SUNY Stony Brook, Stony Brook, NY, 11790 USA. (email: joshua.kartzman@stonybrook.edu)

†Benjamin Robinson is with the Air Force Office of Scientific Research. (email: machine.itel@us.af.mil)

‡Matthew Hale is with the School of Electrical and Computer Engineering at the Georgia Institute of Technology (email: mhale30@gatech.edu)

it requires prior knowledge of f_0 and f_1. The generalized likelihood ratio (GLR) procedure extends CUSUM to the case of an unknown post-change parameter by inserting a maximum likelihood estimator, and it takes the form

$$T_t^{\mathrm{GLR}} = \max_{0 \le k < t} \sup_{\theta \in \Theta} \sum_{i=k+1}^{t} \log \left(\frac{f_\theta(X_i)}{f_0(X_i)} \right), \qquad (2)$$

where Θ is a pre-defined set of possible parameter values. From [9], for an unknown change in the mean of a normal distribution with known variance, the GLR statistic is

$$T_t^{\mathrm{GLR}} = \max_{0 \le k < t} \frac{\left(\sum_{i=k+1}^{t} X_i\right)^2}{2(t-k)}. \qquad (3)$$

The GLR statistic in (2) typically has a per-iteration computational cost of $O(t)$, which can be prohibitive for large t [10]. The FOCuS algorithm from [8] reduces this complexity in the case of detecting a mean change of a Gaussian distribution, and it computes (3) with an $O(\log t)$ per-iteration complexity. At each time t, the FOCuS algorithm maximizes a piecewise-quadratic cost function, where each quadratic corresponds to a possible change-point location. This approach is equivalent to maximizing the CUSUM statistic in (1) over all values of the post-change parameter, and the FOCuS algorithm performs this maximization efficiently by using a pruning strategy. Analysis of the single-stream detection delay of the GLR detection procedure in the case of a univariate mean change of a Gaussian distribution can be found in [9].

2.2 Problem Statement

We consider M independent data sequences. We denote the set of streams as $[M] := \{1, \dots, M\}$. The stream selected and observation generated at time t are denoted as $A_t \in [M]$ and $X_t \in \mathbb{R}$, respectively. Let $A = (A_1, A_2, \dots)$ denote the sampling rule according to which the agent selects the stream to sample at each point in time. The i^{th} observation from stream m is denoted as $X_i^{(m)}$. The known pre-change distribution, whose density is denoted f_0, is $\mathcal{N}(\mu_0, 1)$, where $\mathcal{N}(\mu, \sigma^2)$ denotes a univariate Gaussian distribution with mean μ and variance σ^2. We assume without loss of generality that $\mu_0 = 0$. Let ϕ and Φ denote the standard normal density and distribution functions, respectively. The post-change distribution, whose density is denoted f_1, is $\mathcal{N}(\mu_1, 1)$, where $\mu_1 \neq \mu_0$ and μ_1 is unknown, though the agent knows f_1 is a normal distribution with unit variance. We denote the KL-divergence between f_1 and f_0 as $D(f_1||f_0)$. To ease notation, we assume without loss of generality that stream 1 contains the change-point at time ν. If $t > \nu$ and $A_t = 1$, then $X_t \sim f_1$. Otherwise, $X_t \sim f_0$.

We use T_t to denote the detection statistic that we use, and the stopping time of the change-point detection algorithm is then $\tau = \inf\{t > 0 : T_t \ge \lambda\}$, where λ is the fixed detection threshold that is set when the algorithm is initialized. We denote the σ-algebra generated up to time t as $\mathcal{F}_t := \sigma(A_1, X_1, \dots, A_t, X_t)$. Given M streams and a change-point at time ν in stream 1, we denote the induced probability measure and expected value as $\mathbb{P}_{M,\nu}$ and $\mathbb{E}_{M,\nu}$, respectively. When no change-point exists, we use $\mathbb{P}_{M,\infty}$ and $\mathbb{E}_{M,\infty}$. The following problem adapts the problem formulation from [7].

