ROBUST STATISTICAL PROCESS CONTROL IN BLOCK-RDT FRAMEWORK

Dominique Pastor and Quang-Thang Nguyen†*

Institut MINES-TELECOM; TELECOM Bretagne,
Technopôle Brest Iroise CS 83818, 29238 Brest Cedex 3, France
Université européenne de Bretagne, France

ABSTRACT

Random distortion testing (RDT) introduced in [1] is aimed at detecting any significantly big distortion of a signal with respect to a model of this signal, in presence of noise and without prior knowlegde on the distortion distribution. The RDT formulation makes it possible to state the standard change-in-mean detection problem differently. It leads to the Block-RDT approach that requires no iid assumption and no prior knowledge on the distributions of the observations before and after change. The optimal tests derived in the Block-RDT approach are alternative to Shewhart charts and outperform the latter by accounting for possible model mismatches that cause likelihood theory to fail. Experimental results dedicated to the detection of steps in a random process illustrate our intention.

Index Terms**—** Change detection, control charts, random distortion testing, statistical process control, non-stationary signal

1. INTRODUCTION

Since pionering works of Shewhart, Wald and Page [2–4], change detection is of crucial interest in many research and application areas, including quality control, monitoring, tracking, fault detection, statistical process control (SPC), signal processing, telecommunications, sensor networks and so forth (see [5–17] among others). A standard solution to the change-in-mean detection problem is the Shewhart chart [2]. Basically, the Shewhart chart assumes iid observations with known probability distributions, before and after change. It splits the observed process Y into blocks of N samples. In each block, the change-in-mean detection is posed as the binary hypothesis testing problem with null hypothesis $\mathcal{H}_0 : \xi = \xi_0$ and alternative hypothesis $\mathcal{H}_1 : \xi = \xi_1$, where ξ stands for the common expectation of the samples. The testing is then performed in each block by Neyman-Pearson (NP) test [5, Sec. 2.2.1]. The Shewhart control chart thus suffers from limitations of NP tests. In particular, it is fragile because it is not robust to possible fluctuations around the nominal model. Beyond this drawback, prior knowledge of ξ_1 is actually questionable and composite testing with alternative hypothesis $\xi \neq \xi_0$ should therefore be preferred. More generally, it seems unrealistic to assume prior knowledge on the process distribution, especially when this process is out of control after change. In addition, the iid assumption for observations is not satisfied in many statistical signal processing applications, if only the signal is deterministic. As a consequence, methods issued from the nonparametric change point model framework [7, 18–21] can hardly be directly applied in real world applications, as alternative to Shewhart charts.

This paper presents and discusses an alternative approach to Shewhart charts for the change-in-mean detection problem. This new approach overcomes limitations of Shewhart control charts without requiring the iid assumption and prior knowledge of the sample distributions. This is achieved by stating differently the change-in-mean detection problem thanks to the Random Distortion Testing (RDT) formulation introduced in [1]. More precisely, a change in the mean of the observed process Y is hereafter modeled as a significantly big distortion of the process empirical mean with respect to ξ_0. The change-in-mean detection problem then becomes the RDT problem of detecting such a significantly big distortion in independent and additive noise. Optimal tests exhibited in [1] are then used instead of NP ones to perform the change-in-mean detection in each block of N samples. The resulting procedure is thus called Block-RDT control chart.

After introducing terminology and notation, the main core of the paper is Section 2, where the Block-RDT formulation for the change-in-mean detection problem is introduced and one of the main theoretical results on it is stated. Some experimental results aimed at pedagogically illustrating the theoretical contents of Section 2 are then given in Section 3 before concluding the paper by some discussions and perspectives.

Notation and terminology

All random vectors are assumed to be defined on the same probability space $(\Omega, \mathcal{B}, \mathbb{P})$. With $[\![1, N]\!] = \{1, 2, \ldots, N\}$, $\mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$ denotes the set of all d-dimensional random processes defined on $[\![1, N]\!]$ and valued in $\mathbb{R}^d$. For any $U \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$, we write $U = (U_1, U_2, \ldots, U_N)$ where each U_n is an element of $\mathcal{M}(\Omega, \mathbb{R}^d)$. In the sequel, we confuse each $U \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$ with the $d \times N$-dimensional real random

*The first author is with UMR CNRS 6285 Lab-STICC
†The second author is with INSERM U1101 LaTIM

vector, obtained by stacking the columns of U under each other. The empirical mean of any given $U \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$ is the d-dimensional real random vector defined by $\langle U \rangle = (1/N)\sum_{n=1}^{N} U_n$. The $N \times N$ identity matrix is denoted by I_N.

