Supplementary Materials for "A Distribution-Free Multivariate Control Chart"

Nan Chen[1], Xuemin Zi[2] and Changliang Zou[3*]

[1] *Department of Industrial & Systems Engineering, National University of Singapore, Singapore*

[2] *School of Science, Tianjin University of Technology and Education, China*

[3] *School of Mathematical Sciences, Nankai University, China*

April 27, 2015

1 Technical Proofs

Proof of Proposition 1 By the assumption that F_0 is continuous, we know with probability one all the coordinates of the $\mathbf{X}_i, \mathbf{Y}_j$'s are distinct. Let $\hat{F}^R_{m+n}(\mathbf{t})$ be the E.C.D.F. of the pooled rank samples $\mathcal{R}_{m+n} = \{\mathbf{R}_1, \ldots, \mathbf{R}_{m+n}\}$. By definition, T depends only on the rank $\mathbf{R}_i$, $i = 1, \ldots, m$. Thus, with probability one

$$\mathcal{L}\left(T \mid \hat{F}_{m+n}\right) = \mathcal{L}\left(T \mid \hat{F}^R_{m+n}\right).$$

Furthermore, given $\hat{F}^R_{m+n}$, there are totally $(m+n)!$ permutation of $\{1, \ldots, m+n\}$, resulting in at most C^m_{m+n} possible distinct values of T which do not depend on F_0. The assertion follows immediately. $\square$

Proof of Proposition 2 The first part is obvious by noting

$$\Pr(T \geq C_{\hat{F}}(\alpha)) = E(E(I(T \geq C_{\hat{F}}(\alpha))|\hat{F}_{m+n})) = E(\Pr(T \geq C_{\hat{F}}(\alpha) \mid \hat{F}_{m+n})) = \alpha.$$

Choose $c_{m,n} = 2p\log(m+n)$. Let T^* represent the corresponding counterpart of T which is obtained by a random permutation sample of the sample $\{\mathbf{X}_1, \ldots, \mathbf{X}_m, \mathbf{Y}_1, \ldots, \mathbf{Y}_n\}$. Note

*Corresponding author: nk.chlzou@gmail.com

that the distribution of T_j^* is exactly identical to the classical Wilcoxon rank-sum test statistic and hence $T_j^* \xrightarrow{d} N(0,1)$ as $m, n \to \infty$ (Serfling 1980). Accordingly,

$$\begin{aligned}\Pr(T^* \geq c_{m,n} \mid \hat{F}_{m+n}) &\leq \sum_{j=1}^{p} \Pr(T_j^{*2} \geq p^{-1} c_{m,n} \mid \hat{F}_{m+n}) \\ &= \sum_{j=1}^{p} \Pr(|T_j^*| \geq \sqrt{2\log(m+n)}) \leq 2p(4\pi \log(m+n))^{-1/2} \exp(-\log(m+n)) \to 0.\end{aligned}$$

On the other hand, $C_{\hat{F}}(\alpha)$ is chosen so that $\Pr(T^* \geq C_{\hat{F}}(\alpha) \mid \hat{F}_{m+n}) = \alpha$. Thus, $\Pr(C_{\hat{F}}(\alpha) < c_{m,n}) \to 1$.

Denote G_{0k}, G_{1k} as the marginal CDF of X_k and Y_k, respectively. Define

$$d_1 = \int (1 - G_{1k}(x)) dG_{0k}(x),\ d_2 = \int (1 - G_{1k}(x))^2 dG_{0k}(x),\ d_3 = \int G_{1k}^2(x) dG_{0k}(x).$$

Note that the condition $\Pr(X_k > Y_k) \neq 1/2$ is equivalent to $d_1 \neq 1/2$. Under H_1, $\sum_{i=1}^n R_{ki}$ is asymptotically normal with mean and variance

$$\begin{aligned}e_1 &= m(m+n+1)d_1, \\ v_1 &= mn(d_1 - d_1^2) + mn(n-1)(d_2 - d_1^2) + mn(m-1)(d_3 - d_1^2).\end{aligned}$$

By assumption, $d_1 \neq 1/2$. Without loss of generality, assume $d_1 > 1/2$. As a consequence, under H_1, with probability tending to one

$$\begin{aligned}&\Pr(T \geq C_{\hat{F}}(\alpha)) \geq \Pr(T \geq c_{m,n}) \geq \Pr(T_k \geq c_{m,n}) \\ &\geq \Pr\left(\frac{\sum_{i=1}^n R_{ki} - e_1}{\sqrt{v_1}} \geq \frac{\sqrt{mn(m+n+1)/12}\sqrt{2p\log(m+n)} + m(m+n+1)/2 - e_1}{\sqrt{v_1}} \right) \\ &= \Pr(Z > -\beta n^{1/2}(1 + o(1))) \to 1,\end{aligned}$$

where $\beta > 0$ is a positive constant and Z is a standard normal variable. □

Proof of Theorem 1 For simplicity, in the following argument the dependency of $T_n(w, \lambda)$ on λ and w is suppressed which should not cause any confusion. It is essentially to prove

$$\Pr(T_i < H_i(\alpha), 1 \leq i \leq n \mid \hat{F}_1, \ldots, \hat{F}_n) = (1-\alpha)^n. \tag{S.1}$$

This result can be proved by induction. First, $\Pr(T_1 < H_1(\alpha)) = 1 - \alpha$ by definition and thus (S.1) is true for $n = 1$. Assuming the assertion (S.1) is true for $n-1$, say

$\Pr(T_i < H_i(\alpha), 1 \le i < n \mid \hat{F}_1, \ldots, \hat{F}_{n-1}) = (1-\alpha)^{n-1}$, we have

$$\begin{aligned}
&\Pr(T_i < H_i(\alpha), 1 \le i \le n \mid \hat{F}_1, \ldots, \hat{F}_n) \\
&\quad = \Pr(T_n < H_n(\alpha), T_i < H_i(\alpha), 1 \le i < n \mid \hat{F}_1, \ldots, \hat{F}_n) \\
&\quad = \Pr(T_n < H_n(\alpha) \mid T_i < H_i(\alpha), 1 \le i < n, \hat{F}_1, \ldots, \hat{F}_n) \Pr(T_i < H_i(\alpha), 1 \le i < n \mid \hat{F}_1, \ldots, \hat{F}_n) \\
&\quad = \Pr(T_n < H_n(\alpha) \mid T_i < H_i(\alpha), 1 \le i < n, \hat{F}_n) \Pr(T_i < H_i(\alpha), 1 \le i < n \mid \hat{F}_1, \ldots, \hat{F}_{n-1}) \\
&\quad = (1-\alpha) \cdot (1-\alpha)^{n-1} = (1-\alpha)^n.
\end{aligned}$$

