Online Quickest Change Detection for Multiple Gaussian Sequences Using Stochastic Bandits

Joshua S. Kartzman *, Benjamin D. Robinson †

ABSTRACT

This paper considers the online mutli-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. Streams can model, for example, a suite of sensors or radar antennas. 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 power or resource limitations, so we consider a constrained sampling formulation where the agent only observes one stream during each round. Leveraging ideas from stochastic bandits and reinforcement learning, we examine this problem through the lens of an exploration-exploitation tradeoff. 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 it matches the asymptotic detection delay within a constant factor of single-stream CUSUM in the complete information regime. Compared with previous work, our algorithm relies on considerably fewer assumptions. We envision applications of our work in radar, industrial process management, and fault detection.

1. INTRODUCTION

We propose an algorithm for online multi-stream quickest change detection in the case of an unknown post-change parameter. 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 is constrained to sampling only one stream at each time step, which introduces an exploration-exploitation tradeoff in our problem. We combine ϵ-greedy, which induces a small amount of forced exploration, with a change-point detection procedure known as the Generalized Likelihood Ratio (GLR) statistic.

*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)

While multi-stream quickest change detection in the case of unknown post-change means has been addressed in works such as [11], [10], and [3], our approach differs in significant ways. Compared to [10], where all M streams can be sampled at each time step, our setting constrains the agent to observing a single stream. [11] also considers the same multi-stream constrained sampling formulation and 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 where an enemy agent can slip through detection. [3] also employs an ϵ-greedy based strategy in their multi-dimensional change detection algorithm. However, a limitation of their approach is that it requires the post-change parameter set to be finite, which limits the usability of their algorithm in continuous signal settings.

In this paper, we eliminate these assumptions in the special case of univariate Gaussian sequences with an unknown post-change mean shift. Although this case may seem specialized, it is ubiquitous in radar detection and other remote-sensing problems involving matched filters and outlier detectors. We derive worst-case minimax detection delay results based on a surrogate used in [9].

We provide some background on online quickest change detection. CUSUM, introduced by Page in [6], monitors the maximum of a set of partial sums of log-likelihood ratios to detect a change. The statistic at time t is defined as

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

where X_i denotes the observation at time i and f_0 and f_1 denote the pre-change and post-change densities, respectively. For a threshold λ, a detection is flagged at the earliest time step such that T_t^{CUSUM} exceeds λ. At each time t, (1) can be computed in $O(1)$ time with the recursive formula

$$T_t^{\mathrm{CUSUM}} = \max \left\{ \log \left(\frac{f_1(X_t)}{f_0(X_t)} \right) + T_{t-1}^{\mathrm{CUSUM}}, 0 \right\}. \tag{2}$$

CUSUM was shown to be minimax optimal by Lorden [5]. However it requires prior knowledge of f_0 and f_1.

The GLR procedure extends CUSUM to the case of an unknown post-change parameter by inserting a maximum likelihood estimator:

$$T_t^{\mathrm{GLR}} = \max_{0 \le k \le t} \sup_{\theta \in \boldsymbol{\Theta}} \sum_{i=k+1}^{t} \log \left(\frac{f_\theta(X_i)}{f_0(X_i)} \right). \tag{3}$$

From [9], for a change in mean of a normal distribution with known variance, the statistic is

$$T_t^{\mathrm{GLR}} = \max_{0\le k<t} \frac{\left(\sum_{i=k+1}^{t} X_i\right)^2}{2(t-k)}. \tag{4}$$

(3) typically has a per-iteration computational cost of $O(t)$. The FOCuS algorithm from [8] solves this in the case of detecting a mean change of a Gaussian distribution, and computes (4) with a $O(\log t)$ per-iteration cost. At each time t, the algorithm maximizes a piecewise quadratic cost function $Q_t(\mu)$, where each quadratic corresponds to a possible change-point location. This is equivalent to maximizing the Page-CUSUM statistic in (1) over all values of the post-change parameter. The algorithm performs this efficiently by pruning. Analysis on 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. PROBLEM FORMULATION

