Asynchronous Multi-Sensor Change-Point Detection for Seismic Tremors

Liyan Xie*, Yao Xie*, George V. Moustakides †‡

*School of Industrial & Systems Engineering, Georgia Institute of Technology, {lxie49, yao.xie}@isye.gatech.edu

†Department of Computer Science, Rutgers University, gm463@rutgers.edu

‡Electrical and Computer Engineering department, University of Patras, moustaki@upatras.gr

***Abstract*—We consider the sequential change-point detection for asynchronous multi-sensors, where each sensor observe a signal (due to change-point) at different times. We propose an asynchronous Subspace-CUSUM procedure based on jointly estimating the unknown signal waveform and the unknown relative delays between sensors. Using the estimated delays, we can align signals and use the subspace to combine multiple sensor observations. We derive the optimal drift parameter for the proposed procedure, and characterize the relationship between the expected detection delay, average run length (of false alarms), and the energy of the time-varying signal. We demonstrate the good performance of the proposed procedure using simulation and real data. We also demonstrate that the proposed procedure outperforms the well-known "one-shot procedure" in detecting weak and asynchronous signals.**

I. Introduction

We consider detecting the emergence of a signal, which is observed at multiple sensors with *unknown and different delays and amplitudes*. Such problem arises frequently in sensor network monitoring, where the sensors observe the sudden occurrence of a signal at different times mostly due to propagation delays, as illustrated in Fig. 1. The main application of interest is seismic tremors detection [1]. Tremors are low amplitude ambient vibrations of the ground caused by man-made or atmospheric disturbances; detecting the underlying tremors will enable geophysicists to build better predictive models. Our goal is to detect the emergence of such occurrence (change) by combining observations at multiple sensors.

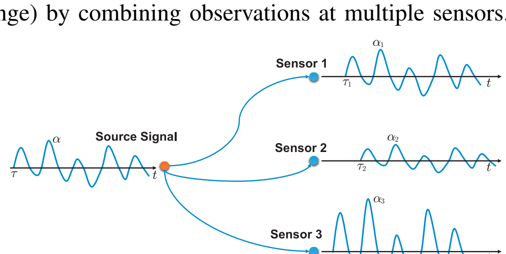


Fig. 1. Illustration of a sensor network with a signal.

Classical sequential change-point detection [7] usually assumes the change happens simultaneously at all or a subset of sensors. The scenario we are proposing above, calls for the development of new methods that can consider delay estimation together with change-point detection. Change-point detection with delays has also been considered in [4] using the so-called one-shot scheme, where the fusion center declares an alarm whenever one sensor detects a change. However, the one-shot scheme relies only on local sensor information, it does not take advantage of the multi-sensor observations and the fact that the change occurs in all sensors but not at the same time. Combining asynchronous sensor observations can be beneficial for the detection problem when the relative delays are small and when the signal is weak, as is the case of seismic tremors. This is because combining observations effectively boosts the SNR.

Even though in our work the time of change in sensors appears to be different, there exists a *deterministic difference* between change-times due to the location geometry of sensors and the location where the tremor occurs. Consequently, our approach will be similar to approaches where there is a simultaneous change in all sensors after, of course, compensating for the fixed but unknown delays.

We propose an asynchronous Subspace-CUSUM procedure, based on jointly estimating the unknown signal waveform and the unknown relative delays. It is related to the Subspace-CUSUM procedure in our prior work [2], [11]. We extend the results therein for the asynchronous case, and develop an optimal choice of the drift parameter, which is essential for CUSUM type of procedures. Our theoretical analysis reveals insights into the relationship between the average energy of the time-varying signal and the expected detection delay. This may potentially allow us to prove the asymptotic optimality of the asynchronous Subspace-CUSUM, by extending the arguments in [2]. We demonstrate the good performance of our procedure using simulated examples and real seismic data. Our procedure outperforms the one-shot scheme especially when the signal amplitude is weak and the relative delays are not too large.

