QUICKEST DETECTION OF COMPOSITE AND NON-STATIONARY CHANGES WITH APPLICATION TO PANDEMIC MONITORING

Yuchen Liang and Venugopal V. Veeravalli

University of Illinois at Urbana-Champaign
ECE Department and Coordinated Science Lab
Urbana, IL

ABSTRACT

The problem of quickest detection of a change in the distribution of a sequence of independent observations is considered. The pre-change distribution is assumed to be known and stationary, while the post-change distributions are assumed to evolve in a pre-determined non-stationary manner with some possible parametric uncertainty. In particular, it is assumed that the cumulative KL divergence between the post-change and the pre-change distributions grows super-linearly with time after the change-point. For the case where the post-change distributions are known, a universal asymptotic lower bound on the delay is derived, as the false alarm rate goes to zero. Furthermore, a window-limited CuSum test is developed, and shown to be asymptotically optimal. For the case where the post-change distributions have parametric uncertainty, a window-limited generalized likelihood-ratio test is developed and is shown to be asymptotically optimal. The analysis is validated through numerical results on synthetic data. The use of the window-limited generalized likelihood-ratio test in monitoring pandemics is also demonstrated.

Index Terms— Quickest change detection (QCD), non-stationary observations, generalized likelihood-ratio (GLR) test, pandemic monitoring, window-limited sequential test.

1. INTRODUCTION

The problem of quickest change detection (QCD) is of fundamental importance in mathematical statistics (see, for example, [1, 2] for an overview). Given a sequence of observations whose distribution changes at some unknown change-point, the goal is to detect the change in distribution as quickly as possible after it occurs, while not making too many false alarms.

In the classical formulations of the QCD problem, it is assumed that the pre- and post-change distributions are known and stationary, and that the pre-change distribution is independent and identically distributed (i.i.d.). In many practical situations, while it is reasonable to assume that we can accurately estimate the pre-change distribution, the post-change distribution is rarely completely known. Furthermore, while it is reasonable to assume that the system is in steady-state before the change-point and producing stationary observations, the post-change distribution may typically be non-stationary. For example, in the pandemic monitoring problem, the distribution of the number of people infected daily might have achieved a steady-state before the start of a new wave. At the onset of the new wave, the post-change distribution is constantly evolving (see Section 4).

In this paper, our main focus is the QCD problem with independent observations, where the pre-change distribution is assumed to be known and stationary, while the post-change distribution is allowed to be non-stationary and have some parametric uncertainty.

There has been prior work on extensions of the classical QCD framework to the case where the pre- and/or the post-change distributions are not stationary. One approach is based on a minimax robust [3] formulation of the QCD problem, where it is assumed that the pre- and post-change distributions come from mutually exclusive uncertainty classes. Under certain conditions, e.g., joint stochastic boundedness [4] and weak stochastic boundedness [5], low-complexity tests that either coincide with [6] or asymptotically approach [7] the optimal test can be found. There is also a body of work on the problem of detecting transient changes (see, e.g., [8]), and persistent changes with some transient dynamics [9].

There have also been extensions of the classical formulation to the case where the pre- and/or post-change distributions are not fully known. In the generalized likelihood ratio (GLR) approach, introduced in [10], it is assumed that the pre- and post-change distributions are i.i.d. and come from one-parameter exponential families, respectively, and the post-change parameter is unknown. The GLR approach is studied in detail for the problem of detecting the change in the mean of a Gaussian distribution with unknown post-change mean in [11]. Both the mixture [12] and the GLR approaches are studied in detail for the case where pre- and post-change distributions are non-i.i.d. and the post-change distribution has parametric uncertainty in [13], where it is assumed that the cumulative Kullback-Leibler (KL) divergence between the post-change and the pre-change distributions grows linearly in the number of observations.

In some application (e.g., the pandemic monitoring problem), the post-change distributions are non-stationary in a way such that the cumulative KL divergence grows super-linearly after the change-point, in which case we say that the post-change distribution is *detection-favorable*. This is the setting we consider in this paper. Our contributions are as follows:

1. We extend the universal lower bound on the worst-case delay given in [13] to the more general detection-favorable setting.
2. We develop a window-limited CuSum test that asymptotically achieves the lower bound on the delay when the post-change distribution is detection-favorable and fully known.
3. We develop and analyze a GLR test that asymptotically achieves the worst-case delay when the post-change distributions are detection-favorable and have parametric uncertainty.