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 switching rule, according to which the agent selects streams. 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 $\mathcal{N}(\mu_0, 1)$. 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 $\mu_1 \neq \mu_0$. μ_1 is unknown, but the agent knows f_1 is a normal distribution with unit variance. Unknown to the agent, we assume without loss of generality that stream 1 contains the change-point at time ν. If $t > \nu$ and $A_t = 1$, $X_t \sim f_1$. Otherwise, $X_t \sim f_0$. The detection statistic is denoted T_t. The stopping time is

$$\tau = \inf\{t > 0 : T_t \ge \lambda\}, \tag{5}$$

where λ is the fixed pre-initialized detection threshold. We denote the σ-algebra generated by the history 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 probability measure and expected value as $\mathbb{P}_{M,\nu}$ and $\mathbb{E}_{M,\nu}$. When no change-point exists, we use $\mathbb{P}_{M,\infty}$ and $\mathbb{E}_{M,\infty}$. Adapting Pollak's formulation [7], we seek to develop a procedure (τ, A) which minimizes the conditional average detection delay

$$\mathrm{CADD}_M(\tau, A) := \sup_{\nu\ge 0} \mathbb{E}_{M,\nu}[\tau - \nu | \tau > \nu], \tag{6}$$

subject to the average run length

$$\mathrm{ARL}_M(\tau, A) := \mathbb{E}_{M,\infty}[\tau]. \tag{7}$$

As done in [9], we consider $\mathbb{E}_{M,0}[\tau]$ as a surrogate for (6).

3. ALGORITHM

We provide a summary of our bandit change detector algorithm, which we name ϵ-FOCuS. $N_t^{(m)}$ denotes the number of observations of stream $m \in [M]$ after time t. The agent maintains a local statistic for each stream $m \in [M]$, which is denoted as $T_t^{(m)}$. This is generated from stream m's observations $X_1^{(m)}, \dots, X_{N_t^{(m)}}^{(m)}$. The statistic for stream $m \in [M]$ at time t is calculated as

$$T_t^{(m)} = \max_{0\le k<N_t^{(m)}} \frac{\left(\sum_{i=k+1}^{N_t^{(m)}} X_i^{(m)}\right)^2}{2\left(N_t^{(m)} - k\right)}, \tag{8}$$

which is identical to the single-stream GLR statistic detailed in (4) except instead of being calculated from t observations its calculated using $N_t^{(m)}$ observations. This can be computed efficiently in $O(\log t)$ time using the FOCuS algorithm [8]. If $N_t^{(m)} = 0$, the statistic corresponds to the empty sum and is set to 0. The agent's detection statistic at time t is equal to the largest local GLR statistic from the M streams:

$$T_t = \max_{m\in[M]} T_t^{(m)}. \tag{9}$$

At time t, an exploration decision G_t is sampled from a Bernoulli(ϵ) distribution, where $\epsilon \in (0, 1)$ corresponds to the probability of exploration initialized at the beginning of the algorithm. If $G_t = 1$ the agent randomly samples from $[M]$. Otherwise the agent selects the stream whose local GLR statistic is largest based on the previous $t - 1$ time steps:

$$A_t = \operatorname*{argmax}_{m\in[M]} T_{t-1}^{(m)} \tag{10}$$

Upon generating X_t from A_t, each $T_t^{(m)}$ is calculated and the algorithm is stopped if $T_t \ge \lambda$.

4. MAIN RESULT

PROPOSITION 1. *Consider an agent playing ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0, 1)$. 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 streams $m \in [M] \setminus \{1\}$ is greater than or equal to stream 1's:*

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

There a.s. exists a finite t_0 such that $\forall t > t_0$, $\mathbf{1}\{H_t\} = 0$.

THEOREM 1. *Consider an agent playing ϵ-FOCuS on M streams. Stream 1 contains a change-point at time $\nu = 0$ shifting 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*

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

where $C_{\epsilon,\mu_1,M}$ is a constant determined by ϵ, μ_1, and M.

THEOREM 2. *Consider an agent playing ϵ-FOCuS on M streams, where all streams are distributed according to f_0. Given a detection threshold $\lambda > 0$ and an exploration parameter ϵ, the expected time until a false detection τ is*

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

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

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

5. DISCUSSION

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

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

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

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

matches the lower bound of any detector,

$$\inf \{\text{CADD}(\tau) : \text{ARL}(\tau) \geq \gamma\} \geq \frac{\log \gamma}{D(f_1 || f_0)}(1 + o(1)), \quad (11)$$

as $\gamma \to \infty$, established in [4]. Here $D(f_1||f_0)$ denotes the KL-divergence and $\mathbb{E}_\nu$ denotes the expected value in the normal single-stream setting with a change-point at ν.