PROBLEM 1. *Develop a procedure (τ, A) that minimizes the conditional average detection delay*

$$\mathrm{CADD}_M(\tau, A) := \sup_{\nu \ge 0} \mathbb{E}_{M,\nu}[\tau - \nu \mid \tau > \nu],$$

subject to the average run length

$$\mathrm{ARL}_M(\tau, A) := \mathbb{E}_{M,\infty}[\tau] \ge \gamma,$$

for a given constant $\gamma > 0$.

3. ALGORITHM

In this section we provide a summary of our bandit change detector algorithm. This algorithm takes in a detection threshold $\lambda > 0$ and an exploration parameter $\epsilon \in (0,1)$ (whose role we define below), and we name the algorithm ϵ-FOCuS. In it, the symbol $N_t^{(m)}$ denotes the number of observations of stream $m \in [M]$ that have been taken at all times up to and including time t, i.e.,

$$N_t^{(m)} = \sum_{k=1}^{t} \mathbf{1}\{A_k = m\},$$

where $\mathbf{1}\{S\}$ is the indicator function for the event S. The agent maintains a local statistic for each stream $m \in [M]$, which is denoted as $T_t^{(m)}$, and this statistic is generated from stream m's observations $X_1^{(m)}, \dots, X_{N_t^{(m)}}^{(m)}$. The local statistic for stream $m \in [M]$ at time t is calculated as

$$T_t^{(m)} = \max_{0 \le k < N_t^{(m)}} \frac{\left(\sum_{i=k+1}^{N_t^{(m)}} X_i^{(m)}\right)^2}{2\left(N_t^{(m)} - k\right)},$$

which is identical to the single-stream GLR statistic detailed in (3), except instead of being calculated from t observations it is calculated using $N_t^{(m)}$ observations since that is the number of observations of stream m that have been made. The value of this statistic can be computed efficiently in $O(\log t)$ time using the FOCuS algorithm [8]. If $N_t^{(m)} = 0$, then the statistic takes the value of the empty sum, which is 0. The agent's detection statistic at time t is equal to the largest local GLR statistic from the M streams, i.e.,

$$T_t = \max_{m \in [M]} T_t^{(m)}.$$

The ϵ-FOCuS algorithm runs as follows. At each time t, an exploration decision G_t is sampled from a Bernoulli distribution with parameter ϵ, where $\epsilon \in (0,1)$ is the probability of exploration that is initialized at the beginning of the algorithm. If $G_t = 1$, then the agent randomly samples a stream index from $[M]$ and then observes the stream with that index. If $G_t = 0$, then the agent selects the stream whose local GLR statistic is largest based on the previous $t-1$ time steps, namely

$$A_t = \operatorname*{argmax}_{m \in [M]} T_{t-1}^{(m)}.$$

Then the stream with index A_t is sampled, which gives the observation X_t. Using that observation, the statistic $T_t^{(A_t)}$ is updated and the statistic T_t is calculated by taking the maximum of all M GLR statistics. The algorithm is stopped if $T_t \ge \lambda$ and continues to run otherwise.

4. MAIN RESULTS

In this section we present our three main results and then discuss how they can be used to calibrate the parameters of the ϵ-FOCuS algorithm.

4.1 Statements of Results

First, we show that the stream with the changepoint has the largest detection statistic after some finite point in time.

PROPOSITION 1. *Consider an agent using ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0, 1)$ and a detection threshold $\lambda > 0$. Suppose without loss of generality that a change-point exists at time $\nu = 0$ in stream 1. Let H_t denote the event that at time t, one of the local statistics of a stream $m \in [M] \setminus \{1\}$ is greater than or equal to stream 1's, expressed as*

$$H_t := \left\{ \max_{m \in [M] \setminus \{1\}} T_t^{(m)} \geq T_t^{(1)} \right\}.$$

Then there a.s. exists a finite t_0 such that for all $t > t_0$, we have $\mathbf{1}\{H_t\} = 0$.

PROOF. See Section 8.3. □

Next, we upper-bound the amount of time required for the stream with the changepoint to be identified as such.