This work was supported in part by the National Science Foundation under grant ECCS-2033900, and by the Army Research Laboratory under Cooperative Agreement W911NF-17-2-0196, through the University of Illinois at Urbana-Champaign.

4. We validate our analysis through numerical results, and demonstrate the use of our approach in monitoring pandemics.

Detailed proofs of all of the theoretical results are given in an extended version of this paper [14]. Generalizations to the case with dependent observations and change-point dependent post-change distributions are also discussed in [14].

2. INFORMATION BOUNDS AND OPTIMAL DETECTION

Let $X_1, \ldots, X_n, \ldots$ be a sequence of independent random variables, and let ν be a change-point. Assume that $X_1, \ldots, X_{\nu-1}$ all have density p_0 with respect to some measure μ. Furthermore, assume that $X_\nu, X_{\nu+1}, \ldots$ have densities $p_{1,0}, p_{1,1}, \ldots$, respectively, with respect to μ, i.e., we are implicitly assuming that the post-change distribution is *time-invariant* with respect to the change-point ν. Note that the distributions of the observations are allowed to be non-stationary after the change-point. Let $\mathbb{P}_\nu$ denote the probability measure on the entire sequence of observations, when the change-point is ν, and let $\mathbb{E}_\nu[\cdot]$ denote the corresponding expectation.

The change-time ν is assumed to be unknown but deterministic. Let τ be a stopping time [4] defined on the observation sequence associated with the detection rule, i.e. τ is the time at which we stop taking observations and declare that the change has occurred.

2.1. Information Bounds for Non-stationary Post-Change Distributions

Lorden [10] proposed solving the following optimization problem to find the best stopping time τ:

$$\inf_{\tau \in \mathcal{C}_\alpha} \mathrm{WADD}(\tau) \tag{1}$$

where

$$\mathrm{WADD}(\tau) := \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathbb{E}_\nu \left[(\tau - \nu + 1)^+ | \mathcal{F}_{\nu-1} \right] \tag{2}$$

characterizes the worst-case delay, $\mathcal{F}_n$ denotes the sigma algebra generated by $X_1, \ldots, X_n$, i.e., $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, and

$$\mathcal{C}_\alpha := \{\tau : \mathrm{FAR}(\tau) \leq \alpha\} \tag{3}$$

with

$$\mathrm{FAR}(\tau) := \frac{1}{\mathbb{E}_\infty[\tau]}.$$

Here, $\mathbb{E}_\infty[\cdot]$ is the expectation operator when the change never happens, and $(\cdot)^+ := \max\{0, \cdot\}$.

In the classical i.i.d. model where the post-change distribution is stationary, the cumulative KL-divergence after the change-point increases linearly in the number of observations. We generalize this condition as follows. Let the growth function g represent the cumulative Kullback-Leibler (KL) divergence under the true distribution. More specifically, let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be increasing and continuous. Note that the inverse of g, denoted by g^{-1}, exists and is also increasing and continuous. It is assumed that the expected sum of the log-likelihood ratios under $\mathbb{P}_\nu$ matches the value of the growth function at all positive integers, i.e.,

$$g(n) = \sum_{i=\nu}^{\nu+n} \mathbb{E}_\nu [Z_{i,\nu}], \forall n \geq 1. \tag{4}$$

Here,

$$Z_{n,k} = \ln \frac{p_{1,n-k}(X_n)}{p_0(X_n)} \tag{5}$$

is the log-likelihood ratio where $n \geq k \geq 1$. Note that the KL divergence is always positive, i.e.,

$$\mathbb{E}_\nu [Z_{i,\nu}] > 0, \forall i \geq \nu.$$

In this paper, we are interested in the case where the post-change distribution is eventually persistently different from the pre-change. Specifically, it is assumed that

$$\exists I > 0,\ s.t.\ g(x) \geq Ix,\ \forall x \text{ sufficiently large.}$$

Lemma 2.1. *Consider the growth function $g(n)$ defined in* (4)*. Suppose that the sum of variance of the log-likelihood ratios satisfies*

$$\sup_{t \geq \nu \geq 1} \sum_{i=t}^{t+n} \mathrm{Var}_\nu (Z_{i,t}) = o(g^2(n)) \tag{6}$$