For our algorithm, as seen in Theorem 1, we use $\mathbb{E}_{M,0}[\tau]$ as a surrogate for $\sup_{\nu \geq 0} \mathbb{E}_\nu[\tau - \nu | \tau > \nu]$. We do this for two reasons: it well-known that CUSUM attains its worst-case detection delay at $\nu = 0$ so it makes for an effective comparison, and [9] uses the same surrogate in their analysis of the single-stream GLR procedure so its well-established in the literature. From Theorem 2, given a threshold $\lambda > 0$, 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 $\lambda \to \infty$. For simplicity of notation we let $C = \frac{\sqrt{\pi}}{\int_0^\infty x g(x)^2 dx}$ since it is constant with respect to the threshold or number of streams. By setting the ARL lower bound equal to γ and solving, we can derive the minimum threshold necessary to attain an ARL at least as large as γ:

$$\lambda = (1 + o(1)) \log \left(\frac{M\gamma}{C} \right) \quad (12)$$

as $\gamma \to \infty$. For a fixed M, we can substitute our threshold in (12) into the expected detection delay bound in Theorem 1 to upper bound our surrogate for the CADD:

$$\begin{aligned} \mathbb{E}_{M,0}[\tau] &\leq \frac{2(1 + o(1)) \log\left(\frac{M\gamma}{C}\right) + C_{\epsilon,\mu_1,M}}{\mu_1^2 (1 - \epsilon)} \\ &\leq \frac{(1 + o(1)) \log (\gamma)}{D(f_1||f_0) (1 - \epsilon)}, \end{aligned} \quad (13)$$

as $\gamma \to \infty$. $\log(M)$, $\log(C)$, and $C_{\epsilon,\mu_1,M}$ are fixed as $\gamma \to \infty$ and are absorbed into $o(1)$. $D(f_1||f_0)$ follows since the KL-divergence between $f_1 \sim \mathcal{N}(\mu_1, 1)$ and $f_0 \sim \mathcal{N}(0, 1)$ is $\frac{\mu_1^2}{2}$.

(13) 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 [4] lower bound of any detector, as seen (11), multiplied by a constant which is independent of the number of streams, making our algorithm within a constant factor of minimax optimality. Moreover, our asymptotic bound is within a constant factor of the performance of CUSUM, an algorithm which requires complete information of both the pre-change and post-change distributions.

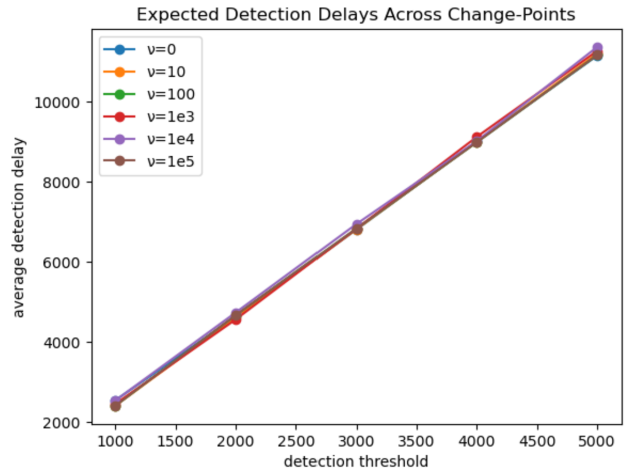

Figure 1: Comparison of expected detection delays of ϵ-FOCuS across different change-point locations and thresholds with $M = 10$, $\mu = 1$, and $\epsilon = 0.1$.

6. SIMULATIONS

We perform Monte Carlo simulations to evaluate the performance of ϵ-FOCuS. In Figure 1 we compare the expected detection delay averaged over 50 simulations across various detection thresholds and values of ν. We set $M = 10$ and $\mu_1 = 1$. We use an exploration parameter of $\epsilon = 0.1$. The results are roughly equivalent for all change-point locations, suggesting the algorithm's performance is invariant to the location of ν. In Figure 2 we compare our algorithm's performance with CUSUM's asymptotic detection delay $\frac{\lambda}{D_{KL}(f_1||f_0)}(1 + o(1)) = \frac{2\lambda}{\mu_1^2}(1 + o(1))$ 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 can be seen in our simulations.

7. 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 based change-point detector algorithm. While our results indicate significant progress over previous approaches, this is still an 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].

8. 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. We would also like to thank Professor Matthew Hale from Georgia Tech for his help with proofreading.