THEOREM 1. *Consider an agent using ϵ-FOCuS on M streams, and suppose that stream 1 contains a change-point at time $\nu = 0$ that shifts its distribution from $\mathcal{N}(0, 1)$ to $\mathcal{N}(\mu_1, 1)$, where $\mu_1 \neq 0$. For a detection threshold $\lambda > 0$ and an exploration parameter $\epsilon \in (0, 1)$, the expected time until detection τ is bounded via*

$$\mathbb{E}_{M,0}[\tau] \leq \frac{1}{1-\epsilon} \left(\frac{2\lambda(1 + o(1))}{\mu_1^2} + C_{\epsilon,\mu_1,M} \right)$$

as $\lambda \to \infty$, where $C_{\epsilon,\mu_1,M} > 0$ is a constant determined by ϵ, μ_1, and M.

PROOF. See Section 8.3. □

Finally, we lower-bound the expected amount of time before which a false detection occurs.

THEOREM 2. *Consider an agent using ϵ-FOCuS on M streams, where all streams are distributed according to f_0, i.e., no change has occurred. Then given a detection threshold $\lambda > 0$ and an exploration parameter $\epsilon \in (0, 1)$, the expected time until a false detection τ is bounded via*

$$\mathbb{E}_{M,\infty}[\tau] \geq \frac{e^{\lambda}\sqrt{\pi}}{M\sqrt{\lambda}\int_0^\infty xg(x)^2 dx}$$

as $\lambda \to \infty$, where $g(x)$ is defined as

$$g(x) = 2x^{-2}\exp\left[-2\sum_{n=1}^{\infty} n^{-1}\Phi\left(-xn^{1/2}/2\right) \right], \ x > 0.$$

PROOF. See Section 8.3. □

4.2 Discussion and Application of Results

The rate at which the $o(1)$ term goes to 0 as $\lambda \to \infty$ in Theorem 1 depends only on the size of μ_1 since it comes from the single-stream asymptotic performance bound of the GLR procedure from [10]. The constant $C_{\epsilon,\mu_1,M}$ grows as M grows, and $C_{\epsilon,\mu_1,M}$ grows as ϵ and μ_1 approach 0, which is intuitive because these changes make it more difficult to identity the correct stream during greedy selection.

Using Theorems 1 and 2, we compare ϵ-FOCuS in the M-stream setting with CUSUM in the single-stream setting when f_0 and f_1 are both known to CUSUM. CUSUM is optimal for Pollak's minimax formulation [7] in the single-stream setting, which implies that its worst-case expected detection delay

$$\text{CADD}(\tau) := \sup_{\nu \geq 0} \mathbb{E}_{1,\nu}[\tau - \nu \mid \tau > \nu]$$

is minimized, subject to the average run length (the expected stopping time when there is no change-point), namely

$$\text{ARL}(\tau) := \mathbb{E}_{1,\infty}[\tau] \geq \gamma,$$

for some constant $\gamma > 0$. This optimality implies that it matches the asymptotic lower bound of any detector, namely

$$\inf\left\{\text{CADD}(\tau) : \text{ARL}(\tau) \geq \gamma\right\} \geq \frac{\log\gamma}{D(f_1||f_0)}(1 + o(1)) \quad (4)$$

as $\gamma \to \infty$ [4]. As seen in Theorem 1, we use $\mathbb{E}_{M,0}[\tau]$ as a surrogate for the CADD. We do this for two reasons: (i) it well-known that CUSUM attains its worst-case detection delay at $\nu = 0$, so this surrogate makes for an effective comparison [10], and (ii) [9] uses the same surrogate in their analysis of the single-stream GLR procedure. From Theorem 2, the expected stopping time when no change-point occurs is lower bounded as $\mathbb{E}_{M,\infty}[\tau] \geq \frac{e^{\lambda}C}{M\sqrt{\lambda}}$ as the threshold $\lambda \to \infty$. For simplicity of notation we let $C = \frac{\sqrt{\pi}}{\int_0^\infty xg(x)^2 dx}$ since it is constant with respect to the threshold and number of streams. By setting the ARL lower bound equal to γ and solving, we derive the minimum threshold needed to attain an ARL at least as large as γ to be

$$\lambda = (1 + o(1))\log\left(\frac{M\gamma}{C}\right) \quad (5)$$

as $\gamma \to \infty$. For a fixed M, we can substitute our threshold in (5) into the expected detection delay bound in Theorem 1 to upper bound our surrogate for the CADD, which gives