The rest of the paper is organized as follows. In section II, we introduce the background of sequential subspace detection and briefly summarize the Subspace-CUSUM procedure. In section III, we first consider the case where the delays are perfectly known and form the detection statistic. Next, we propose a detector that combines sensor synchronization with detection when the delays are unknown. The theoretical result about how to properly select the drift parameter is discussed in section IV. In section V we evaluate our method by applying it to both simulated signals and real seismic data and observe its performance under strong and weak signals.

II. Background: Subspace-CUSUM procedure

We first introduce the Subspace-CUSUM procedure which will be the basis of our subsequent discussion. Consider the change-point detection problem where the covariance changes from an identity matrix $\sigma^2 I_k$ to a spiked matrix $\Sigma = \sigma^2 I_k + \theta u u^\intercal$, here k is the dimension, $\theta > 0$ is the signal strength, $u \in \mathbb{R}^{k\times 1}$ represents a basis for the subspace with $\|u\|_2 = 1$, and σ^2 is the noise power. We can define the SNR as $\rho = \theta/\sigma^2$. Assume the sequentially observed data are as follows

$$\begin{aligned} &x_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2 I_k), && t = 1, 2, \ldots, \tau, \\ &x_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2 I_k + \theta u u^\intercal), && t = \tau+1, \tau+2, \ldots \end{aligned} \tag{1}$$

where τ is the unknown change-point that we would like to detect as quickly as possible. We assume that the subspace u is unknown since it represents anomaly or new information.

The well-known cumulative sum (CUSUM) test [5], [6] cumulates the *log-likelihood ratio* and declares an alarm whenever the cumulation exceeds a threshold. For the data model in (1), the log-likelihood ratio for each sample is derived as the equation (7) in [11]:

$$\log \frac{f_0(x_t)}{f_\infty(x_t)} = \frac{1}{2\sigma^2}\frac{\rho}{1+\rho}\Big\{(u^\intercal x_t)^2 - \sigma^2\Big(1+\frac{1}{\rho}\Big)\log(1+\rho)\Big\},$$

where f_∞ and f_0 denote the probability density function before and after the change. Based on this, we can form the CUSUM statistic as

$$S_t = (S_{t-1})^+ + (u^\intercal x_t)^2 - \sigma^2\left(1+\frac{1}{\rho}\right)\log(1+\rho),$$

where $(x)^+ = \max\{x, 0\}$. When u is unknown, in the Subspace-CUSUM procedure, u is replaced with a sequential estimate $\hat{u}_t$ (will be explained later):

$$\mathcal{S}_t = (\mathcal{S}_{t-1})^+ + (\hat{u}_t^\intercal x_t)^2 - d. \tag{2}$$

Following the analogy in [3], we require the *drift* d to satisfy

$$\mathbb{E}_\infty[(\hat{u}_t^\intercal x_t)^2] < d < \mathbb{E}_0[(\hat{u}_t^\intercal x_t)^2], \tag{3}$$

where $\mathbb{E}_\infty$ and $\mathbb{E}_0$ denote the expectation under nominal and alternative measure respectively.

The Subspace-CUSUM procedure can be defined through the following stopping time

$$\mathcal{T} = \inf\{t > 0 : \mathcal{S}_t \geq b\}, \tag{4}$$

where b is a pre-specified threshold set to control the false alarm rate. This test is known to be asymptotically optimum for the stationary case (constant θ) [2].

To obtain the estimate $\hat{u}_t$, we form the sample covariance matrix using the observations $\{x_{t+1}, \ldots, x_{t+w}\}$ that lie in the future of t,

$$\Sigma_t = \sum_{j=t+1}^{t+w} x_j x_j^\intercal, \tag{5}$$

then the *unit-norm* singular vector corresponding to the largest singular value of Σ_t can be viewed as an estimator for u at time t. The usage of observations from the future is always possible by properly delaying the data. In particular, if we stop at time $\mathcal{T} = t$, this implies that we used data from times up to $t+w$ and, consequently, $t+w$ is the true time we stop and not t. The main advantage of this idea is that it provides the estimator $\hat{u}_t$ that is independent of x_t.