where $f(n) = o(g^2(n))$ is equivalent to $f(n)/g^2(n) \to 0$ as $n \to \infty$. Further, suppose that

$$\mathbb{E}_\nu [Z_{i,\nu}] \leq \mathbb{E}_\nu [Z_{i+\Delta,\nu+\Delta}] \tag{7}$$

for all positive integers Δ. Then,

$$\sup_{\nu \geq 1} \mathbb{P}_\nu \left\{ \max_{t \leq n} \sum_{i=\nu}^{\nu+t} Z_{i,\nu} \geq (1+\delta) g(n) \right\} \xrightarrow{n \to \infty} 0 \tag{8}$$

and

$$\sup_{t \geq \nu \geq 1} \mathbb{P}_\nu \left\{ \sum_{i=t}^{t+n} Z_{i,t} \leq (1-\delta) g(n) \right\} \xrightarrow{n \to \infty} 0 \tag{9}$$

for any $\delta \in (0,1)$.

Example 2.1. Consider the Gaussian exponential mean-change detection problem as follows. Denote by $\mathcal{N}(\mu_0, \sigma_0^2)$ the Gaussian distribution with mean μ_0 and variance σ_0^2. Let $X_1, \ldots, X_{\nu-1}$ be distributed as $\mathcal{N}(\mu_0, \sigma_0^2)$, and for all $n \geq \nu$, let X_n be distributed as $\mathcal{N}(\mu_0 e^{c(n-\nu)}, \sigma_0^2)$. Here c is some positive fixed constant. The log-likelihood ratio is given by:

$$Z_{n,t} = \ln \frac{p_{1,n-t}(X_n)}{p_0(X_n)} = \frac{\mu_0}{\sigma_0^2}(e^{c(n-t)} - 1) X_n - \frac{\mu_0^2 (e^{2c(n-t)} - 1)}{2\sigma_0^2}. \tag{10}$$

Now, the growth function can be calculated as

$$g(n) = \sum_{i=\nu}^{\nu+n} \mathbb{E}_\nu [Z_{i,\nu}] = \sum_{i=0}^{n} \frac{\mu_0^2}{2\sigma_0^2} (e^{ci} - 1)^2. \tag{11}$$

Conditions (6) and (7) are checked in [14]. □

The following theorem gives an asymptotic lower bound on the worst-case delay as $\alpha \to 0$.

Theorem 2.2. *Suppose that* (8) *holds. Then,*

$$\inf_{\tau \in \mathcal{C}_\alpha} \mathrm{WADD}(\tau) \geq g^{-1}(|\ln \alpha|)(1 + o(1)) \tag{12}$$

where $o(1) \to 0$ as $\alpha \to 0$.

2.2. Asymptotically Optimal Detection with Non-stationary Post-Change Distributions

Under the classical setting, Page's CuSum test is optimal [15] and has the following structure:

$$\Lambda(n) = \max_{1 \le k \le n+1} \sum_{i=k}^{n} Z_i = (\Lambda(n-1) + Z_n)^+$$
$$\tau_{\text{Page}}(b) = \inf \{n : \Lambda(n) \ge b\}. \quad (13)$$

In the above,

$$Z_n = \ln \frac{p_1(X_n)}{p_0(X_n)} \quad (14)$$

is the log-likelihood ratio when the post-change distributions are stationary. When the post-change distributions are potentially non-stationary, we modify the CuSum stopping rule as:

$$\tau_C(b) := \inf \left\{ n : \max_{1 \le k \le n+1} \sum_{i=k}^{n} Z_{i,k} \ge b \right\} \quad (15)$$

where $Z_{i,k}$ represents the log-likelihood ratio between densities $p_{1,i-k}$ and p_0 for observation X_i (defined in (5)).