$$\begin{aligned}
\mathbb{E}_{M,0}[\tau] &\leq \frac{1}{1-\epsilon}\left(\frac{2\lambda(1+o(1))}{\mu_1^2} + C_{\epsilon,\mu_1,M}\right) \\
&= \frac{1}{1-\epsilon}\left(\frac{2(1+o(1))\log\left(\frac{M\gamma}{C}\right)(1+o(1))}{\mu_1^2} + C_{\epsilon,\mu_1,M}\right) \\
&= \frac{(1+o(1))\log(\gamma)}{D(f_1||f_0)(1-\epsilon)} \qquad (6)
\end{aligned}$$

as $\gamma \to \infty$. Here, $\log(M)$, $\log(C)$, and $C_{\epsilon,\mu_1,M}$ are fixed as $\gamma \to \infty$ and are absorbed into $o(1)$ and $(1+o(1))(1+o(1)) = (1+o(1))$. Thus, (6) implies that, as $\gamma \to \infty$, the expected detection delay of our algorithm (given an ARL greater than γ) is upper-bounded by Lai's lower bound of any detector [4], as seen in (4), multiplied by a constant which is independent of the number of streams. Therefore, our algorithm is within a constant factor of minimax optimality. Moreover, our bound is within a constant factor of the performance of CUSUM, which requires complete knowledge of both the pre-change and post-change distributions, even though we do not require knowledge of the post-change distribution.

5. SIMULATIONS

We perform Monte Carlo simulations to evaluate the performance of ϵ-FOCuS for $M = 10$ streams and a post-change

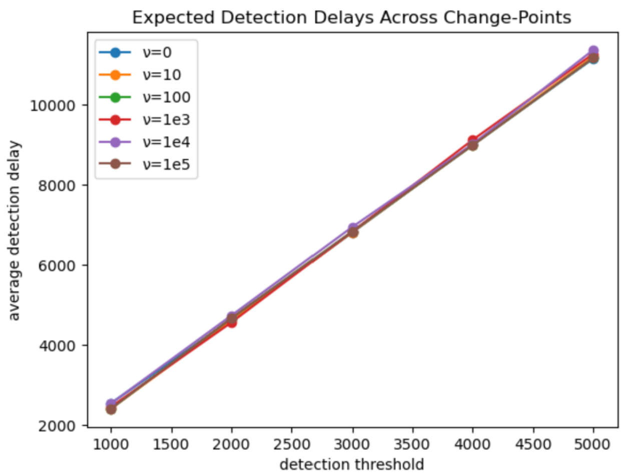

Figure 1: Comparison of expected detection delays of ϵ-FOCuS across different change-point locations and thresholds with $M = 10$, $\mu_1 = 1$, and $\epsilon = 0.1$.

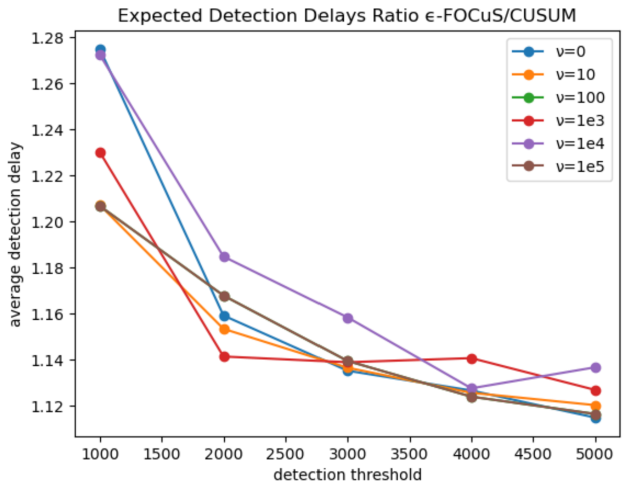

Figure 2: Ratio of asymptotic expected detection delays of ϵ-FOCuS to CUSUM across different change-point locations with $M = 10$, $\mu_1 = 1$, and $\epsilon = 0.1$.

mean of $\mu_1 = 1$. We use an exploration parameter of $\epsilon = 0.1$, and in Figure 1 we compare the expected detection delay across various detection thresholds λ and values of ν. Specifically, Figure 1 shows the value of the detection delay averaged over 50 simulations that were done for each (λ, ν) pair. The results are roughly equivalent for all change-point locations, suggesting the algorithm's performance depends only weakly on the location of ν.