The third last equality holds because $\hat{F}_n$ is a sufficient statistic if only $\mathcal{X}^n_{-m_0+1}$ is observed. Therefore given $\hat{F}_n$, the distribution of T_n would be independent of $\hat{F}_1, \ldots, \hat{F}_{n-1}$. In addition, since $T_i, i = 1, 2, \ldots, n-1$ is independent of $\mathbf{X}_n$, they are independent of $\hat{F}_n$ conditional on $\hat{F}_i, i = 1, 2, \ldots, n-1$. Finally,

$$\begin{aligned}
&\Pr(RL = n) = E[\Pr(RL = n \mid \hat{F}_1, \ldots, \hat{F}_n)] \\
&\quad = E[\Pr(T_n > H_n(\alpha), T_i < H_i(\alpha), 1 \le i < n \mid \hat{F}_1, \ldots, \hat{F}_n)] \\
&\quad = E[\Pr(T_n > H_n(\alpha) \mid T_i < H_i(\alpha), 1 \le i < n, \hat{F}_n) \Pr(T_i < H_i(\alpha), 1 \le i < n \mid \hat{F}_1, \ldots, \hat{F}_{n-1}) \\
&\quad = E[\alpha(1-\alpha)^{n-1}] = \alpha(1-\alpha)^{n-1}.
\end{aligned}$$

□

Proof of Proposition 3 Because the other cases are similar, we only show that $T_{jn}(w, \lambda)$ and $T_{jk}(w, \lambda)$ are independent. It suffices to prove $R_{jnk_1}, n-w+1 \le k_1 \le n$ are independent of $R_{j,n-w,k_2}, k_2 \le n-w$. Note that $R_{j,n,k_1} = \sum_{i=-m_0+1}^{n-w} I(X_{jk_1} \ge X_{ji}) + \sum_{i=n-w+1}^{n} I(X_{jk_1} \ge X_{ji})$. The first part is independent of $R_{j,n-w,k_2}$ because $R_{j,n-w,k_2}$ is determined only by the order statistics of $\mathcal{X}^{n-w}_{-m_0+1,j}$. In addition, since $X_{jl}, l = n-w+1, \ldots, n$ is independent of X_{jk_2}, the second part is also independent of $R_{j,n-w,k_2}$. This completes the proof. □

2 Extensions of the Proposed Chart

Although we focus on the DFEWMA chart, its idea can be easily extended to construct other control charts as long as the test statistic is based on the marginal empirical distributions so that the distribution-free property holds.

For instance, if the change-point detection method is considered, we may simply replace $T_n(w, \lambda)$ by $\max_{1 \le k \le n} T_n(k, 0)$ or $\sum_{k=1}^{n} T_n(k, 0)$ which correspond to the generalized likelihood ratio method (Lai 1995) and Shiryaev-Roberts procedure (Shiryaev 1961), respectively.

To alleviate computational effort, the window-limited schemes like

$$\max_{n-w_2+1\leq k\leq n} T_n(k,0) \quad \text{and} \quad \sum_{k=n-w_2+1}^{n} T_n(k,0),$$

where w_2 is a positive integer, are often used instead (Hawkins, Qiu, and Kang 2003).

In addition, when m_0 is sufficiently large or F_0 is known, $T_{jn}(w,\lambda)$ would essentially become

$$\sum_{i=n-w+1}^{n} (1-\lambda)^{n-i} \frac{V_{ji}-0.5}{\sqrt{1/12}},$$

where V_{ji} is the (approximate) marginal CDF of X_{ji}, 0.5 and 1/12 are simply the mean and variance of V_{ji} for standardization. In this situation, we may simplify the procedure by using the traditional EWMA recursive form, say

$$S_{jn} = (1-\lambda)S_{j,n-1} + \lambda V_{jn}. \tag{S.2}$$

Accordingly, the charting statistic would be $\sum_{j=1}^{p} S_{jn}^2$. Also, the data-dependent dynamic control limits are not required because we can always use re-sampling method to find the control limits for this particular F_0 so that the desired IC run length distribution is attained.

Moreover, the current version of the proposed scheme is designed to detect location shifts only. We believe that, after certain modifications, the proposed method should be able to effectively detect a general change in distribution over time. It is amount to obtain change information from each marginal distribution by some univariate powerful nonparametric goodness-of-test statistics (c.f. Zou and Tsung 2010) instead of the Wilcoxon rank-sum test statistic $T_{jn}(w,\lambda)$ and then integrate them together as done in $T_n(w,\lambda)$.

To end this subsection, we note that when a group of g observations $\{\mathbf{X}_{i1},\ldots,\mathbf{X}_{ig}\}$ are taken sequentially from the process at each time point i, the DFEWMA chart can be readily defined in a similar way to (2) in the paper by using the following modified empirical distribution function,

$$\hat{F}_n^{(g)}(\mathbf{t}) = \frac{1}{(n+m_0)g} \sum_{i=-m_0+1}^{n} \sum_{k=1}^{g} I(\mathbf{X}_{ik} \leq \mathbf{t}).$$

The corresponding ranks can be defined in terms of $\hat{F}_n^{(g)}(\mathbf{t})$ and testing statistic is obtained accordingly.

3 Skewness and kurtosis of multivariate distributions

Table S.1: Skewness and Kurtosis of multivariate normal, multivariate $t_{p,5}$, and multivariate Gamma distribution of different dimension. * in the parentheses indicates that the values are obtained based on sample estimation.