As shown in (13), Page's classical CuSum algorithm admits a recursive way to compute its test statistic. Unfortunately, despite independent observations, the test statistic in (15) cannot be computed recursively. For computational tractability, we therefore consider a window-limited version of the test in (15):

$$\tilde{\tau}_C(b) := \inf \left\{ n : \max_{n-m_\alpha \le k \le n+1} \sum_{i=k}^{n} Z_{i,k} \ge b \right\} \quad (16)$$

where m_α is the window size. Throughout this paper, we require that m_α satisfy the following conditions:

$$\liminf m_\alpha / g^{-1}(|\ln \alpha|) > 1 \quad \text{and} \quad \ln m_\alpha = o(|\ln \alpha|). \quad (17)$$

Since the range for the maximum is smaller in $\tilde{\tau}_C(b)$ than in $\tau_C(b)$, given any realization of $X_1, X_2, \ldots$, if the test statistic of $\tilde{\tau}_C(b)$ crosses the threshold b at some time n, so does that of $\tau_C(b)$. Therefore, for any fixed threshold $b > 0$,

$$\tau_C(b) \le \tilde{\tau}_C(b) \quad a.s. \quad (18)$$

In the following, we first control the asymptotic false alarm rate of $\tilde{\tau}_C(b)$ with an appropriately chosen threshold in Lemma 2.3. Then, we upper bound the asymptotic delay of $\tilde{\tau}_C(b)$ in Lemma 2.4. Finally, we combine these two lemmas and provide an asymptotically optimal solution to the problem in (1) in Theorem 2.5.

Lemma 2.3. *Suppose that* $b_\alpha = |\ln \alpha| + \ln(2m_\alpha)$. *Then,*

$$\mathbb{E}_\infty [\tilde{\tau}_C(b_\alpha)] \ge \alpha^{-1}(1 + o(1))$$

where $o(1) \to 0$ *as* $\alpha \to 0$.

Remark. If $\ln m_\alpha = o(|\ln \alpha|)$, then $b_\alpha = |\ln \alpha| (1 + o(1))$.

Lemma 2.4. *Suppose that* $b_\alpha = |\ln \alpha| (1 + o(1))$. *Further, suppose that* (9) *holds for* $Z_{n,k}$ *when* $n \ge k$. *Then,*

$$\text{WADD}(\tilde{\tau}_C(b_\alpha)) \le g^{-1}(|\ln \alpha|)(1 + o(1))$$

where $o(1) \to 0$ *as* $\alpha \to 0$.

Theorem 2.5. *Let* $b_\alpha = |\ln \alpha| + \ln(2m_\alpha)$. *Suppose that* (8) *and* (9) *hold for* $Z_{n,k}, \forall n \ge k$. *Then, the stopping rule in* (16) *solves the problem in* (1) *asymptotically as* $\alpha \to 0$, *and*

$$\text{WADD}(\tilde{\tau}_C(b_\alpha)) = g^{-1}(|\ln \alpha|)(1 + o(1)) \quad (19)$$

where $o(1) \to 0$ *as* $\alpha \to 0$.

Example 2.2. Consider the same setting as in Example 2.1. We have shown that conditions (6) and (7) hold, and thus (8) and (9) follow by Lemma 2.1. Considering $g(n)$ in (11) as $n \to \infty$, we obtain

$$\text{WADD}(\tilde{\tau}_C(b_\alpha)) = \frac{1}{2c} \ln \left(\frac{2\sigma_0^2(1 - e^{-2c})}{\mu_0^2} |\ln \alpha| \right) (1 + o(1))$$
$$= O\left(\frac{1}{2c} \ln(|\ln \alpha|) \right) \quad (20)$$

where $o(1) \to 0$ as $\alpha \to 0$, and b_α is as defined in Theorem 2.5.

3. WINDOW-LIMITED GLR WITH UNKNOWN PARAMETERS

We now study the case where the evolution of the post-change distribution is parametrized by $\theta \in \mathbb{R}^d$. Let $X_\nu, X_{\nu+1}, \ldots$ be distributed as $\mathbb{P}_\nu^{\theta_1}$, where the corresponding densities are $p_{1,0}^{\theta_1}, p_{1,1}^{\theta_1}, \ldots$ with respect to the common measure μ. Let $\Theta \subset \mathbb{R}^d$ be the parameter set and $\theta_1 \in \Theta$. Note that Θ does not need to be compact. The true post-change parameter θ_1 is assumed to be unknown but deterministic. Let the log-likelihood ratio be re-defined as

$$Z_{n,k}^\theta = \ln \frac{p_{1,n-k}^\theta(X_n)}{p_0(X_n)} \quad (21)$$

for any $n \ge k$ and $\theta \in \Theta$. Here X_n is drawn from the distribution with true change-point ν and true post-change parameter θ_1. The problem is to solve (1) asymptotically as $\alpha \to 0$ under parameter uncertainty.