In Figure 2 we compare our algorithm's performance with CUSUM's asymptotic detection delay, which is $\frac{\lambda}{D_{KL}(f_1||f_0)}(1+o(1)) = \frac{2\lambda}{\mu_1^2}(1 + o(1))$ as $\lambda \to \infty$. From Theorem 1, the asymptotic detection delay should be worse by at most a factor of $1/(1 - \epsilon) \approx 1.11$, which we indeed see in Figure 2 as λ grows.

6. CONCLUSION

We presented a novel use case of reinforcement learning for the setting of online multi-stream quickest change detection. The majority of our work was in characterizing the asymptotic performance of our ϵ-greedy change-point detection algorithm. While our results indicate significant progress over previous approaches, this is still ongoing work. The eventual goal will be to replace ϵ-greedy with a stochastic bandit algorithm which attains sub-linear regret such as UCB [2] or Thompson Sampling [1].

7. ACKNOWLEDGEMENTS

This work was supported by the Air Force Office of Scientific Research under grants FA9550-19-1-0169 and RYCOR036, and by the Air Force Research Lab Sensors Directorate.

References

[1] Shipra Agrawal and Navin Goyal. "Near-Optimal Regret Bounds for Thompson Sampling". In: 64.5 (2017).

[2] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. "Finite-time Analysis of the Multiarmed Bandit Problem". In: *Machine learning* 47.2-3 (2002), pp. 235–256.

[3] Aditya Gopalan, Braghadeesh Lakshminarayanan, and Venkatesh Saligrama. "Bandit Quickest Changepoint Detection". In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 29064–29073.

[4] Tze Leung Lai. "Information bounds and quick detection of parameter changes in stochastic systems". In: *IEEE Transactions on Information Theory* 44.7 (1998), pp. 2917–2929.

[5] Gary Lorden. "Procedures for Reacting to a Change in Distribution". In: *The Annals of Mathematical Statistics* 42.6 (1971), pp. 1897–1908.

[6] Ewan S. Page. "Continuous Inspection Schemes". In: *Biometrika* 41.1/2 (1954), pp. 100–115.

[7] Moshe Pollak. "Optimal Detection of a Change in Distribution". In: *The Annals of Statistics* (1985), pp. 206–227.

[8] Gaetano Romano et al. "Fast Online Changepoint Detection via Functional Pruning CUSUM statistics". In: *Journal of Machine Learning Research* 24 (2023), pp. 1–36.

[9] D. Siegmund and E. S. Venkatraman. "Using the Generalized Likelihood Ratio Statistic for Sequential Detection of a Change-Point". In: *The Annals of Statistics* 23.1 (1995), pp. 255 –271.

[10] Haoyun Wang and Yao Xie. "Sequential change-point detection: Computation versus statistical performance". In: *WIREs Computational Statistics* 16.1 (2024).

[11] Qunzhi Xu and Yajun Mei. "Multi-Stream Quickest Detection with Unknown Post-Change Parameters Under Sampling Control". In: *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE. 2021, pp. 112–117.

8. APPENDIX

8.1 Basic Facts

For the Chernoff-Hoeffding bound, we use the formulation given by [1], which we restate here.

FACT 1 (CHERNOFF-HOEFFDING BOUND). *Let $X_1, \dots, X_n$ be independent random variables with support $[0,1]$ such that $\mathbb{E}[X_i] = \mu, \forall i$. Let $S_n = \sum_{i=1}^n X_i$. For all $a \geq 0$, we have*

$$\mathbb{P}(S_n \geq n\mu + a) \leq e^{-2a^2/n}, \quad \mathbb{P}(S_n \leq n\mu - a) \leq e^{-2a^2/n}.$$

FACT 2 (BOREL-CANTELLI LEMMA). *For a sequence of events $A_1, A_2, \dots,$ if $\sum_{n=1}^\infty \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(A_n \text{ i.o.}) = 0$, i.e., only finitely many events A_n occur a.s.*