		Skewness			Kurtosis		
	p	Multinormal	$t_{p,5}$	$\text{Gam}_{p,3}$	Multinormal	$t_{p,5}$	$\text{Gam}_{p,3}$
Univariate	1	0	0	$\sqrt{8/3}$	3	9	7
Multivariate	10	0	0	28(*)	120	360	170(*)
	30	0	0	86(*)	960	2880	1103(*)

4 Comparisons with adaptive charts

Tsung and Wang (2010) reviewed adaptive control charts and classified the adaptive control charts into two categories: the ones with adaptive sampling parameters (e.g., sample size, sampling interval); and the ones with adaptive design parameters (e.g. reference parameter in CUSUM or smoothing parameter in EWMA). These charts are mainly designed for the following cases: a) detecting unknown or mixed-range shifts; b) monitoring processes with dynamic means; and c) detecting dynamic shifts from a stable process. Our approach on the surface is similar to adaptive control charts, in the sense that it changes control limits "adaptively". However, there are also some fundamental differences. First of all, in terms of charting performance, our chart can always attain the target IC run length distribution regardless of the distribution or dimension of F_0. In contrast, to our best knowledge existing adaptive charts cannot have such robust performance. In particular, in order to optimize the charting performance, they often need the knowledge of F_0 to find the optimal adaptive design setting. Secondly, the adaptive parameters are at most *distribution-dependent*. That is, given F_0, the adaptive charts, including its design parameters (or the optimal way it is changing) are fixed. On the other hand, our chart is rather *data-dependent*. It changes along the data instead of the distribution of the data. This is also the key success factor that our chart has exact IC performance compared with other adaptive charts. In the supplementary material, we provide simulation comparisons with typical adaptive charts to demonstrate these points.

In this section, we compare with some adaptive control charts. Two classical and effective adaptive control charts are compared. The first one (termed Adaptive T^2) was proposed by Zhu and Jiang (2009). It adaptively identifies a set of potentially shifted variables, and uses

Table S.2: IC performance of the DFEWMA, Adaptive T^2, and AMEWMA charts with ten-dimensional multivariate normal observations

m_0	Method	ARL_0	SDRL	FAR
	Geometric	200	200	0.140
50	DFEWMA ($\lambda = 0.1$)	201	194	0.133
	DFEWMA ($\lambda = 0.05$)	202	197	0.137
	Adaptive T^2	20.7	24.9	0.78
	AMEWMA	20.0	17.3	0.79
100	DFEWMA ($\lambda = 0.1$)	197	193	0.135
	DFEWMA ($\lambda = 0.05$)	200	194	0.128
	Adaptive T^2	52.8	58.4	0.45
	AMEWMA	45.0	39.1	0.43

regression adjusted method to construct a T^2 chart. The second one (termed AMEWMA, Mahmoud and Zahran 2010) is to adjust the smoothing parameter λ in MEWMA chart based on the difference between current observations and previous charting statistic. If the difference is small, a smaller λ is used; if the difference is large, a larger λ is used. The design parameters of the adapting rule is optimized to have satisfactory performance in detecting both small shifts and large shifts.

Similar to existing results in the paper, we compare them in terms of IC and OC performance. Since both adaptive T^2 and AMEWMA requires inverse of covariance matrix to compute the charting statistics, we expect that their IC performance is not satisfactory when m_0 is small and p is large. In fact, similar to the results in Table S.2, even when the data are normally distributed, both adaptive charts cannot attain specified run length distribution. Their IC performance are similar to those of MEWMA or MSEWMA. It is expected that its IC performance could be even poorer if data are not normally distributed.

We also compared their OC performance when data follow different distributions. Similar to the settings in the paper, we adjust the control limits such that their IC ARL equals 200. Their OC performance is shown in Table S.3. It is clear from the table that AMEWMA generally perform well in detecting large shifts, while not as good as DFEWMA in detecting small shifts. When the data are t_5 or Gamma distributed, AMEWMA also has worse performance. Regarding adaptive T^2 chart, its performance is similar to other T^2 charts: good at detecting large shifts while insensitive in detecting small shifts. Moreover, when data are not normally distributed, the performance is far from satisfactory.

In fact, since most of the adaptive charts need knowledge of F_0 to optimize its parameters, its IC performance is not robust when F_0 is unknown and m_0 is small. This feature also distinguish the adaptive charts from our chart, at least in terms of IC performance.

5 Other real data applications

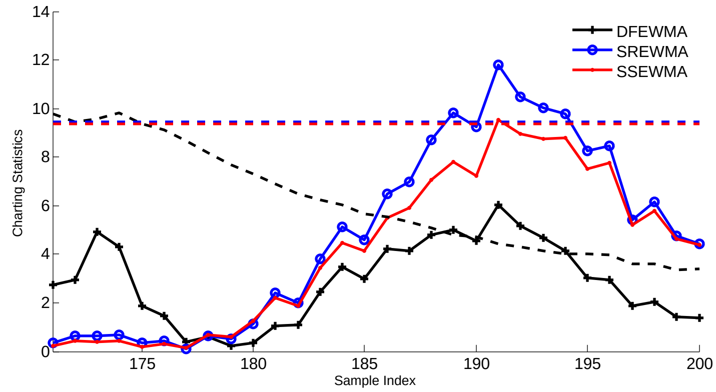
Although DFEWMA is most advantageous and robust when p is large and m_0 is relatively small, it also performs satisfactorily in monitoring data with low dimensions. This section demonstrates this point through two well studied datasets.

The first dataset came from an aluminum electrolytic capacitor (AEC) manufacturing process. It contains 200 observations. Each observation is composed of three quality characteristics of an AEC, namely the capacitance, loss tangent, and leakage current level. The first 170 observations were believed to be IC samples, while the remaining 30 samples are used to illustrate the chart performance. For a detailed description of the dataset, please refer to Zou and Tsung (2011). We compare DFEWMA with two alternatives, self-starting EWMA chart (SSEWMA, Hawkins and Maboudou-Tchao 2007) and spatial rank EWMA chart (SREWMA, Zou et al. 2012). The control limits of SSEWMA and SREWMA are adjusted such that the ARL_0 is close to 200 through resampling of the 170 IC samples. Correspondingly, α is set to 0.005 for DFEWMA chart. Figure S.1a compares the operation of three charts on the remaining 30 observations. It shows DFEWMA signals the change at the same time as SREWMA chart, while slightly earlier than SSEWMA chart. However, DFEWMA does not require extensive tuning of control limits through resampling, and the IC performance is guaranteed.