For any $\rho \in [0, \infty)$, $\mathcal{R}(\rho, \cdot)$ denotes the cumulative distribution function (cdf) of the square root of any random variable that follows the non-central χ^2 distribution with d degrees of freedom and non-centrality parameter ρ^2. Given $\gamma \in (0, 1]$, $\lambda_\gamma(\rho)$ is univoquely defined for every $\rho \in [0, \infty)$ by the equality $\mathcal{R}(\rho, \lambda_\gamma(\rho)) = 1 - \gamma$.

2. CHANGE-IN-MEAN DETECTION BY BLOCK-RDT

We introduce the Block-RDT formulation for the change-in-mean detection problem in gaussian noise. We then define the performance criteria to devise the optimal tests for this formulation. These performance criteria are similar to the standard notions of size and power [22]. Their differences with standard ones rely mainly on the fact that the Block-RDT formulation is not a usual hypothesis testing problem. Indeed, it concerns random events and not hypotheses on deterministic parameters or random parameters with known distributions. This is the reason why such a problem is said to be an event testing problem, as in [1]. The optimality criterion chosen is then based on the natural invariance of the event testing problem under consideration. Subsection 2.3 introduces this optimality criterion and its optimal solution.

2.1. Problem statement

Let us consider some d-dimensional real random process $Y = \Xi + W$ where Ξ and W are elements of $\mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$, with $W \sim \mathcal{N}(0, \mathrm{I}_N \otimes \mathrm{C})$, where C is positive definite. Furthermore, W_n and Ξ_n are assumed to be independent for each $n \in [\![1, N]\!]$. In our formulation, Ξ models some distortion around the nominal model $\xi_0 \in \mathbb{R}^d$ so that, in absence of any distortion, Ξ would be equal to ξ_0, as in the model considered by the Shewhart chart. Due to distortions, we do not have $\Xi = \xi_0$ anymore. However, it can be expected that the empirical mean $\langle \Xi \rangle$ remains close to ξ_0 in the nominal situation and drifts significantly from this value in case of an abrupt change. This is the reason why our proposition is to test whether $\langle \Xi \rangle$ deviates significantly from ξ_0 or not. In order to compensate variations due to C, the deviation between $\langle \Xi \rangle$ and ξ_0 is measured by the Mahalanobis norm [23] $\|\langle \Xi \rangle - \xi_0\|$ of $\langle \Xi \rangle - \xi_0$, with $\|x\| = \sqrt{x^{\mathrm{T}} \mathrm{C}^{-1} x}$ for $x \in \mathbb{R}^d$. The role of tolerance τ is to distinguish small distortions from large ones of actual significance. This Block-RDT formulation is then summarized by:

$$
\begin{cases}
\textbf{Observation: } Y = \Xi + W \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]} \\
\text{with } \begin{cases} \Xi \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}, W \sim \mathcal{N}(0, \mathrm{I}_N \otimes \mathrm{C}), \\ \Xi_n \text{ and } W_n \text{ independent for each } n \in [\![1, N]\!], \end{cases} \\
\textbf{Null event : } \big[\|\langle \Xi \rangle - \xi_0\| \leqslant \tau \big], \\
\textbf{Alternative event : } \big[\|\langle \Xi \rangle - \xi_0\| > \tau \big].
\end{cases}
\tag{1}
$$

The following performance criteria are then suitable for solving this event testing problem.