III. Problem Setup

In this section, we first consider the case when delays are known, and show how the Subspace-CUSUM procedure should be applied to detect the signal. When the delays are unknown, we develop a method that simultaneously synchronizes sensors by estimating their relative delays and detects the change.

Consider k sensors as in Fig. 1. Suppose we have sequential observations at each sensor, $x_1(t), x_2(t), \ldots, x_k(t)$. Before the emergence of the signal, the observations are noises that are assumed to be *i.i.d.* normal random variables. We will also assume that their powers are known which implies that, without loss of generality, we can assume that these powers are all equal to σ^2 since this is always possible with proper normalization. When a source signal $s(t)$ occurs, the observations at different sensors will capture it but with different delays and amplitudes. We assume the signal is causal, i.e., $s(t) = 0, \forall t < 0$. Denote the time point of onset of the signal in ith sensor as τ_i, and define the change-point as

$$\tau = \min_{1\leq i\leq k} \tau_i.$$

For the ith sensor, this means:

$$\begin{aligned} &x_i(t) = e_i(t), && t = 1, 2, \ldots, \tau, \\ &x_i(t) = \alpha_i s(t-\tau_i) + e_i(t), && t = \tau+1, \tau+2, \ldots, \end{aligned} \tag{6}$$

where $e_i(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ are random noises, and α_i is the *unknown* amplitude of the change at the ith sensor. Assume that the source signal $s(t)$ is *unknown*. We also define the relative delay between the ith and jth sensor as $\tau_{ij} = \tau_j - \tau_i$. Note that the delays with respect to the source signal are always nonnegative, but the relative delays can be either positive or negative.

A. Known Delays

To help build intuition, we first consider the ideal case where the relative delays τ_{ij} are known. Without loss of generality, we assume $\tau_{1i} \geq 0, \forall i$. If this is the case we can construct the time-shifted version of the observations

$$\tilde{x}_i(t) = x_i(t+\tau_{1i}), \quad i = 1, 2, \ldots, k, \tag{7}$$

and combine them to form a k-dimensional vector

$$\tilde{x}_t = \begin{bmatrix}\tilde{x}_1(t) & \tilde{x}_2(t) & \cdots & \tilde{x}_k(t)\end{bmatrix}^\intercal. \tag{8}$$

Note that when there is a signal, after the data transformation in (7), we can show that the vectorized observations can be written as

$$\tilde{x}_t = s(t)\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_k\end{bmatrix}^\intercal + e_t,\ e(t) \sim \mathcal{N}(0, \sigma^2 I_k). \tag{9}$$

Therefore the covariance structure of $\tilde{x}_t$ will undergo a similar change as in the model in (1), namely, be the identity matrix

$\sigma^2 I_k$ before the emergence of the signal $s(t)$, and become the spiked model $\sigma^2 I_k + \theta(t)uu^\intercal$ after. Here the subspace u, represents the unknown *normalized* post-change amplitudes

$$u = \frac{\alpha}{\|\alpha\|}, \tag{10}$$

where, for simplicity, we denote $\alpha = [\alpha_1, \ldots, \alpha_k]^\intercal$, and the *time-varying* signal strength $\theta(t)$ is given by

$$\theta(t) = s^2(t)\|\alpha\|^2.$$

Note that (9) differs from the previous model (1) in that $\theta(t)$ is now *time-varying*, since the signal $s(t)$ changes with time. However, the change-point detection is still to detect the transition of the covariance matrix from an identity to a spiked covariance matrix. So, we can still adapt the Subspace-CUSUM in section II to be used in this case. We should add that optimum detection schemes for time-varying models are difficult to derive. For existing results please refer to [7].

Since the normalized post-change amplitude vector u is unknown, we can estimate it by forming the sample covariance matrix Σ_t introduced in (5) but with x_t replaced by $\tilde{x}_t$ in (8). If we apply the singular value decomposition on Σ_t then the singular vector corresponding to the largest singular value provides the desired estimate $\hat{u}_t$. Then, we plug $\hat{u}_t$ back into the detection statistic (2) and obtain

$$\mathcal{S}_t = (\mathcal{S}_{t-1})^+ + (\hat{u}_t^\intercal \tilde{x}_t)^2 - d, \tag{11}$$

and the stopping time is defined as in (4). Of course, there still remains the question of selecting the proper d. We defer the discussion of this issue until Section IV.