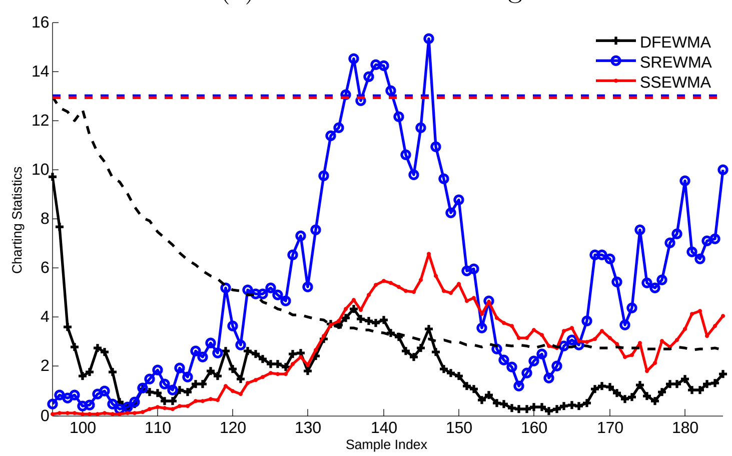
We also studied a dataset from aluminum smelter process. The dataset contains five variables (the contents of SiO_2, Fe2O3, MgO, CaO, and Al2O3), and 185 observations. It was used widely in the MSPC literature (e.g. Qiu and Hawkins 2001; Zamba and Hawkins 2006; Hawkins and Maboudou-Tchao 2007). Similar to their settings, we use the first 95 observations as IC reference samples. The control limits are set to 9.373 and 9.450 for SSEWMA and SREWMA respectively, such that ARL_0 is 200. Figure S.1b compares their performance in monitoring the last 90 observations. We can observe that DFEWMA signals change slightly earlier than SREWMA chart around 135th observation. In contrast, SSEWMA does not produce any alarm throughout the run. This also demonstrates that with normality assumption, SSEWMA is not robust and sensitive to detect changes for non-normal distributions.

Table S.3: OC ARL comparison of the DFEWMA, Adaptive T^2, and AMEWMA charts when m_0 is not sufficiently large; numbers in parentheses are SDRL values.

			DFEWMA			
	(m_0, p)	δ	$\lambda = 0.1$	$\lambda = 0.05$	Adaptive T^2	AMEWMA
Normal	$(100, 10)$	0.5	69.3(90.9)	44.5(46.1)	153 (150)	70.6(69.9)
		1.0	14.3(9.08)	14.4(6.67)	72.0(72.9)	20.3(10.7)
		2.0	6.84(2.11)	7.75(2.50)	9.45(9.12)	6.73(3.53)
		4.0	5.45(1.46)	6.53(1.88)	1.16(0.43)	1.37(0.70)
	$(200, 30)$	0.5	22.7(17.7)	22.4(13.0)	177 (180)	102 (95.2)
		1.0	8.47(3.49)	9.55(3.69)	122 (129)	29.7(17.5)
		2.0	4.65(1.35)	5.71(1.68)	21.6(24.5)	7.06(4.00)
		4.0	3.91(1.00)	4.76(1.43)	1.20(0.53)	1.21(0.46)
t_5	$(100, 10)$	0.5	101(137)	64.8(85.2)	202 (212)	85.4(89.8)
		1.0	18.2(14.0)	17.3(9.02)	197 (213)	25.3(16.2)
		2.0	7.73(2.64)	9.04(2.92)	166 (173)	9.10(2.72)
		4.0	5.80(1.53)	6.80(1.99)	79.4(93.2)	4.30(0.82)
	$(200, 30)$	0.5	37.7(45.5)	28.0(20.2)	199 (208)	72.8(59.3)
		1.0	10.4(4.51)	11.3(4.64)	195 (205)	25.2(9.88)
		2.0	5.36(1.63)	6.26(1.97)	182 (189)	10.8(2.20)
		4.0	4.14(1.07)	5.05(1.40)	139 (141)	5.39(0.77)
Gam_3	$(100, 10)$	0.5	56.8(90.8)	36.5(55.7)	189 (222)	108 (134)
		1.0	12.1(5.49)	12.7(5.33)	168 (200)	40.0(51.2)
		2.0	6.39(1.90)	7.51(2.33)	97.5(130)	8.76(5.89)
		4.0	4.94(1.35)	6.22(1.79)	20.6(32.4)	3.07(0.78)
	$(200, 30)$	0.5	19.0(13.3)	18.1(8.24)	200 (235)	100 (95.7)
		1.0	7.40(2.48)	8.57(2.70)	172 (208)	29.8(25.7)
		2.0	4.43(1.16)	5.39(1.58)	70.5(94.0)	5.62(0.33)
		4.0	3.52(0.95)	4.42(1.23)	5.87(7.69)	2.13(0.39)



(a) AEC manufacturing



(b) Aluminum smelter process

Figure S.1: Comparisons between DFEWMA chart and SREWMA, SSEWMA charts in monitoring two real data. The solid lines are charting statistics as indicated in the legend. The dashed lines are corresponding control limits for each method.

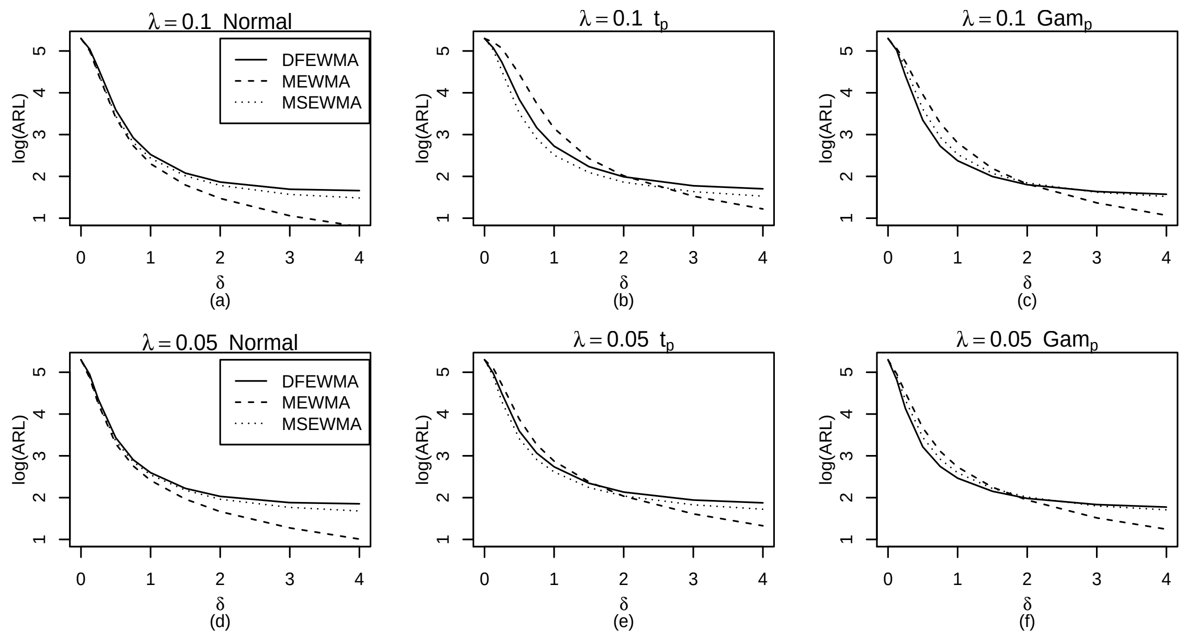


Figure S.2: OC ARL comparison of the DFEWMA, MSEWMA and MEWMA charts using $\lambda = 0.05$ and 0.1 when $p = 10$ and $m_0 = 100,000$