Consider the following window-limited GLR stopping:

$$\tilde{\tau}_G(b) := \inf \left\{ n : \max_{n-m_\alpha \le k \le n+1} \sup_{\theta \in \Theta_\alpha} \sum_{i=k}^{n} Z_{i,k}^\theta \ge b \right\} \quad (22)$$

where $\Theta_\alpha \nearrow \Theta$ as $\alpha \searrow 0$. Therefore, it is guaranteed that $\theta_1 \in \Theta_\alpha$ for all small enough α. Further, let $\Theta_\alpha \subset \mathbb{R}^d$ be compact for each α, and thus the maximizing θ given the pair (k, n) at the false-alarm rate α, denoted by $\hat{\theta}_{n,k}$, is contained in Θ_α.

If Θ_α is discrete-valued, the $\sup$ in (22) becomes $\max$, and the stopping time is equivalent to running $|\Theta_\alpha|$ CuSum algorithms simultaneously, where $\tilde{\tau}_G$ stops whenever one of the algorithms stops. Therefore, we only consider the case where Θ_α is continuous.

Finally, it is assumed that the largest absolute eigenvalue of the Hessian matrix of $Z_{n,k}^\theta$ exists and is finite in the neighborhood of $\hat{\theta}_{n,k}$ when the false alarm rate is small. Specifically, there exists $\epsilon > 0$ such that for any $\hat{\theta} \in \Theta$ and any large enough $b > 0$,

$$\sup_{\theta : \|\theta - \hat{\theta}\| < b^{-\frac{\epsilon}{2}}} \lambda_{\max} \left(-\nabla_\theta^2 \sum_{i=k}^{n} Z_{i,k}^\theta \right) \le 2b^\epsilon \quad (23)$$

where $\lambda_{\max}(A)$ represents the maximum eigenvalue of a matrix A.

In the following, we first upper bound the asymptotic delay of $\tilde{\tau}_G(b)$ in Lemma 3.1. Next, we control the asymptotic false alarm rate of $\tilde{\tau}_G(b)$ with some proper threshold in Lemma 3.2. Finally, we combine these two lemmas and establish asymptotic optimality in Theorem 3.3.

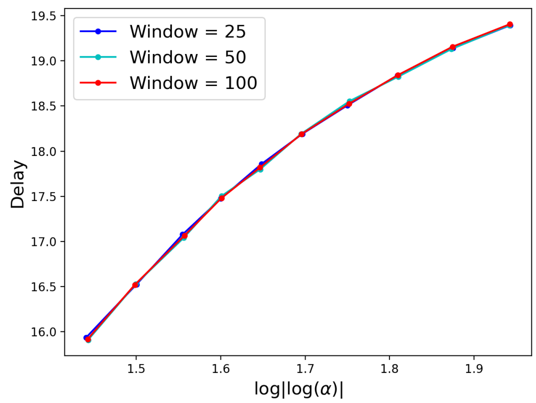

Fig. 1: Performances of tests with different sizes of windows. The Gaussian exponential mean-change problem is considered, with $\mu_0 = 0.1, \sigma_0^2 = 10000$, and $c = 0.4$. The change-point $\nu = 1$.

Lemma 3.1. *If $\theta_1 \in \Theta_\alpha$, then for any threshold $b > 0$,*

$$\text{WADD}\left(\tilde{\tau}_G(b)\right) \leq \text{WADD}\left(\tilde{\tau}_C(b)\right).$$

Corollary 3.1.1. *Suppose that $b_\alpha = |\ln \alpha|\,(1 + o(1))$. Since $\theta_1 \in \Theta_\alpha$ for all small α's, the worst-case delay of $\tilde{\tau}_G(b)$ satisfies*

$$\text{WADD}\left(\tilde{\tau}_G(b_\alpha)\right) \leq g^{-1}(|\ln \alpha|)(1 + o(1))$$

asymptotically as $\alpha \to 0$.

Lemma 3.2. *Suppose that b_α satisfies*

$$2 m_\alpha C_d^{-1} b_\alpha^{\frac{\epsilon d}{2}} e^{1 - b_\alpha} = \alpha \tag{24}$$

where C_d is a constant that depends only on d. Then,

$$\mathbb{E}_\infty\left[\tilde{\tau}_G(b_\alpha)\right] \geq \alpha^{-1}(1 + o(1))$$

where $o(1) \to 0$ as $\alpha \to 0$.