8.2 Supporting Lemmas

LEMMA 1. *Consider an agent using ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0,1)$. Suppose without loss of generality that a change-point exists at time $\nu = 0$ in stream 1. Then*

$$\sum_{t=0}^\infty \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right) < \infty.$$

PROOF. We denote the sample mean of stream 1's observations at time t as $\hat{\mu}_t^{(1)} = \frac{\sum_{i=1}^{N_t^{(1)}} X_i^{(1)}}{N_t^{(1)}}$. At time t,

$$\mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right)$$
$$\leq \mathbb{P}_{M,0}\left(\frac{\left(\sum_{i=1}^{N_t^{(1)}} X_i^{(1)}\right)^2}{2N_t^{(1)}} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right)$$
$$= \mathbb{P}_{M,0}\left(|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{N_t^{(1)} M}}, N_t^{(1)} > \frac{\epsilon t}{2M}\right), \quad (7)$$

where the first inequality follows because

$$T_t^{(1)} = \max_{0 \leq k < N_t^{(1)}} \frac{\left(\sum_{i=k+1}^{N_t^{(1)}} X_i^{(1)}\right)^2}{2(N_t^{(1)} - k)} \geq \frac{\left(\sum_{i=1}^{N_t^{(1)}} X_i^{(1)}\right)^2}{2N_t^{(1)}}.$$

Lowering bounding $N_t^{(1)}$ as $\frac{\epsilon t}{2M}$, we bound (7) as

$$\mathbb{P}_{M,0}\left(|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{N_t^{(1)} M}}, N_t^{(1)} > \frac{\epsilon t}{2M}\right)$$
$$\leq \mathbb{P}_{M,0}\left(|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|}{\sqrt{2}}, N_t^{(1)} > \frac{\epsilon t}{2M}\right).$$

By the strong law of large numbers, $|\hat{\mu}_t^{(1)}| \xrightarrow{a.s.} |\mu_1|$. Therefore, as $N_t^{(1)}$ grows, $\left\{|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|}{\sqrt{2}}\right\}$ holds for finitely many t, giving $\sum_{t=0}^\infty \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right) < \infty$. ☐

LEMMA 2. *Consider an agent using ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0,1)$, and suppose without loss of generality that a change-point exists at time $\nu = 0$ in stream 1. Then*

$$\sum_{t=0}^\infty \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} \leq \frac{\epsilon t}{2M}\right) < \infty.$$

PROOF. At time t, independent of $\mathcal{F}_{t-1}$, the probability of picking stream 1 is at least $\frac{\epsilon}{M}$ simply from exploration. Then the Chernoff-Hoeffding (Fact 1) gives

$$\mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} \leq \frac{\epsilon t}{2M}\right)$$
$$\leq \mathbb{P}_{M,0}\left(N_t^{(1)} \leq \frac{\epsilon t}{2M}\right) \leq \exp\left(-\frac{\epsilon^2 t}{2M^2}\right). \quad (8)$$

Since (8) is exponentially decreasing, its sum from $t = 0$ to ∞ is finite, and the result follows. ☐

LEMMA 3. *Consider an agent using ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0,1)$, and suppose without loss of generality that a change-point exists at time $\nu = 0$ in stream 1. Then*

$$\sum_{t=0}^\infty \mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}, T_t^{(1)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right) < \infty.$$

PROOF. At time t,

$$\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}, T_t^{(1)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right) \leq \mathbb{P}_{M,0}\left(T_t^{(2)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right)$$
$$\leq \mathbb{P}_{M,0}\left(\max_{0 \leq i < j \leq t} \frac{\left|\sum_{k=i+1}^j X_k^{(2)}\right|}{\sqrt{j-i}} \geq \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right), \quad (9)$$

where (9) follows since the event

$$\left\{\max_{0 \leq k < N_t^{(2)}} \frac{\left(\sum_{k=i+1}^{N_t^{(2)}} X_k^{(2)}\right)^2}{2\left(N_t^{(2)} - i\right)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right\},$$

implies the event

$$\left\{\max_{0 \leq i < j \leq t} \frac{\left|\sum_{k=i+1}^j X_k^{(2)}\right|}{\sqrt{j-i}} \geq \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right\}.$$