B. Unknown Delays

In this section, we consider the case where delays are unknown. Generally, the exact delay is not possible to obtain beforehand since it depends on the location of the tremor epicenter. Therefore, we need to come up with a method that will achieve sensor synchronization in order to apply the Subspace-CUSUM procedure. In fact, this will be performed continuously and in a sequential manner in parallel with the change-detection task.

We select one sensor as reference, and attempt to synchronize all other sensors with respect to this sensor. Synchronization can be implemented based on the maximum likelihood approach to estimate the relative delay, on a sensor-by-sensor basis or simultaneously for all sensors. Without loss of generality, we regard the data of the first sensor $x_1(t)$ as the reference and compute the relative delays with respect to $x_1(t)$. Assume that we have available an upper bound $\tau_{\max}$ on the unknown relative delays, so they are restricted in the interval $[-\tau_{\max}, \tau_{\max}]$.

For the ith sensor, the log-likelihood function of the observations $\{x_i(t+1), \ldots, x_i(t+w)\}$ after change can be written as

$$\begin{aligned}
&\ell_{s,\tau_i}\big(x_i(t+1), \ldots, x_i(t+w)\big) = \\
&\quad \frac{\alpha_i}{\sigma^2} \sum_{j=t+1}^{t+w} x_i(j)s(j-\tau_i) - \frac{\alpha_i^2}{2\sigma^2} \sum_{j=t+1}^{t+w} s^2(j-\tau_i) \\
&\qquad - \frac{1}{2\sigma^2} \sum_{j=t+1}^{t+w} x_i^2(j) - \frac{w}{2}\log(2\pi\sigma^2).
\end{aligned}$$

Therefore for any given signal waveform $s(t)$, the maximum likelihood estimator of τ_i at ith sensor is given by

$$\hat{\tau}_i = \underset{-\tau_{\max} \le z \le \tau_{\max}}{\arg\max} \Big|\sum\nolimits_{j=t+1}^{t+w} x_i(j)s(j-z)\Big|. \tag{12}$$

Based on the maximum likelihood estimator (12), we propose Algorithm 1 which performs the joint estimation of signal waveform and relative delay iteratively. Once we obtain the

Algorithm 1 Joint estimate of signal waveform and delay

Input: δ, $n_{\max}$

Initialize: $n \leftarrow 1$; $\hat{s}^{(1)} \leftarrow x_1$; $\hat{\tau}_{1i}^{(0)} = \infty$, $\hat{\tau}_{1i}^{(1)} = 0, \forall i$
while $\max_{i\ge 2} |\hat{\tau}_{1i}^{(n)} - \hat{\tau}_{1i}^{(n-1)}| \ge \delta$ and $n \le n_{\max}$ **do**
  $n \leftarrow n+1$
  for $i = 2, \ldots, k$ **do**
    $\hat{\tau}_{1i}^{(n)} = \underset{-\tau_{\max} \le z \le \tau_{\max}}{\arg\max} |\sum_{j=t+1}^{t+w} x_i(j)\hat{s}^{(n-1)}(j-z)|$
  end for
  Form sample vector (8) using delay estimate $\hat{\tau}_{1i}^{(n)}$
  Find $\hat{u}$, the singular vector corresponding to the largest singular value of the sample covariance matrix (5)
  $\hat{s}^{(n)}(t) \leftarrow \sum_{i=1}^{k} \hat{u}_i x_i(t + \hat{\tau}_{1i}^{(n)})$
end while

estimates of the delays we can then use (7) with τ_{1i} replaced by $\hat{\tau}_{1i}$ and then apply the Subspace-CUSUM as described above with x_t replaced by $\tilde{x}_t$.