Remark. Re-arranging the terms, (24) becomes:

$$b_\alpha = |\ln \alpha| + \ln(2 m_\alpha C_d^{-1} e) + \frac{\epsilon d}{2} \ln b_\alpha. \tag{25}$$

Since $\ln m_\alpha = o(|\ln \alpha|)$, $b_\alpha = |\ln \alpha|\,(1 + o(1))$ as $\alpha \to 0$.

Theorem 3.3. *Suppose that b_α is as defined in Lemma 3.2 and that m_α satisfies (17). Further, suppose that (8), (9) and (23) hold for $Z_{n,k}^\theta$ when $n \geq k$. Then, $\tilde{\tau}_G(b_\alpha)$ solves the problem in (1) asymptotically as $\alpha \to 0$, and*

$$\text{WADD}\left(\tilde{\tau}_G(b_\alpha)\right) = g^{-1}(|\ln \alpha|)(1 + o(1)) \tag{26}$$

where $o(1) \to 0$ as $\alpha \to 0$.

4. NUMERICAL RESULTS AND DISCUSSION

In Fig. 1, we study the performance of the proposed tests through simulations for the Gaussian exponential mean-change problem (see Example 2.1). It is observed that the delay at $\nu = 1$ is $O(\ln(|\ln \alpha|))$ for all sizes of windows considered, as described in (20).

Next, we apply our GLR algorithm to monitoring the spread of COVID-19 using new case data from various counties in the US [16]. The goal is to detect the onset of a new wave of the pandemic based on the incremental daily cases. The problem is modeled as one of detecting a change in the mean of a Beta distribution. The Beta distribution model is used because the daily incremental fraction is bounded between 0 and 1; models such as Gaussian, with unbounded support may not be appropriate. Let $\mathcal{B}(a, b)$ denote the Beta distribution with shape parameters a and b. Let

$$\begin{aligned} p_0 &= \mathcal{B}(a_0, b_0) \\ p_{1,n-k} &= \mathcal{B}(a_0 h(n-k), b_0), \forall n \geq k \end{aligned} \tag{27}$$

Here, h is a parametric function such that $h \geq 1$. Note that if $a_0 \ll b_0$ and $h(n - \nu)$ is not too large,

$$\mathbb{E}_\nu\left[X_n\right] = \frac{a_0 h(n-\nu)}{a_0 h(n-\nu) + b_0} \approx \frac{a_0}{b_0} h(n-\nu) \tag{28}$$

for all $n \geq \nu$. Therefore, h is designed to capture the shape of the average fraction of daily incremental cases. Let

$$h(\Delta) = 1 + \frac{10^{c_0}}{c_2} \exp\left(-\frac{(\Delta - c_1)^2}{2c_2^2}\right) \tag{29}$$

where $c_0, c_1, c_2 \geq 0$ are all parameters. This specific choice of h has two advantages: 1) It guarantees a rapid growth during the start of a new epidemic wave. When $n - \nu$ is small, $h(n - \nu)$ grows like the left edge of a Gaussian density if c_1 is large. 2) It guarantees that daily incremental cases will eventually vanish at the end of the current epidemic wave, i.e., $h(n - \nu) \to 0$ as $n \to \infty$.