2.2. Size and power of tests for Block-RDT

Given some natural number k, a k-dimensional test is any measurable map of $\mathbb{R}^k$ into $\{0, 1\}$. Let $\mathfrak{T}$ be some $d \times N$-dimensional test. For any signal $\Xi \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$ such that $\mathbb{P}\big[\|\langle \Xi \rangle - \xi_0\| \leqslant \tau\big] \neq 0$, the *size of* $\mathfrak{T}$ *for testing the empirical mean of* Ξ is then defined by:

$$
\alpha_\Xi(\mathfrak{T}) = \mathbb{P}\big[\mathfrak{T}(\Xi + W) = 1 \big| \|\langle \Xi \rangle - \xi_0\| \leqslant \tau\big]. \tag{2}
$$

Similarly, the *power of* $\mathfrak{T}$ *for testing the empirical mean of* Ξ is defined for every Ξ such that $\mathbb{P}\big[\|\langle \Xi \rangle - \xi_0\| > \tau\big] \neq 0$ by:

$$
\beta_\Xi(\mathfrak{T}) = \mathbb{P}\big[\mathfrak{T}(\Xi + W) = 1 \big| \|\langle \Xi \rangle - \xi_0\| > \tau\big]. \tag{3}
$$

Similarly to Neyman-Pearson's approach in binary hypothesis testing, we are looking for tests with guaranteed size and optimal power in a suitable sense. Accordingly, we search for tests whose size, for testing the mean of any given $\Xi \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$ such that $\mathbb{P}\big[\|\langle \Xi \rangle - \xi_0\| \leqslant \tau\big] \neq 0$, is upper-bounded by a specified level $\gamma \in (0, 1)$. If we define the *size* of a given $d \times N$-dimensional $\mathfrak{T}$ by:

$$
\alpha(\mathfrak{T}) = \sup_{\Xi \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]} : \mathbb{P}[\|\langle \Xi \rangle - \xi_0\| \leqslant \tau] \neq 0} \alpha_\Xi(\mathfrak{T}), \tag{4}
$$

we are thus looking for tests $\mathfrak{T}$ such that $\alpha(\mathfrak{T}) \leqslant \gamma$. Mimicking standard terminology in statistical inference, such tests are said to have level γ and the class of these tests with level γ is denoted by $\mathcal{K}_\gamma$. We also say that $\mathfrak{T}$ has size γ if $\alpha(\mathfrak{T}) = \gamma$.

On the other hand, there does not exist $\mathfrak{T}^* \in \mathcal{K}_\gamma$ such that $\beta_\Xi(\mathfrak{T}^*) \geqslant \beta_\Xi(\mathfrak{T})$ for any $\mathfrak{T} \in \mathcal{K}_\gamma$ and any $\Xi \in \mathcal{M}(\Omega, \mathbb{R}^d)^{[\![1,N]\!]}$ such that $\mathbb{P}\big[\|\langle \Xi \rangle - \xi_0\| > \tau\big] \neq 0$. However, as seen below, the event testing problem (1) exhibits some invariance properties. Thus, we restrict our attention to a certain class $\mathcal{C}$ of invariant tests and exhibit $\mathfrak{T}_{\text{opt}} \in \mathcal{C} \cap \mathcal{K}_\gamma$ such that $\beta_\Xi(\mathfrak{T}_{\text{opt}}) \geqslant \beta_\Xi(\mathfrak{T})$ for any $\mathfrak{T} \in \mathcal{C} \cap \mathcal{K}_\gamma$ and any Ξ with $\mathbb{P}\big[\|\langle \Xi \rangle - \xi_0\| > \tau\big] \neq 0$. Test $\mathfrak{T}_{\text{opt}}$ will be said to be UMP in $\mathcal{C} \cap \mathcal{K}_\gamma$ for the change-in-mean detection problem (1).