References

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

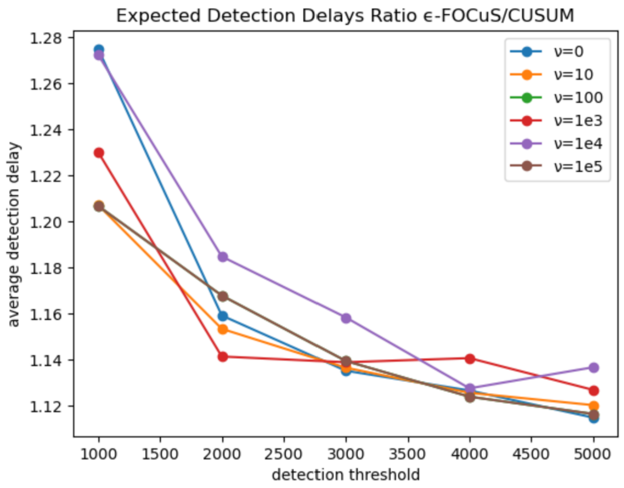

Figure 2: Ratio of asymptotic expected detection delays of ϵ-FOCuS to CUSUM across different change-point locations with $M = 10$, $\mu = 1$, and $\epsilon = 0.1$.

[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] Yao Xie and David Siegmund. "Sequential Multi-Sensor Change-Point". In: *The Annals of Statistics* 41.2 (2013), pp. 670 –692.

[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.

9. APPENDIX

9.1 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 $\mathbb{E}[X_t|X_1,\dots,X_{t-1}] = \mu$. Let $S_n = \sum_{i=1}^n X_n$. For all $a \geq 0$,*

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

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

9.2 Supporting Lemmas

LEMMA 1. *Consider an agent playing on $M > 1$ streams with an exploration parameter $\epsilon \in (0,1)$. A change-point exists at time $\nu = 0$ in stream 1. The following holds:*

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

PROOF OF LEMMA 1. We denote the sample mean of the observations of stream 1 up to time t as

$$\hat{\mu}_t^{(1)} = \frac{\sum_{i=1}^{N_t^{(1)}} X_i^{(1)}}{N_t^{(1)}}.$$

At time t,

$$\begin{aligned} &\mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right) \\ &\leq \mathbb{P}_{M,0}\left(\frac{\left(\sum_{i=1}^{N_t^{(1)}} X_i^{(1)}\right)^2}{2N_t^{(1)}} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right) \\ &= \mathbb{P}_{M,0}\left(|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{N_t^{(1)} M}}, N_t^{(1)} > \frac{\epsilon t}{2M}\right), \end{aligned} \quad (14)$$

where the first inequality follows because

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

Since the event

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

implies the event $\left\{|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|}{\sqrt{2}}\right\}$ from lowering bounding $N_t^{(1)}$ in $\frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{N_t^{(1)} M}}$, we bound (14) as

$$\begin{aligned} &\mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right) \\ &\leq \mathbb{P}_{M,0}\left(|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|}{\sqrt{2}}, N_t^{(1)} > \frac{\epsilon t}{2M}\right). \end{aligned}$$

By the strong law of large numbers, $|\hat{\mu}_t^{(1)}| \xrightarrow{a.s.} |\mu_1|$. Therefore, as $N_t^{(1)}$ grows, $\left\{|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|}{\sqrt{2}}\right\}$ can hold for at most finitely many t, implying

$$\begin{aligned}&\sum_{t=0}^{\infty} \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right)\\ &\quad\leq \mathbb{E}_{M,0}\left[\sum_{t=0}^{\infty} \mathbf{1}\left\{|\hat{\mu}_t^{(1)}| < \frac{|\mu_1|}{\sqrt{2}}, N_t^{(1)} > \frac{\epsilon t}{2M}\right\}\right] < \infty.\end{aligned}$$

□

Lemma 2. *Consider an agent playing ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0,1)$. A change-point exists at time $\nu = 0$ in stream 1. The following holds:*

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

Proof of Lemma 2. At time t, independent of $\mathcal{F}_{t-1}$, the probability of picking stream 1 is at least $\frac{\epsilon}{M}$ simply from exploration. From the Chernoff-Hoeffding bound (Fact 1):

$$\begin{aligned}&\mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} \leq \frac{\epsilon t}{2M}\right)\\ &\quad\leq \mathbb{P}_{M,0}\left(N_t^{(1)} \leq \frac{\epsilon t}{2M}\right) \leq \exp\left(-\frac{\epsilon^2 t}{2M^2}\right). \qquad (15)\end{aligned}$$