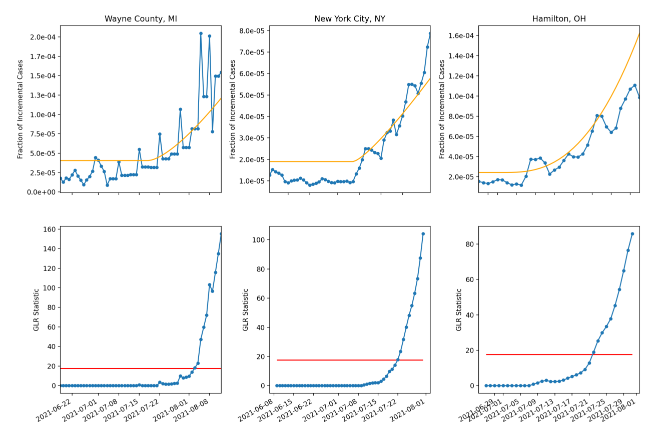

Fig. 2: COVID-19 monitoring example. The upper row shows the four-day moving average of the daily new cases of COVID-19 as a fraction of the population in Wayne County, MI (left), New York City, NY (middle), and Hamilton County, OH (right). A pre-change $\mathcal{B}(a_0, b_0)$ distribution is estimated using data from the previous 20 days (from May 26, 2021 to June 14, 2021). The mean of the Beta distributions with the hypothesized change-point and estimated parameters from the GLR algorithm is also shown (in orange). The lower row shows the evolution of the GLR test statistic (defined in (22)), respectively. The FAR threshold α is set to 0.01, and the corresponding GLR test threshold is also shown (in red). The post-change distribution at time n with hypothesized change-point k is modeled as $\mathcal{B}(a_0 h(n-k), b_0)$, where h is defined in (29). The parameters c_0, c_1 and c_2 are assumed to be unknown. The window size $m_\alpha = 20$. The threshold is set using equation (24).

In Fig. 2, we illustrate the use our GLR algorithm with the distribution model in (27) in the detection of the onset of a new wave of COVID-19. We assumed a start date of June 15th, 2021 for the monitoring, at which time the pandemic appeared to be in a steady state with incremental cases staying relatively flat. We observe that the GLR statistic significantly and persistently crosses the test-threshold around late July in all counties, which is strong indication of a new wave of the pandemic. More importantly, unlike the raw observations which are highly varying, the GLR statistic shows a clear dichotomy between the pre- and post-change settings, with the statistic staying near zero before the purported onset of the new wave, and taking off nearly vertically after the onset.

5. REFERENCES

[1] V. V. Veeravalli and T. Banerjee, "Quickest change detection," in *Academic press library in signal processing: Array and statistical signal processing*. Cambridge, MA: Academic Press, 2013.

[2] L. Xie, S. Zou, Y. Xie, and V. V. Veeravalli, "Sequential (quickest) change detection: Classical results and new directions," *arXiv preprint arXiv:2104.04186*, 2021.

[3] P. J. Huber, "A robust version of the probability ratio test," *The Annals of Mathematical Statistics*, vol. 36, no. 6, pp. 1753–1758, Dec. 1965.

[4] P. Moulin and V. V. Veeravalli, *Statistical Inference for Engineers and Data Scientists*. Cambridge, UK: Cambridge University Press, 2018.

[5] T. L. Molloy and J. J. Ford, "Misspecified and asymptotically minimax robust quickest change detection," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5730–5742, 2017.

[6] T. L. Molloy and J. J. Ford, "Minimax robust quickest change detection in systems and signals with unknown transients," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2976–2982, July 2019.

[7] Y. Liang and V. V. Veeravalli, "Non-parametric quickest detection of a change in the mean of an observation sequence," in *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, 2021, pp. 1–6.

[8] B. K. Guépié, L. Fillatre, and I. Nikiforov, "Sequential detection of transient changes," *Sequential Analysis*, vol. 31, no. 4, pp. 528–547, 2012. [Online]. Available: https://doi.org/10.1080/07474946.2012.719443

[9] S. Zou, G. Fellouris, and V. V. Veeravalli, "Quickest change detection under transient dynamics: Theory and asymptotic analysis," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1397–1412, 2019.

[10] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1897–1908, Dec. 1971.

[11] D. Siegmund and E. S. Venkatraman, "Using the generalized likelihood ratio statistic for sequential detection of a change-point," *The Annals of Statistics*, vol. 23, no. 1, pp. 255–271, Feb. 1995.

[12] M. Pollak, "Optimality and almost optimality of mixture stopping rules," *Annals of Statistics*, vol. 6, no. 4, pp. 910–916, Jul. 1978.

[13] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, November 1998.

[14] Y. Liang and V. V. Veeravalli, "Quickest change detection with non-stationary and composite post-change distribution," *arXiv preprint arXiv:2110.01581*, 2021.

[15] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Annals of Statistics*, vol. 14, no. 4, pp. 1379–1387, Dec. 1986.

[16] N. Y. Times. Coronavirus in the U.S.: Latest Map and Case Count. [Online]. Available: https://www.nytimes.com/interactive/2021/us/covid-cases.html