2.3. Block-RDT control chart

Let us first describe the relevant invariance properties of problem (1). First, we have the following eigenvector decomposition $\mathrm{C} = \mathrm{U}\Lambda\mathrm{U}^{\mathrm{T}}$, where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of C and U is a $d \times d$ orthogonal matrix. We now set: $\Phi = \Lambda^{-1/2}\mathrm{U}^{\mathrm{T}}$. Let $\mathcal{G}$ be the group of all transforms g_{R} such that $g_{\mathrm{R}}(x) = \Phi^{-1}\mathrm{R}\Phi(x - \xi_0) + \xi_0$ for any $x \in \mathbb{R}^d$, where R is an $d \times d$ orthogonal matrix. We now introduce the group action [24, Definition 6.3, p. 186] π that associates to each given $g \in \mathcal{G}$ the map $\pi_g : \mathbb{R}^{d\times N} \to \mathbb{R}^{d\times N}$ defined for every $x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^{d\times N}$ by: $\pi_g(x) = (g(x_1), g(x_2), \ldots, g(x_N))$. Some routine algebra shows that the event testing problem (1) is invariant under the action of π in that $\pi_g(Y) = \pi_g(\Xi) + W'$ with $W' = (W'_1, W'_2, \ldots, W'_N) \sim \mathcal{N}(0, \mathrm{I}_N \otimes \mathrm{C})$ and $g(\Xi_n)$ is independent of W'_n for each $n \in [\![1, N]\!]$. Therefore, $\pi_g(Y)$ satisfies the same hypotheses as Y. It can also be easily proved that $\|\langle\pi_g(\Xi)\rangle - \xi_0\| = \|\langle\Xi\rangle - \xi_0\|$. Hence, the change-in-mean detection problem (1) remains unchanged by substituting $\pi_g(\Xi)$ for Ξ and W' for W. It is thus natural to seek π-invariant $d \times N$-dimensional tests, that is, $d \times N$-dimensional tests $\mathfrak{T}$ invariant under the action of π: $\mathfrak{T}(\pi_g(x)) = \mathfrak{T}(x)$ for any $g \in \mathcal{G}$ and any $x \in \mathbb{R}^{d\times N}$.

Since the multiple d-dimensional samples in the same block gives the opportunity to reduce the noise variance by averaging observations, we consider π-invariant integrator tests, that is, π-invariant $d \times N$-dimensional tests $\mathfrak{T}$ for which there exists some d-dimensional test $\overline{\mathfrak{T}}$, henceforth called the reduced form of $\mathfrak{T}$, such that $\mathfrak{T}(x) = \overline{\mathfrak{T}}(\langle x\rangle)$ for any $x \in \mathbb{R}^{d\times N}$. Reduced forms of π-invariant integrator tests are $\mathcal{G}$-invariant: $\overline{\mathfrak{T}}(g(x)) = \overline{\mathfrak{T}}(x)$ for any $x \in \mathbb{R}^d$ and any $g \in \mathcal{G}$.

The following result derives from the foregoing and [1, Theorem 2], after noticing that the event testing problem (1) is an RDT problem [1].

Proposition 1 *Given $\gamma \in (0,1)$ and $\tau \in [0,\infty)$, put $\eta_N = \lambda_\gamma(\tau\sqrt{N})/\sqrt{N}$. Let $\mathfrak{T}_{\eta_N}$ be the $d \times N$-dimensional test defined for every $x \in \mathbb{R}^d$ by:*

$$\mathfrak{T}_{\eta_N}(x) = \begin{cases} 0 & \text{if} \quad \|\langle x\rangle - \xi_0\| \leqslant \eta_N \\ 1 & \text{otherwise,} \end{cases} \tag{5}$$

Test $\mathfrak{T}_{\eta_N}$ has size γ and is UMP among all π-invariant integrator tests with level γ for the change-in-mean detection problem (1).

3. EXPERIMENTAL RESULTS

In this work, we consider the detection of a change-in-mean in gaussian noise when some model mismatch is introduced. This example is particularly useful to highlight the added-value brought by the Block-RDT chart over the conventional Shewhart chart with respect to practical issues.

More specifically, let us begin with the following perfect model $Y_n = \Xi_n + W_n$, for $n \in \mathbb{N}$, where the clean signal Ξ_n is deterministic with $\Xi_n = \xi_0$ before change point and $\Xi_n = \xi_1$ after change point. We assume that $W_n \sim \mathcal{N}(0, \mathrm{C})$ for any $n \in \mathbb{N}$, where C is positive definite. As recalled in the introduction, the Shewhart chart proceeds per blocks of N samples. In each block, we test the null hypothesis $\mathcal{H}_0$ that the expectation of each sample is ξ_0 against the alternative hypothesis $\mathcal{H}_1$ that the expectation of each sample is ξ_1. As long as $\mathcal{H}_0$ is accepted, the hypotheses are tested on consecutive blocks of observations. Given $\gamma \in (0,1)$, $\mathcal{H}_0$ can be tested against $\mathcal{H}_1$ in each block by a Neyman-Pearson test with size γ. Unfortunately, there may be some unexpected mismatch between the model and the actual signal observed in practice. In fact, clean signal is generally not constant under each hypothesis because of unavoidable perturbations that can hardly be modelled. Because of such mismatch, the Shewhart control chart may fail to keep the false alarm rate bounded by γ. By construction, the Block-RDT chart is not affected by such limitations.