IV. Theoretical Analysis

Our previously introduced condition (3) for the drift parameter is necessary to guarantee that Subspace-CUSUM will behave similarly as the regular CUSUM. In other words, the increment will have a negative mean under the nominal model resulting in multiple restarts while under the alternative regime the increment will be positive on average leading the statistic $\mathcal{S}_t$ to exceed the threshold. The aforementioned property is crucial and it ensures that the detection delay is proportional to the threshold while the expected duration between false alarms is an exponential function of the threshold.

When the post-change statistical behavior is not stationary, (3) is no longer applicable and we need to consider time as an additional source of variability. Since in our problem $s(t)$ is time-varying, the expectation over $s(t)$ must be replaced with the average *over time*. Specifically we have

$$\lim_{W\to\infty}\frac{1}{W}\sum_{j=0}^{W-1}\mathbb{E}_\infty[(\hat{u}_{t+j}^\intercal x_{t+j})^2]<d$$
$$<\lim_{W\to\infty}\frac{1}{W}\sum_{j=0}^{W-1}\mathbb{E}_0[(\hat{u}_{t+j}^\intercal x_{t+j})^2]. \quad (13)$$

In other words we assume that (3) is valid *on average* over time. Let us now see how this translates in our specific problem. We first make the following assumption.

Assumption 1. *There exists a positive constant E_0 such that the following limit is valid*

$$\lim_{W\to\infty}\frac{1}{W}\sum_{i=1}^{W}s^2(i)=E_0>0. \quad (14)$$

Quantity E_0 denotes the average energy of the signal $s(t)$.

To compute the expectations of $(\hat{u}_t^\intercal\tilde{x}_t)^2$, especially under the alternative regime, it is necessary to be able to describe the statistical behavior of our estimate $\hat{u}_t$. We will assume that the window size w is sufficiently large so that Central Limit Theorem type approximations are possible. Explicit formulas are given in the following Lemma.

Lemma 1. *If Assumption 1 is true, and $e_i(t)\overset{\text{iid}}{\sim}\mathcal{N}(0,\sigma^2)$, we have that under the pre-change regime,*

$$\mathbb{E}_\infty\left[(\hat{u}_t^\intercal\tilde{x}_t)^2\right]=\sigma^2,$$

and under the post-change regime:

$$\mathbb{E}_0\left[(\hat{u}_t^\intercal\tilde{x}_t)^2\right]\;=\;\sigma^2\left[1+\frac{s^2(t)\rho}{E_0}\left(1-\frac{1+\rho}{w\rho^2}(k-1)\right)\right],$$

where

$$\rho=\frac{E_0}{\sigma^2}\|\alpha\|^2,$$

can be viewed as the average SNR over time.

Proof. The proof is given in the Appendix. □

Using Lemma 1 and (13) we can immediately deduce that the drift d must satisfy

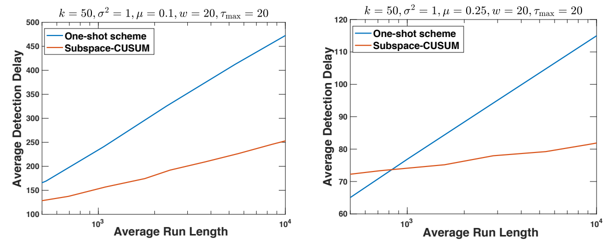
$$\sigma^2<d<\sigma^2\left[1+\rho\left(1-\frac{1+\rho}{w\rho^2}(k-1)\right)\right],$$

which is similar to the stationary condition imposed in [2] but with the average energy E_0 replacing the constant energy of the stationary version.

V. Numerical results

A. Simulation

In this section, we perform simulations to compare the performance of our subspace-based test with the one-shot scheme proposed in [4]. We adopt the same setting as in [4], where the distribution of the data stream at each sensor changes from $\mathcal{N}(0,\sigma^2)$ to $\mathcal{N}(\mu,\sigma^2)$ asynchronously. Indeed this is a special case of the model (6) by letting the signal $s(t)=1$, and the amplitude $\alpha_i=\mu$ for all sensors. In the numerical experiments, the noise level $\sigma^2=1$, the number of sensors $k=50$, the maximal relative delay $\tau_{\max}$ is set to 20, and the window size w is set to 20.