Since (15) is exponentially decreasing,

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

□

Lemma 3. *Consider an agent playing ϵ-FOCuS on $M > 1$ streams with an exploration parameter $\epsilon \in (0,1)$. A change-point exists at time $\nu = 0$ in stream 1. The following holds:*

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

Proof of Lemma 3. At time t,

$$\begin{aligned}&\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}, T_t^{(1)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right)\\ &\quad\leq \mathbb{P}_{M,0}\left(T_t^{(2)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right)\\ &\quad\leq \mathbb{P}_{M,0}\left(\max_{0\leq i<j\leq t} \frac{\left|\sum_{k=i+1}^{j} X_i^{(2)}\right|}{\sqrt{j-i}} \geq \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right). \qquad (16)\end{aligned}$$

(16) follows since the event

$$\left\{\max_{0\leq k<N_t^{(2)}} \frac{\left(\sum_{k=i+1}^{N_t^{(2)}} X_i^{(2)}\right)^2}{2\left(N_t^{(2)} - i\right)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right\}, \qquad (17)$$

implies

$$\left\{\max_{0\leq i<j\leq t} \frac{\left|\sum_{k=i+1}^{j} X_i^{(2)}\right|}{\sqrt{j-i}} \geq \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right\} \qquad (18)$$

because $N_t^{(2)} \leq t$. Applying Proposition 1 from [9], there exists a constant $C > 0$ such that

$$\begin{aligned}&\mathbb{P}_{M,0}\left(T_t^{(2)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right)\\ &\quad\leq \mathbb{P}_{M,0}\left(\max_{0\leq i<j\leq t} \frac{|\sum_{k=i+1}^{j} X_i^{(2)}|}{\sqrt{j-i}} \geq \frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right)\\ &\quad\sim Ct\left(\frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right)\phi\left(\frac{|\mu_1|\sqrt{\epsilon t}}{2\sqrt{M}}\right). \qquad (19)\end{aligned}$$

Since (19) is exponentially decreasing,

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

□

9.3 Proofs of Main Results

Proof of Proposition 1. Applying the union bound, we bound 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 $T_t^{(1)}$:

$$\begin{aligned}\mathbb{P}_{M,0}(H_t) &= \mathbb{P}_{M,0}\left(\max_{m\in[M]\setminus\{1\}} T_t^{(m)} \geq T_t^{(1)}\right)\\ &\leq \sum_{m=2}^{M} \mathbb{P}_{M,0}\left(T_t^{(m)} \geq T_t^{(1)}\right)\\ &= (M-1)\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}\right). \qquad (20)\end{aligned}$$

(20) 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

$$\begin{aligned}\mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}\right) &\leq \mathbb{P}_{M,0}\left(T_t^{(2)} \geq T_t^{(1)}, T_t^{(1)} \geq \frac{\mu_1^2 \epsilon t}{8M}\right)\\ &\quad + \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} > \frac{\epsilon t}{2M}\right)\\ &\quad + \mathbb{P}_{M,0}\left(T_t^{(1)} < \frac{\mu_1^2 \epsilon t}{8M}, N_t^{(1)} \leq \frac{\epsilon t}{2M}\right). \qquad (21)\end{aligned}$$

Applying Lemmas 1, 2, and 3:

$$\sum_{t=0}^{\infty} \mathbb{P}_{M,0}(H_t) < \infty.$$

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 amount of time observing stream 1 and the amount of time spent exploring and exploiting other streams:

$$\begin{aligned}\mathbb{E}_{M,0}[\tau] &= \mathbb{E}_{M,0}\left[N_\tau^{(1)}\right] + \mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 0, A_t \neq 1\}\right]\\ &\quad + \mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 1, A_t \neq 1\}\right]. \qquad (22)\end{aligned}$$

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 $C > 0$ such that

$$\mathbb{E}_{M,0}\left[\sum_{t=1}^{\tau} \mathbf{1}\{G_t = 0, A_t \neq 1\}\right] = C.$$

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

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

The expected number of observations from stream 1 before a detection is flagged within stream 1 is equal to the expected value of τ when $M = 1$. Since a detection can be flagged from another stream, in which case stream 1 would have generated fewer samples, the expected value of $N_\tau^{(1)}$ is upper bounded by the expected value of τ when $M = 1$:

$$\mathbb{E}_{M,0}\left[N_\tau^{(1)}\right] \leq \mathbb{E}_{1,0}[\tau] \tag{23}$$