Therefore, instead of dealing with the perfect and somewhat unrealistic model described above, consider the case where $Y_n = \Xi_n + W_n$ for all $n \in \mathbb{N}$, with $\Xi_n = \xi_0 + \Delta_n$ before change and $\Xi_n = \xi_1 + \Delta_n$ after change. The random vectors Δ_n are additive distortions of the model with unknown distributions. Let us assume the amplitude of each Δ_n to be bounded by some positive value τ: $\mathbb{P}[\|\Delta_n\| \leq \tau] = 1, \forall n \in \mathbb{N}$. With this assumption, the testing on a given block can be performed, even in presence of distortions, thanks to Proposition 1.

To illustrate the foregoing, we carried out experiments for $d = 2$ and $\mathrm{C} = \sigma^2\mathrm{I}_d$. We chose $\Delta_n \sim \mathcal{N}(0, \sigma_\Delta^2\mathrm{I}_d)$ with $\sigma_\Delta = \tau/2$ for each n. Because $\mathbb{P}[\|\Delta_n\| \leq \tau] = 86.47\%$, we do not have $\|\Delta_n\| \leq \tau$. However, this does not really impact the results below. These ones were obtained with a Signal-to-Noise ratio (SNR) $\frac{\|\xi_1 - \xi_0\|}{\sigma}$ that varied up to 10dB and a Signal-to-maximum-Distortion Ratio (SDR) $\frac{\|\xi_1 - \xi_0\|}{\tau}$ fixed to 17dB. Fig.1 displays the false-alarm and detection rates ($\mathrm{P_{fa}}$ and $\mathrm{P_d}$, respectively) yielded by the Shewhart and Block-RDT charts for a given block. Different block-sizes values were considered to get these results. It turns out that Shewhart chart cannot guarantee any specified level $\gamma < 0.5$, although the distortion is quite small. In contrast, Block-RDT maintains the false-alarm rate below the specified value γ. Note that such results would remain exactly the same if the noise covariance matrix C were not scaled-identity, simply because the use of the Mahalanobis norm amounts to performing noise whitening before using the standard Euclidean norm.

To complete these results, figures Fig.2 and Fig.3 plot the mean run-length between false-alarms ($\mathrm{ARL_0}$) and the mean detection delay ($\mathrm{ARL_1}$) for different block-sizes. These two indices directly reflect the behavior of $\mathrm{P_{fa}}$ and $\mathrm{P_d}$ [5, 7]. Indeed, according to Figure 3, Block-RDT entails longer change decision delay than Shewhart chart for small SNR's. When the SNR is above 0 dB, the change decision delays of the Block-RDT and Shewhart charts are similar. As long as $\mathrm{ARL_0}$ is concerned, the duration between false-alarms is smaller by using Block-RDT than Shewhart chart, which implies that more conformed products can be produced by the former than the latter in quality control context. It also follows from these performance curves that the value of the