(a) (b)

Fig. 2. Comparison of Average detection delay as a function of Average run length for Subspace-CUSUM and One-shot detection scheme.

In Fig. 2 we compare the average detection delay of the one-shot scheme (blue line) and our subspace-based test (red line) as a function of the average run length, which is the average period between false alarms. Fig. 2 shows the results for (a) $\mu=0.1$ and (b) $\mu=0.25$, respectively. Our method can detect the weak signal with much smaller delays compared with the one-shot scheme. When μ is relatively large, the one-shot scheme outperform our method for small average run length values. However the proposed method performs better when the average run length increases. This suggests that combing multi-sensor observations can improve performance significantly.

B. Seismic Data

In this example, we consider the seismic tremor signal detection problem. The tremor signals are useful for geophysical study and prediction of potential earthquakes. Usually, the tremor signals are very weak to detect using data at any individual sensor; therefore, network detection methods have been developed which essentially use covariance information of the data [8]. This network-based detection problem can also be solved by our Subspace-CUSUM scheme discussed in III-B. Note that the tremor signal is abrupt but transient, therefore the detection statistic will decrease when the tremor disappears, and increase again when a new signal appears.

The seismic dataset we use is the records at Parkfield, California from 2am to 4am on 12/23/2004, which consist of 13 seismic sensors that simultaneously record a continuous stream of signals. The sampling frequency of the raw data is 250Hz. In the data preprocessing step, we normalize the observations at each sensor by subtracting the average value and dividing the maximal absolute value. The raw data after preprocessing is shown in Fig. 3. From the published catalog ([1]), we see three small earthquakes as shown in table I. There are also many low-frequency tremor records, mainly at time 2:34 $\sim$ 2:35, 2:42 $\sim$ 2:53, 3:24, 3:26, and 3:39.

[1]Northern California Earthquake Data Center: http://www.ncedc.org/ncedc/catalog-search.html

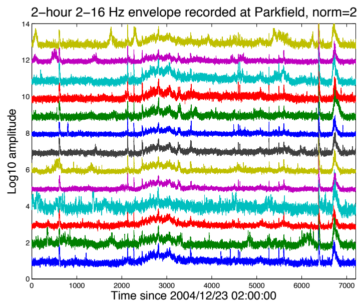


Fig. 3. Raw data from 13 sensors.

TABLE I
EARTHQUAKE CATALOG AT PARKFIELD DURING 2004/12/23 02-04 UT.

Date	Time	Lat	Lon	Mag	Event ID
2004/12/23	02:09:54.01	35.4593	-120.7500	1.47	21429343
2004/12/23	02:35:23.70	36.0368	-120.6088	1.10	30229299
2004/12/23	03:46:09.23	35.9290	-120.4797	1.47	21429365

In this example, we assume that the maximum delay $\tau_{\max} = 100$ (namely, 0.4 sec). The window size $w = 200$, which corresponds to 0.8 sec. We use the data within the first 500 sec (pre-change period) to set the drift parameter d numerically, which is 1.5 times the mean value of $(\hat{u}_t^\intercal \boldsymbol{x_t})^2$. We computed the Subspace-CUSUM statistic in (11), as shown in Fig. 4. It can be seen that the three main peaks are at 603.6 sec, 2127.0 sec, 6370.0 sec respectively, which is quite close to the true earthquake time (recall that these times 594.0 sec, 2123.7 sec, and 6369.2 sec). There are some small and continuous peaks within the time period 2500 sec $\sim$ 3200 sec, which match the tremor catalog of 2:42 $\sim$ 2:53.

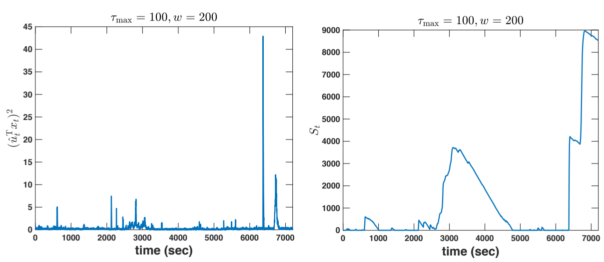


Fig. 4. Left: the increment $(\hat{u}_t^\intercal \tilde{x}_t)^2$. Right: the Subspace-CUSUM detection statistics $\mathcal{S}_t$ over time.