From [9], the expected stopping time in the single-stream setting given a detection threshold $\lambda > 0$ is

$$\mathbb{E}_{1,0}[\tau] \approx \frac{2\lambda - 3}{\mu_1^2} + \frac{4\rho}{\mu_1},$$

where $\rho \approx 0.583$. Substituting all of the above into (22):

$$\mathbb{E}_{M,0}[\tau] \leq \frac{2\lambda - 3}{\mu_1^2} + \frac{4\rho}{\mu_1} + C + \frac{\epsilon(M-1)}{M}\mathbb{E}_{M,0}[\tau],$$

which implies

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

where we express the sum of the constants as $C_{\epsilon,\mu_1,M}$ to denote its dependence on ϵ, μ_1, and M. $\square$

Proof of Theorem 2. Given a threshold $\lambda > 0$, for any $t_0 \in \mathbb{N}$,

$$\begin{aligned}
&\mathbb{P}_{M,\infty}(\tau > t_0) \\
&= \mathbb{P}_{M,\infty}\left(\bigcap_{m=1}^{M}\left\{\max_{0 \leq i < j \leq N_{t_0}^{(m)}} \frac{\left(\sum_{k=i+1}^{j} X_i^{(m)}\right)^2}{2(j-i)} < \lambda\right\}\right) \\
&\geq \mathbb{P}_{M,\infty}\left(\bigcap_{m=1}^{M}\left\{\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_i^{(m)}\right)^2}{2(j-i)} < \lambda\right\}\right) \\
&\geq \prod_{m=1}^{M} \mathbb{P}_{M,\infty}\left(\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_i^{(m)}\right)^2}{2(j-i)} < \lambda\right) \\
&\geq \mathbb{P}_{M,\infty}\left(\max_{0 \leq i < j \leq t_0} \frac{\left(\sum_{k=i+1}^{j} X_i^{(1)}\right)^2}{2(j-i)} < \lambda\right)^M.
\end{aligned} \tag{24}$$

The first line of (24) follows since the stopping time is only greater than t_0 if all of the GLR statistics produced up to time t_0 by each stream are less than λ. The next step follows since $N_{t_0}^{(m)} \leq t_0, \forall m \in [M]$, 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, it implies 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 flagging a detection after t_0 samples. The next line of (24) follows since the events are independent. The dependence between each of the streams' statistics is present due to the correlation between the values of $N_t^{(m)}$ given a fixed time t. Otherwise the streams have no impact on each other. We removed the dependence on t and $N_t^{(m)}$ and considered the statistics after each stream has been observed a fixed t_0 times. This can be thought of as calculating M independent statistics from M independent samples of size t_0. The last line of (24) 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

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

is equal to the probability of the event $\{\tau > t_0\}$ when $M = 1$ and $\nu = \infty$. Inserting this into (24), the following relationship holds between the distributions of τ in the multi-stream and single-stream no-change settings:

$$\mathbb{P}_{M,\infty}(\tau > t_0) \geq \mathbb{P}_{1,\infty}(\tau > t_0)^M. \tag{25}$$

From Theorem 1 of [9], τ is asymptotically exponentially distributed with expected value

$$\mathbb{E}_{1,\infty}[\tau] = \frac{e^\lambda\sqrt{\pi}}{\sqrt{\lambda}\int_0^\infty xg(x)^2dx}, \tag{26}$$

as $\lambda \to \infty$ when $M = 1$ and $\nu = \infty$. Applying the formula for the expected value of a non-negative random variable:

$$\begin{aligned}
\mathbb{E}_{M,\infty}[\tau] &= \sum_{t_0=0}^{\infty} \mathbb{P}_{M,\infty}(\tau > t_0) \geq \sum_{t_0=0}^{\infty} \mathbb{P}_{1,\infty}(\tau > t_0)^M \\
&\sim \int_0^\infty \exp(-Mt_0/\mathbb{E}_{1,\infty}[\tau])dt_0 \\
&= \frac{\mathbb{E}_{1,\infty}[\tau]}{M} = \frac{e^\lambda\sqrt{\pi}}{M\sqrt{\lambda}\int_0^\infty xg(x)^2dx},
\end{aligned} \tag{27}$$

as $\lambda \to \infty$.

$\square$