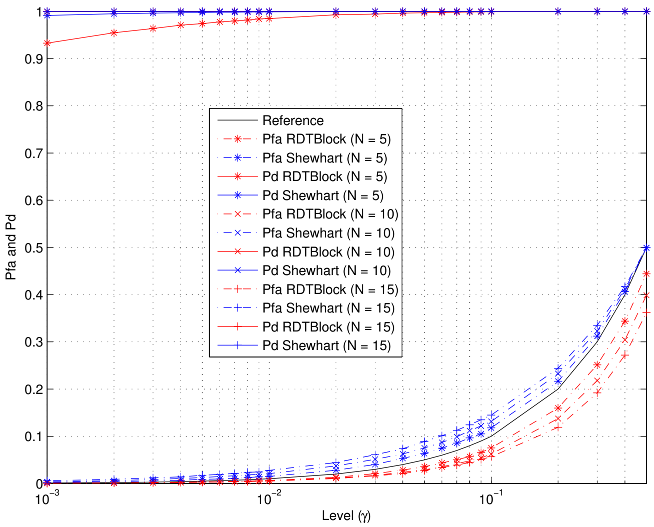

Fig. 1: Probabilities of false-alarm (P_{fa}) and good detection (P_d) yielded by Shewhart and Block-RDT charts for different values of block-size (N). The SNR and SDR were set to 5 dB and 17 dB respectively. The distortion was thus of small amplitude.

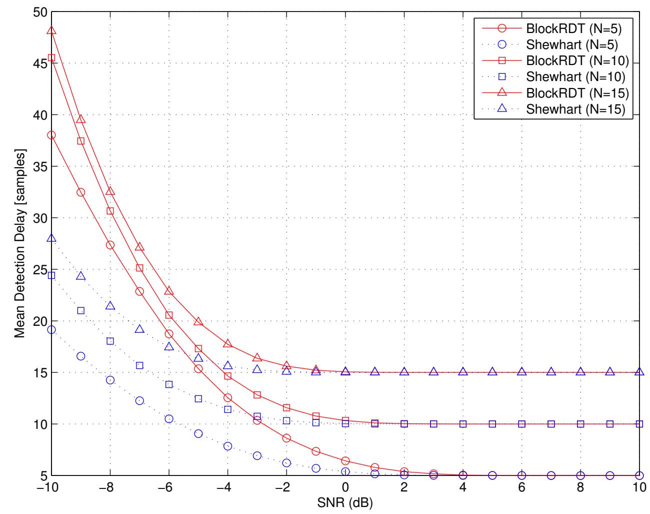

Fig. 3: Mean run-length between false-alarms (ARL_1) by Shewhart and Block-RDT charts for different values of block-size (N). The SDR was set to SNR $= 5$ dB as always. The level was fixed to $\gamma = 0.05$.

block-size must be made so as to achieve a trade-off between ARL_0 and ARL_1, with respect to the requirements for the control process. A sequential analysis approach could bypass this issue.

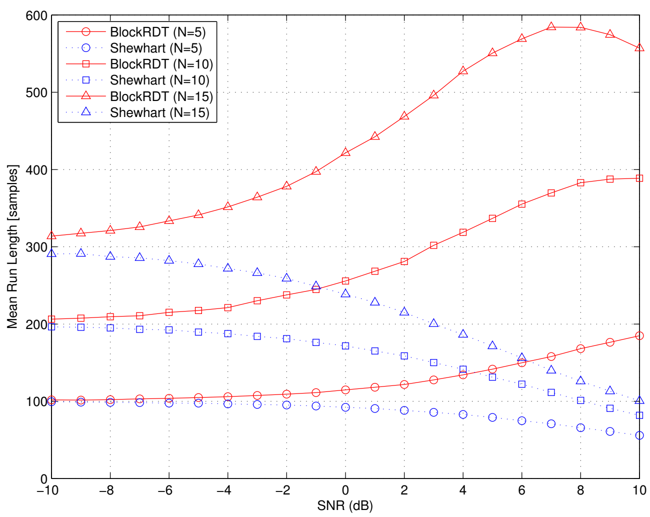

Fig. 2: Mean run-length between false-alarms (ARL_0) by Shewhart and Block-RDT charts for different values of block-size (N). The SDR was set to SNR $= 5$ dB as always. The level was fixed to $\gamma = 0.05$.

4. CONCLUSION AND PERSPECTIVES

The RDT formulation [1] has yielded a new model and a new control chart, namely, the Block-RDT control chart, for the detection of a change-in-mean in a d-dimensional real random signal observed in Gaussian noise. Block-RDT control chart performs blockwise detection of a change-in-mean, without iid assumption and prior knowledge on the signal distribution. It is an alternative to Shewhart control chart for situations where likelihood theory does not apply by lack of prior knowledge or may fail because of possible model mismatch.