6 Additional Simulation Results

6.1 Large m_0 comparisons with MSEWMA, MEWMA

Figure S.2 compares the OC ARL of the DFEWMA, MSEWMA and MEWMA charts using $\lambda = 0.05$ and 0.1 when $p = 30$ and $m_0 = 100,000$.

6.2 Complete Tables of Section 4.2.2

Table S.4 to Table S.6 compare the OC performance among DFEWMA, SREWMA, SSEWMA, and RTC when the mean shifts in different patterns. m_0 is set to 100, and different data distributions and data dimensions are considered.

6.3 Comparisons with SREWMA, SSEWMA, RTC in other shift scenarios

Table S.7 to Table S.9 summarize the results when the mean shift according to different patterns, and when the marginal distribution of the p-variate are not identical.

Table S.4: OC ARL comparison of the DFEWMA, SREWMA, SSEWMA, and RTC charts with multivariate normal observations; numbers in parentheses are SDRL values.

		$\lambda = 0.1$			$\lambda = 0.05$			
(m_0, p)	δ	DFEWMA	SREWMA	SSEWMA	DFEWMA	SREWMA	SSEWMA	RTC
(50,10)	0.5	86.4(115)	76.0(127)	72.8(115)	60.4(79.3)	52.8(79.7)	50.9(68.8)	71.2(90.1)
	1.0	16.1(16.6)	12.6(8.19)	13.4(10.7)	16.0(8.68)	13.7(7.10)	14.0(7.35)	15.7(13.3)
	2.0	6.97(2.08)	5.51(1.98)	5.00(1.93)	8.23(2.55)	6.80(2.49)	6.20(2.32)	6.61(1.89)
	4.0	5.56(1.44)	3.61(1.05)	2.58(0.82)	6.75(1.93)	4.61(1.48)	3.36(1.14)	5.07(1.22)
(100,10)	0.5	69.3(90.9)	49.2(73.4)	52.6(71.9)	44.5(46.1)	36.9(43.9)	37.9(42.4)	59.1(67.4)
	1.0	14.3(9.08)	11.4(6.38)	11.6(6.81)	14.4(6.67)	12.6(5.85)	12.6(5.85)	14.0(9.86)
	2.0	6.84(2.11)	5.24(1.77)	4.66(1.65)	7.75(2.50)	6.48(2.25)	5.80(2.03)	6.39(1.74)
	4.0	5.45(1.46)	3.44(0.97)	2.44(0.75)	6.53(1.88)	4.35(1.30)	3.10(0.98)	4.92(1.14)
(50,30)	0.5	44.0(82.6)	95.7(227)	72.0(115)	29.4(27.7)	45.3(88.9)	47.4(66.3)	38.4(44.0)
	1.0	9.00(3.83)	11.4(13.5)	12.8(10.8)	10.4(4.30)	12.0(6.85)	13.4(6.39)	9.50(4.73)
	2.0	4.89(1.45)	5.00(2.01)	4.94(1.79)	5.98(1.86)	6.24(2.62)	6.32(2.23)	5.55(1.41)
	4.0	4.06(1.07)	3.28(1.13)	2.67(0.85)	5.04(1.45)	4.23(1.59)	3.61(1.16)	4.38(1.06)
(100,30)	0.5	28.3(32.9)	44.7(99.6)	50.5(74.6)	24.3(15.3)	28.7(35.1)	35.8(36.9)	33.1(32.3)
	1.0	8.53(3.65)	9.40(4.80)	10.7(5.45)	10.0(3.69)	10.9(4.89)	12.0(4.98)	8.96(3.96)
	2.0	4.79(1.36)	4.62(1.62)	4.50(1.51)	5.80(1.69)	5.84(2.04)	5.78(1.87)	5.29(1.36)
	4.0	3.98(1.04)	3.02(0.90)	2.42(0.71)	4.84(1.35)	3.88(1.24)	3.18(0.94)	4.25(1.01)

References

Hawkins, D. M. and Maboudou-Tchao, E. M. (2007), "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting," *Technometrics*, 49, 199–209.

Hawkins, D. M., Qiu, P., and Kang, C. W. (2003), "The Changepoint Model for Statistical Process Control," *Journal of Quality Technology*, 35, 355–366.

Lai, T. L. (1995), "Sequential Changepoint Detection in Quality Control and Dynamical Systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 613–658.

Mahmoud, M. A. and Zahran, A. R. (2010), "A Multivariate Adaptive Exponentially Weighted Moving Average Control Chart," *Communications in Statistics - Theory and Methods*, 39, 606–625.

Table S.5: OC ARL comparison of the DFEWMA, SREWMA, SSEWMA and RTC charts with multivariate $t_{p,5}$ observations; numbers in parentheses are SDRL values.