VI. CONCLUSION

We study the sequential asynchronous multi-sensor change-point detection. In the proposed asynchronous Subspace-CUSUM algorithm, observations at different sensors are synchronized and combined together to construct the detection statistic. We derive the optimal drift parameter for the proposed procedure, and characterize the relationship between the expected detection delay, average run length, and the energy of the time-varying signal. The good performance of the proposed procedure is presented using simulated signals and real seismic data. We also demonstrate that the proposed procedure outperforms the one-shot procedure in detecting weak and asynchronous signals.

ACKNOWLEDGMENT

The work of George Moustakides was supported by the US National Science Foundation under Grant CIF 1513373, through Rutgers University. The work of Liyan Xie and Yao Xie was supported by the US National Science Foundation under Grant CCF 1442635, CMMI 1538746, and a career award CCF 1650913. The authors would like to thank Dr. Zhigang Peng at Georgia Institute of Technology on his helpful suggestions for seismic data processing.

REFERENCES

[1] F.-C. Lin, M. P. Moschetti, and M. H. Ritzwoller, "Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps," *Geophysical Journal International*, vol. 173, no. 1, pp. 281–298, 2008.

[2] L. Xie, G. V. Moustakides, and Y. Xie, "First-order optimal sequential subspace change-point detection," *arXiv preprint arXiv:1806.10760*, 2018.

[3] T. Oskiper and H. V. Poor, "Quickest detection of a random signal in background noise using a sensor array," *EURASIP Journal on Advances in Signal Processing*, 360150, 2005.

[4] O. Hadjiliadis, H. Zhang, and H. V. Poor, "One shot schemes for decentralized quickest change detection," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3346–3359, 2009.

[5] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.

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

[7] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. Chapman and Hall/CRC, 2014.

[8] Z. Li, Z. Peng, D. Hollis, L. Zhu, and J. McClellan, "High-resolution seismic event detection using local similarity for large-n arrays," *Scientific Report*, vol. 8, no. 1646, 2018.

[9] T. W. Anderson, "Asymptotic theory for principal component analysis," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 122–148, 1963.

[10] D. Paul, "Asymptotics of sample eigenstructure for a large dimensional spiked covariance model," *Statistica Sinica*, pp. 1617–1642, 2007.

[11] L. Xie, Y. Xie, and G. V. Moustakides, "Sequential subspace changepoint detection," *arXiv preprint arXiv:1811.03936*, 2018.

APPENDIX

Proof of Lemma 1. We follow the proof of Lemma 1 in [2], the only difference being that the SNR is now time varying. Based on Assumption 1, as $w \to \infty$, the average SNR over time equals to ρ.

Similar to [2], let ω_t denotes the *un-normalized* eigenvector and $v_t = \omega_t - u$ denotes the estimation error. Using Central Limit Theorem arguments [9], [10], we end up with

$$\begin{aligned}
\mathbb{E}_0[(\hat{u}_t^\intercal \boldsymbol{x}_t)^2] &= \sigma^2 + s^2(t)\|\alpha\|^2\mathbb{E}_0[(\hat{u}_t^\intercal u)^2] \\
&= \sigma^2 + s^2(t)\|\alpha\|^2\mathbb{E}_0\left[\frac{1}{1+\|v_t\|^2}\right] \\
&\approx \sigma^2 + s^2(t)\|\alpha\|^2\mathbb{E}_0[1-\|v_t\|^2] \\
&= \sigma^2 + s^2(t)\|\alpha\|^2\left(1-\frac{1+\rho}{w\rho^2}(k-1)\right).
\end{aligned}$$

For the approximate equality we used the fact that to a first order approximation we can write $1/(1+\|v_t\|^2) \approx 1-\|v_t\|^2$ because $\|v_t\|^2$ is of the order of $1/w$ while the approximation error is of higher order. This completes the proof. $\square$