Such results complete those already obtained within the RDT framework [1, 12, 25] and open prospects in detection of abrupt changes, in continuation to standard approaches such as those exposed in [7]. Especially, the RDT framework is basically intended to account for model mismatches, multi-dimensional and not necessarily iid observations. As such, extension of this framework and its applications match concrete issues encountered in real-world applications, inasmuch as Block-RDT has very low complexity. An extension to sequential framework should also be of interest for an adaptive choice of the optimal block size.

5. REFERENCES

[1] D. Pastor and Q.-T. Nguyen, “Random Distortion Testing and Optimality of Thresholding Tests,” *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4161 – 4171, Aug. 2013.

[2] W. A. Shewhart, *Economic Control of Quality Manufactured Product*. D.Van Nostrand Reinhold, Princeton, NJ, 1931.

[3] A. Wald, “Sequential tests of statistical hypotheses,” *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 06 1945.

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

[5] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes - Theory and Application*. Prentice-Hall, Inc., 1993.

[6] I. Hwang, S. Kim, Y. Kim, and C. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 636–653, May 2010.

[7] A. Tartakovsky, M. Basseville, and I. Nikiforov, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. Chapman & Hall/CRC, 2014.

[8] F. Desobry, M. Davy, and C. Doncarli, "An online kernel change detection algorithm," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2961–2974, Aug 2005.

[9] N. Vaswani, "Additive change detection in nonlinear systems with unknown change parameters," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 859–872, March 2007.

[10] J. Geng and L. Lai, "Non-bayesian quickest change detection with stochastic sample right constraints," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 5090–5102, Oct 2013.

[11] E. Brooks, R. Wynne, V. Thomas, C. Blinn, and J. Coulston, "On-the-fly massively multitemporal change detection using statistical quality control charts and landsat data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3316–3332, June 2014.

[12] Q.-T. Nguyen, D. Pastor, and E. L'Her, "Automatic detection of autopeep during controlled mechanical ventilation," *BioMedical Engineering OnLine*, vol. 11, no. 1, p. 32, 2012.

[13] R. Jana and S. Dey, "Change detection in teletraffic models," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 846–853, Mar 2000.

[14] T. He, S. Ben-David, and L. Tong, "Nonparametric change detection and estimation in large-scale sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1204–1217, April 2006.

[15] M. Xie, S. Han, B. Tian, and S. Parvin, "Anomaly detection in wireless sensor networks: A survey," *Journal of Network and Computer Applications*, vol. 34, no. 4, pp. 1302 – 1325, 2011.

[16] S. Zou, Y. Liang, P. H.V., and X. Shi, "Kernel-based nonparametric anomaly detection," in *15th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2014), Toronto, Canada, June 22-25*, 2014.

[17] A. Tartakovsky, B. Rozovskii, R. Blazek, and H. Kim, "A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3372–3382, Sept 2006.

[18] G. Ross, D. Tasoulis, and N. Adams, "Nonparametric monitoring of data streams for changes in location and scale," *Technometrics*, vol. 53, no. 4, pp. 379–389, 2011.

[19] G. J. Ross and N. M. Adams, "Two nonparametric control charts for detecting arbitrary distribution changes," *Journal of Quality Technology*, vol. 44, no. 2, pp. 102 – 116, 2012.

[20] S. Liu, M. Yamada, N. Collier, and M. Sugiyama, "Change-point detection in time-series data by relative density-ratio estimation," *Neural Networks*, vol. 43, no. 0, pp. 72 – 83, 2013.

[21] D. M. Hawkins and K. D. Zamba, "Statistical process control for shifts in mean or variance using a change-point formulation," *Technometrics*, vol. 47, no. 2, pp. 164 – 173, 2005.

[22] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses, 3rd edition*. Springer, 2005.

[23] P. Mahalanobis, "On the generalised distance in statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49 – 55, Feb. 1936.

[24] M. L. Eaton, *Multivariate statistics. A vector space approach*. Wiley, 1983.

[25] F.-X. Socheleau and D. Pastor, "Testing the energy of random signals in a known subspace: An optimal invariant approach," *IEEE Signal Processing Letters*, vol. 21, no. 10, pp. 1182–1186, Oct 2014.