		$\lambda = 0.1$			$\lambda = 0.05$			
(m_0, p)	δ	DFEWMA	SREWMA	SSEWMA	DFEWMA	SREWMA	SSEWMA	RTC
(50,10)	0.5	117(148)	89.5 (142)	140(158)	88.4(111)	64.7(100)	101(131)	86.7(116.)
	1.0	23.8(30.2)	16.8 (20.5)	54.6(96.4)	19.7(13.0)	16.2(13.1)	24.0(28.3)	19.6(18.9)
	2.0	8.25(2.79)	6.25 (2.60)	8.49(4.15)	9.27(3.24)	7.51(2.98)	8.73(3.61)	7.29(2.26)
	4.0	5.83(1.61)	3.85 (1.21)	3.75(1.18)	7.04(2.10)	4.86(1.63)	4.37(1.47)	5.49(1.47)
(100,10)	0.5	101(137)	69.6 (103)	132(152)	64.8(85.2)	49.6(59.9)	82.0(105)	73.1(87.7)
	1.0	18.2(14.0)	14.5 (9.61)	37.4(55.4)	17.3(9.02)	15.0(7.62)	20.7(15.2)	17.2(14.0)
	2.0	7.73(2.64)	6.12 (2.19)	8.06(3.35)	9.04(2.92)	7.39(2.67)	8.21(3.04)	7.19(2.09)
	4.0	5.80(1.53)	3.78 (1.09)	3.59(1.07)	6.80(1.99)	4.76(1.44)	4.11(1.29)	5.35(1.40)
(50,30)	0.5	70.1(100)	124 (221)	145(138)	45.2(64.4)	65.9(129)	94.8(112)	56.8(80.8)
	1.0	11.5(6.21)	17.4 (36.1)	69.4(99.7)	12.8(5.63)	14.5(9.13)	23.7(21.9)	11.8(7.80)
	2.0	5.55(1.66)	5.86 (2.45)	9.97(11.4)	6.75(2.18)	7.11(3.02)	8.92(3.26)	6.17(1.75)
	4.0	4.31(1.10)	3.63 (1.21)	4.24(1.25)	5.27(1.45)	4.61(1.70)	4.80(1.50)	4.83(1.34)
(100,30)	0.5	47.6(71.8)	71.3 (138)	140(143)	31.9(36.6)	42.2(59.8)	81.1(97.4)	51.6(66.0)
	1.0	11.1(5.38)	12.6 (8.59)	51.7(75.3)	11.8(4.62)	13.4(6.52)	20.2(12.3)	11.0(6.72)
	2.0	5.40(1.66)	5.45 (1.94)	8.80(4.64)	6.48(2.10)	6.70(2.37)	8.39(2.80)	6.10(1.66)
	4.0	4.15(1.15)	3.36 (0.99)	3.87(1.03)	5.06(1.46)	4.26(1.35)	4.32(1.24)	4.79(1.25)

Qiu, P. and Hawkins, D. (2001), "A Rank-Based Multivariate CUSUM Procedure," *Technometrics*, 43, 120–132.

Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, New York: John Wiley & Sons.

Shiryaev, A. N. (1961), "The Problem of the Most Rapid Detection of a Disturbance in a Stationary Process," *Soviet Mathmatics-Doklady*, 2, 795–799.

Tsung, F. and Wang, K. (2010), *Adaptive Charting Techniques: Literature Review and Extensions*, Springer, pp. 19–35.

Zamba, K. D. and Hawkins, D. M. (2006), "A Multivariate Change-Point Model for Statistical Process Control," *Technometrics*, 48, 539–549.

Table S.6: OC ARL comparison of the DFEWMA, SREWMA, SSEWMA, and RTC charts with multivariate $\mathrm{Gam}_{p,3}$ observations; numbers in parentheses are SDRL values.

		$\lambda = 0.1$			$\lambda = 0.05$			
(m_0, p)	δ	DFEWMA	SREWMA	SSEWMA	DFEWMA	SREWMA	SSEWMA	RTC
(50,10)	0.5	78.2(122)	68.3(99.7)	78.1(114)	49.7(75.2)	90.9(134)	107(140)	77.8(107.)
	1.0	13.6(9.40)	17.1(11.7)	20.9(18.9)	13.9(6.56)	18.9(25.4)	32.0(55.3)	13.3(12.0)
	2.0	6.61(1.98)	7.69(2.88)	8.04(3.44)	7.99(2.53)	6.40(2.42)	7.39(3.67)	6.67(1.83)
	4.0	5.13(1.40)	4.94(1.55)	4.08(1.43)	6.27(1.86)	3.94(1.14)	3.40(1.13)	5.32(1.34)
(100,10)	0.5	56.8(90.8)	48.8(58.2)	65.1(89.0)	36.5(55.7)	70.3(103)	91.2(120)	64.6(76.9)
	1.0	12.1(5.49)	15.3(7.58)	17.9(11.6)	12.7(5.33)	15.2(9.88)	23.7(31.4)	11.5(7.20)
	2.0	6.39(1.90)	7.39(2.50)	7.56(2.90)	7.51(2.33)	6.19(2.09)	6.81(2.80)	6.59(1.68)
	4.0	4.94(1.35)	4.72(1.39)	3.82(1.26)	6.22(1.79)	3.78(1.03)	3.19(0.99)	5.13(1.25)
(50,30)	0.5	29.7(49.9)	102(238)	79.9(114)	21.9(15.2)	47.8(92.1)	52.5(74.3)	39.6(54.2)
	1.0	7.97(2.85)	12.8(21.6)	17.8(25.3)	9.40(3.30)	12.5(7.14)	14.5(7.90)	7.90(2.38)
	2.0	4.67(1.24)	5.15(2.05)	5.67(2.25)	5.79(1.60)	6.44(2.66)	6.72(2.51)	5.49(1.44)
	4.0	3.69(0.94)	3.37(1.10)	2.95(0.94)	4.67(1.25)	4.34(1.59)	3.83(1.30)	4.34(1.12)
(100,30)	0.5	21.9(21.2)	49.2(109)	63.0(86.9)	19.4(9.88)	31.9(36.9)	41.7(49.3)	33.0(34.0)
	1.0	7.55(2.62)	10.4(5.92)	13.7(10.6)	9.06(3.06)	11.6(5.27)	13.5(6.07)	7.72(2.05)
	2.0	4.49(1.23)	4.91(1.66)	5.18(1.83)	5.57(1.61)	6.08(2.08)	6.23(2.11)	5.39(1.33)
	4.0	3.65(0.92)	3.13(0.88)	2.70(0.80)	4.46(1.22)	4.02(1.22)	3.40(1.05)	4.25(1.01)

Zhu, Y. and Jiang, W. (2009), "An Adaptive T2 Chart for Multivariate Process Monitoring and Diagnosis," *IIE Transactions*, 41, 1007–1018.

Zou, C. and Tsung, F. (2010), "Likelihood Ratio-Based Distribution-Free EWMA Control Charts," *Journal of Quality Technology*, 42, 174–196.

— (2011), "A Multivariate Sign EWMA Control Chart," *Technometrics*, 53, 84–97.