From Proposition 1 in [9], a constant $C > 0$ exists such that

$$\mathbb{P}_{M,0}\left(\max_{0 \leq i < j \leq t} \frac{|\sum_{k=i+1}^j X_k^{(2)}|}{\sqrt{j-i}} \geq \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right)$$
$$\sim Ct\left(\frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right)\phi\left(\frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right). \quad (10)$$

Since (10) is exponentially decreasing, its sum from $t = 0$ to ∞ is finite, and the result follows. ☐

8.3 Proofs of Main Results

PROOF OF PROPOSITION 1. Applying the union bound, we split the probability of H_t at time t into the sum of the probabilities that each of the statistics for streams $m \neq 1$ exceeds $T_t^{(1)}$, which gives

$$\mathbb{P}_{M,0}(H_t) = \mathbb{P}_{M,0}\left(\max_{m \in [M]\setminus\{1\}} T_t^{(m)} \geq T_t^{(1)}\right)$$
$$\leq \sum_{m=2}^M \mathbb{P}_{M,0}\left(T_t^{(m)} \geq T_t^{(1)}\right) = (M-1)\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}\right), \quad (11)$$

where (11) follows since the statistics for streams $m \neq 1$ are identically distributed. We can bound $\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}\right)$

as

$$\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}\right) \leq \mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}, T_t^{(1)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right) + \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right) + \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} \leq \frac{\epsilon t}{2M}\right).$$

Applying Lemmas 1, 2, and 3 gives $\sum_{t=0}^{\infty} \mathbb{P}_{M,0}(H_t) < \infty$. From the Borel-Cantelli lemma (Fact 2), H_t a.s. holds true for only finitely many t. □

PROOF OF THEOREM 1. The detection delay can be partitioned into the time spent observing stream 1 and the time spent observing other streams as

$$\mathbb{E}_{M,0}[\tau] = \mathbb{E}_{M,0}\left[N_\tau^{(1)}\right] + \mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 0, A_t \neq 1\}\right] + \mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 1, A_t \neq 1\}\right]. \quad (12)$$

From Proposition 1, there a.s. exists a finite time step after which streams $m \neq 1$ are not selected during exploitation, so there exists a finite constant $C_{\epsilon,\mu_1,M} > 0$ (which depends on ϵ, μ_1, and M) such that

$$\mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 0, A_t \neq 1\}\right] \leq C_{\epsilon,\mu_1,M}. \quad (13)$$

At any time t, $G_t = 1$ happens with probability ϵ. Given $G_t = 1$, $A_t \neq 1$ happens with probability $\frac{M-1}{M}$. Therefore

$$\mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 1, A_t \neq 1\}\right] = \frac{\epsilon(M-1)}{M}\mathbb{E}_{M,0}[\tau]. \quad (14)$$

When $M = 1$, the expected number of observations from stream 1 before a detection is flagged within stream 1 is equal to the expected value of τ. However, in the multi-stream ($M > 1$) setting, detection can be flagged from a stream other than stream 1, in which case stream 1 would have generated fewer samples than it did in the single-stream setting. Then the expected value of $N_\tau^{(1)}$ in the multi-stream ($M > 1$) setting is upper-bounded by the expected value of τ in the single-stream ($M = 1$) setting. Formally,

$$\mathbb{E}_{M,0}\left[N_\tau^{(1)}\right] \leq \mathbb{E}_{1,0}[\tau]. \quad (15)$$

Since the GLR procedure is asymptotically optimal in the single-stream setting [10], we have

$$\mathbb{E}_{1,0}[\tau] \leq \sup_{\nu \geq 0} \mathbb{E}_{1,\nu}[\tau - \nu \mid \tau > \nu] = \frac{\lambda(1 + o(1))}{D(f_1||f_0)}, \quad (16)$$

as the threshold $\lambda \to \infty$. Using (13), (14), (15), and (16) in (12) gives

$$\begin{aligned}\mathbb{E}_{M,0}[\tau] &\leq \frac{\lambda(1 + o(1))}{D(f_1||f_0)} + C_{\epsilon,\mu_1,M} + \frac{\epsilon(M-1)}{M}\mathbb{E}_{M,0}[\tau] \\ &\leq \frac{2\lambda(1 + o(1))}{\mu_1^2} + C_{\epsilon,\mu_1,M} + \epsilon\mathbb{E}_{M,0}[\tau],\end{aligned}$$

which follows since $D(f_1||f_0) = \frac{\mu_1^2}{2}$. The result follows from solving for $\mathbb{E}_{M,0}[\tau]$. □