Zou, C., Wang, Z., and Tsung, F. (2012), "A Spatial Rank-Based Multivariate EWMA Control Chart," *Naval Research Logistics*, 59, 91–110.

Table S.7: OC ARL comparison of the DFEWMA, SREWMA, SSEWMA, and RTC charts when $m_0 = 100$ and all components shift with equal size; numbers in parentheses are SDRL values.

			$\lambda = 0.1$			$\lambda = 0.05$			
	p	δ	DFEWMA	SREWMA	SSEWMA	DFEWMA	SREWMA	SSEWMA	RTC
MVN	10	0.5	24.4(25.3)	51.2(75.4)	54.5(80.9)	23.5(19.2)	37.8(43.4)	38.5(40.9)	35.4(34.8)
		1.0	7.99(3.78)	11.8(6.60)	11.9(6.72)	9.29(4.03)	13.0(6.23)	12.8(5.95)	10.0(5.24)
		2.0	3.97(1.32)	5.37(1.86)	4.77(1.72)	4.92(1.63)	6.61(2.32)	5.98(2.08)	5.43(1.65)
		4.0	2.75(0.77)	3.50(0.98)	2.47(0.77)	3.39(0.99)	4.46(1.34)	3.15(1.00)	3.85(1.12)
	30	0.5	18.8(15.8)	27.6(49.9)	36.6(52.4)	19.2(11.9)	22.3(16.8)	27.0(22.9)	25.2(23.3)
		1.0	6.63(2.77)	7.83(3.47)	8.71(3.73)	7.97(3.11)	9.43(3.86)	10.0(3.82)	7.95(2.95)
		2.0	3.42(1.02)	4.16(1.34)	3.92(1.27)	4.23(1.32)	5.31(1.73)	5.05(1.57)	4.62(1.34)
		4.0	2.28(0.60)	2.89(0.82)	2.18(0.62)	2.82(0.77)	3.77(1.13)	2.88(0.86)	3.20(0.94)
t_5	10	0.5	33.4(42.6)	72.7(118)	133(155)	28.2(26.8)	48.8(59.5)	81.1(105)	46.6(51.2)
		1.0	9.52(5.02)	14.9(10.8)	39.2(63.8)	10.7(5.12)	15.2(7.85)	20.7(13.9)	11.4(7.27)
		2.0	4.49(1.60)	6.14(2.20)	8.14(3.41)	5.43(2.04)	7.42(2.66)	8.24(3.04)	6.04(1.88)
		4.0	2.98(0.87)	3.81(1.09)	3.63(1.09)	3.63(1.17)	4.77(1.47)	4.17(1.31)	4.44(1.35)
	30	0.5	29.7(40.8)	126(137)	49.6(110)	23.4(16.8)	59.5(70.8)	30.2(32.3)	39.2(44.3)
		1.0	8.16(3.67)	33.7(54.6)	9.93(5.23)	9.27(3.77)	16.2(7.63)	11.2(4.98)	9.39(4.80)
		2.0	3.89(1.18)	7.10(2.37)	4.76(1.60)	4.69(1.56)	7.17(2.25)	5.94(2.00)	5.40(1.65)
		4.0	2.54(0.68)	3.42(0.88)	3.17(0.91)	3.07(0.88)	3.89(1.09)	4.05(1.23)	3.91(1.22)
Gam_3	10	0.5	19.8(16.7)	39.8(58.5)	51.2(72.2)	18.7(10.9)	30.6(26.3)	36.7(37.2)	40.0(45.1)
		1.0	6.92(2.47)	10.6(4.78)	13.0(8.23)	8.03(2.80)	11.6(4.59)	12.7(6.27)	7.85(2.42)
		2.0	3.74(1.02)	5.08(1.47)	5.18(1.96)	4.60(1.37)	6.20(1.90)	5.97(2.16)	5.13(1.40)
		4.0	2.65(0.71)	3.49(0.90)	2.61(0.81)	3.30(0.94)	4.39(1.24)	3.15(1.02)	4.00(1.17)
	30	0.5	16.1(9.73)	21.6(24.4)	29.7(36.8)	15.3(6.71)	19.5(11.7)	23.0(14.8)	38.1(38.5)
		1.0	6.15(1.79)	7.53(3.06)	8.61(3.96)	7.03(2.10)	8.88(3.39)	9.63(3.85)	7.41(2.12)
		2.0	3.24(0.79)	4.07(1.25)	3.93(1.29)	3.94(1.03)	5.11(1.63)	4.83(1.60)	4.36(1.16)
		4.0	2.22(0.52)	2.92(0.82)	2.21(0.65)	2.73(0.69)	3.76(1.15)	2.81(0.88)	3.23(0.93)

Table S.8: OC ARL comparison of the DFEWMA, SREWMA, SSEWMA, and RTC charts when $m_0 = 100$ and shift direction is uniformly distributed on the p-sphere; numbers in parentheses are SDRL values.