PROOF OF THEOREM 2. Given a threshold $\lambda > 0$, for any $t_0 \in \mathbb{N}$ we have

$$\begin{aligned}&\mathbb{P}_{M,\infty}(\tau > t_0) \\ &= \mathbb{P}_{M,\infty}\left(\bigcap_{m=1}^{M}\left\{\max_{0 \leq i < j \leq N_{t_0}^{(m)}} \frac{\left(\sum_{k=i+1}^{j} X_k^{(m)}\right)^2}{2(j-i)} < \lambda\right\}\right) \\ &\geq \mathbb{P}_{M,\infty}\left(\bigcap_{m=1}^{M}\left\{\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_k^{(m)}\right)^2}{2(j-i)} < \lambda\right\}\right) \\ &= \prod_{m=1}^{M} \mathbb{P}_{M,\infty}\left(\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_k^{(m)}\right)^2}{2(j-i)} < \lambda\right) \\ &= \mathbb{P}_{M,\infty}\left(\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_k^{(1)}\right)^2}{2(j-i)} < \lambda\right)^M. \quad (17)\end{aligned}$$

The first line of (17) follows since the stopping time is only greater than t_0 if all of the GLR statistics produced up to time t_0 are less than λ. The next step follows since $N_{t_0}^{(m)} \leq t_0$ for all $m \in [M]$, which is the case since $\sum_{m \in [M]} N_{t_0}^{(m)} = t_0$. If all $\{(i,j) : 0 \leq i < j \leq t_0\}$ produce statistics less than λ, then all $\left\{(i,j) : 0 \leq i < j \leq N_{t_0}^{(m)}\right\}$ produce statistics less than λ. The converse may not be true, so the probability of not flagging a detection in stream $m \in [M]$ after $N_{t_0}^{(m)}$ samples is more likely than the probability of not flagging a detection after t_0 samples from stream $m \in [M]$. The next line of (17) follows since the events are independent. The last line of (17) follows since all of the streams are identically distributed as f_0 in the no-change scenario. From Proposition 1 in [9], the probability of the event

$$\left\{\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_k^{(1)}\right)^2}{2(j-i)} < \lambda\right\}$$

is equal to the probability of the event $\{\tau > t_0\}$ when $M = 1$ and $\nu = \infty$. Using this equivalence in (17), we see that

$$\mathbb{P}_{M,\infty}(\tau > t_0) \geq \mathbb{P}_{1,\infty}(\tau > t_0)^M.$$

From Theorem 1 of [9], when $M = 1$ and $\nu = \infty$ the value of τ is asymptotically exponentially distributed with expected value

$$\mathbb{E}_{1,\infty}[\tau] = \frac{e^{\lambda}\sqrt{\pi}}{\sqrt{\lambda}\int_0^{\infty} xg(x)^2 dx}$$

as $\lambda \to \infty$. Calculating the formula for the expected value of an exponential random variable using tail integration gives

$$\begin{aligned}\mathbb{E}_{M,\infty}[\tau] &= \sum_{t_0=0}^{\infty} \mathbb{P}_{M,\infty}(\tau > t_0) \geq \sum_{t_0=0}^{\infty} \mathbb{P}_{1,\infty}(\tau > t_0)^M \\ &\sim \int_0^{\infty} \exp(-Mt_0/\mathbb{E}_{1,\infty}[\tau]) dt_0 \\ &= \frac{\mathbb{E}_{1,\infty}[\tau]}{M} = \frac{e^{\lambda}\sqrt{\pi}}{M\sqrt{\lambda}\int_0^{\infty} xg(x)^2 dx}\end{aligned}$$

as $\lambda \to \infty$. □