			$\lambda = 0.1$			$\lambda = 0.05$			
	p	δ	DFEWMA	SREWMA	SSEWMA	DFEWMA	SREWMA	SSEWMA	RTC
MVN	10	0.5	134(157)	110(143)	115(142)	99.9(120)	88.9(113)	87.4(107)	100(111)
		1.0	33.5(53.5)	33.5(55.6)	34.4(52.5)	23.4(15.4)	26.0(23.0)	26.9(23.7)	26.2(23.9)
		2.0	8.19(2.88)	8.98(5.12)	8.78(5.00)	9.35(3.07)	10.3(4.65)	10.1(4.68)	7.75(2.55)
		4.0	4.42(1.15)	4.57(1.52)	3.91(1.47)	5.36(1.46)	5.69(1.88)	4.80(1.77)	5.05(1.37)
	30	0.5	155(176)	100(193)	97.4(126)	130(151)	62.1(103)	70.7(88.1)	140(161)
		1.0	60.1(89.3)	18.2(34.2)	22.3(25.7)	36.0(34.7)	17.1(9.98)	20.0(11.5)	51.5(51.8)
		2.0	10.6(4.24)	6.42(2.57)	6.82(2.70)	11.9(4.24)	7.94(3.02)	8.35(3.07)	10.7(5.53)
		4.0	4.84(1.24)	3.68(1.13)	3.29(1.02)	5.83(1.58)	4.74(1.53)	4.39(1.36)	5.63(1.48)
t_5	10	0.5	142 (165)	127(161)	164(165)	116(139)	104(137)	136(154)	117(143)
		1.0	47.9(68.7)	45.8(71.4)	110(140)	30.9(28.5)	33.6(38.8)	56.1(74.7)	36.8(37.4)
		2.0	9.68(3.91)	10.9(6.60)	22.6(34.9)	10.8(3.93)	11.8(5.81)	15.5(9.68)	8.87(3.53)
		4.0	4.96(1.39)	5.19(1.83)	6.20(2.56)	5.98(1.69)	6.29(2.21)	6.71(2.59)	5.59(1.55)
	30	0.5	171(187)	175(155)	176(255)	144(157)	153(151)	140(204)	158(183)
		1.0	91.2(123)	151(148)	87.7(161)	55.9(79.8)	93.7(112)	51.6(83.2)	81.7(100)
		2.0	14.1(7.38)	68.6(95.2)	15.0(11.5)	14.3(5.66)	24.6(18.1)	15.2(8.03)	15.5(11.5)
		4.0	5.66(1.58)	10.5(5.45)	6.04(2.35)	6.71(1.84)	9.41(3.33)	7.28(2.66)	6.55(1.71)
Gam_3	10	0.5	86.7(115)	149(179)	167(180)	63.4(79.4)	119(154)	139(158)	71.4(81.9)
		1.0	17.2(14.1)	62.5(98.4)	109(143)	17.1(9.01)	43.8(54.3)	59.1(76.7)	18.0(16.4)
		2.0	6.91(2.56)	13.4(9.28)	21.5(25.7)	8.05(2.87)	14.1(7.08)	16.7(9.17)	6.86(2.01)
		4.0	4.00(1.17)	5.84(2.08)	6.47(2.55)	5.02(1.47)	7.04(2.48)	7.24(2.76)	4.83(1.28)
	30	0.5	125(152)	198(278)	172(175)	92.8(116)	150(211)	157(169)	105(118)
		1.0	28.4(33.8)	119(206)	133(154)	22.3(13.1)	77.8(128)	95.9(118)	30.1(29.1)
		2.0	8.24(3.12)	26.8(75.5)	43.6(63.9)	9.59(3.37)	19.4(12.0)	25.7(18.3)	8.03(2.65)
		4.0	4.27(1.17)	7.45(3.19)	9.06(3.99)	5.27(1.47)	8.86(3.46)	9.97(3.67)	5.06(1.29)

Table S.9: OC ARL comparison of the DFEWMA, SREWMA, SSEWMA, and RTC charts when $m_0 = 100$ for mixed distribution. The first $p/2$ components follow t_1 distribution, and the last $p/2$ components follow ξ_1^2 distribution. Different mean shift scenarios are compared. Numbers in parentheses are SDRL values.

			$\lambda = 0.1$			$\lambda = 0.05$			
	p	δ	DFEWMA	SREWMA	SSEWMA	DFEWMA	SREWMA	SSEWMA	RTC
$p/5$ shifts	10	0.5	113.(159.)	185 (186)	192(170)	98.6(120.)	174 (172)	191(174)	101 (113)
		1.0	34.4(60.7)	176 (182)	191(170)	29.1(23.8)	159 (166)	191(175)	31.7(31.8)
		2.0	8.89(9.71)	137 (166)	191(172)	12.8(5.73)	107 (129)	187(175)	9.52(4.08)
		4.0	4.82(6.95)	68.3(112)	186(169)	8.40(2.86)	51.4(83.5)	173(167)	6.52(1.70)
	30	0.5	85.4(137.)	200 (221)	182(132)	62.9(83.8)	177 (201)	182(138)	64.6(68.0)
		1.0	14.1(14.7)	188 (217)	185(137)	17.3(8.57)	157 (189)	180(140)	15.3(11.1)
		2.0	5.37(6.71)	139 (192)	184(137)	8.67(3.05)	93.3(137)	175(141)	6.98(1.87)
		4.0	3.12(5.67)	48.3(106)	177(137)	5.95(1.76)	28.1(47.6)	154(135)	5.27(1.34)
all shifts	10	0.5	27.3(62.4)	44.7(57.9)	185(166)	17.7(15.6)	32.8(28.1)	177 (170)	20.1(20.7)
		1.0	6.88(7.87)	11.7(5.81)	179(166)	7.96(8.13)	12.4(5.06)	129 (143)	7.41(1.80)
		2.0	3.16(6.15)	5.62(1.67)	156(158)	4.25(6.35)	6.69(2.09)	26.3(26.9)	5.64(1.47)
		4.0	1.62(5.64)	3.86(1.08)	59.6(114)	2.44(5.78)	4.73(1.41)	8.88(2.34)	4.6 (1.21)
	30	0.5	11.0(9.86)	38.5(58.5)	177(132)	10.8(9.43)	26.2(19.2)	160(129)	10.7(5.46)
		1.0	4.46(6.15)	10.0(4.81)	169(133)	5.38(6.68)	10.8(4.45)	117(110)	6.22(1.46)
		2.0	2.17(5.22)	4.97(1.63)	149(129)	2.86(5.63)	6.00(2.03)	36.5(46.0)	4.57(1.16)
		4.0	1.16(4.89)	3.45(1.07)	122(123)	1.59(5.24)	4.29(1.42)	10.5(2.67)	3.59(0.94)
random shifts	10	0.5	40.3(78.4)	136(167)	189(170)	36.6(44.0)	100(124)	188(171)	23.4(41.8)
		1.0	11.2(13.6)	54.7(91.9)	188(165)	15.4(8.75)	36.7(45.5)	180(166)	9.21(5.83)
		2.0	4.96(7.34)	12.8(12.0)	179(163)	8.50(3.38)	13.2(9.21)	142(145)	6.24(1.72)
		4.0	2.60(6.05)	5.89(2.42)	157(157)	5.62(1.81)	6.95(2.63)	38.0(61.9)	4.98(1.38)
	30	0.5	51.3(85.8)	187(209)	183(131)	41.2(60.7)	149(175)	181(136)	24.2(25.2)
		1.0	12.5(16.5)	133(179)	183(134)	16.2(8.49)	81.6(113)	174(133)	9.56(5.79)
		2.0	5.09(7.50)	36.2(68.4)	179(134)	8.78(3.26)	22.3(15.8)	158(130)	6.22(1.69)
		4.0	2.46(6.20)	8.83(3.96)	166(131)	5.52(1.59)	9.78(4.08)	108(106)	4.80(1.21)