An estimator of the number of change points based on a weak invariance principle

Christoph Kühn[a]

[a] Center for Mathematical Sciences, Munich University of Technology, D-80290 Munich, Germany
e-mail: kuehn@mathematik.tu-muenchen.de, fax: +49 89 289 28464

Abstract

We study an estimator of the number of change points in the drift of a stochastic process based on the Schwarz criterion. In a general statistical model where the additive measurement noise satisfies a certain weak invariance principle (examples included are partial sums, renewal processes, and linear processes in time series analysis) consistency can be shown under the condition that the number of jumps is not greater than a given upper bound.

AMS subject classification: 62F12; 60F17.
Keywords: Change points; Change in the mean; Schwarz criterion; Invariance principle; Partial sum; Renewal process; Linear process.

1 Introduction

The aim of this paper is to discuss the problem of estimating the *number* of change points in a general context. It extends work of Yao (1988) and Lee (1995) concerning only sequences of independent normal random variables. In particular, we combine an auxiliary result of Yao (1988) for i.i.d.-normal random variables with an invariance principle for partial sums to improve a later result of Yao and Au (1989) where an i.i.d.-sequence with a more restrictive moment condition was considered.

As in many other asymptotic studies of change-point problems it turns out that certain invariance principles are very useful and allow for reducing the statistical analysis to that of an asymptotic Gaussian model. Recently, Horváth and Steinebach (1999) pursued this idea in a general vein by simply taking advantage of a weak invariance principle in the *testing* of a change in the mean or variance of a rather general stochastic process. In particular, cases of dependent observations are included in this framework.

Assume that we observe a stochastic sequence $(Z_n(i))_{i=1,\dots,n}$ having the following structure:

$$Z_n(i) := \begin{cases} a_1 i + b_1 Y_n^{(1)}(i) & : \quad 1 \le i \le n_1 \\ Z_n(n_{j-1}) + a_j(i - n_{j-1}) + b_j Y_n^{(j)}(i - n_{j-1}) & : \quad n_{j-1} < i \le n_j,\ j = 2, \dots, l+1 \end{cases} \tag{1.1}$$

where $a_j \neq a_{j-1}$, b_j and $n_j = [\theta_j n]$, $0 = \theta_0 < \theta_1 < \dots < \theta_l < \theta_{l+1} = 1$, are unknown parameters, and where $\left(Y_n^{(j)}(i)\right)_{i=1,\dots,n_j-n_{j-1}}$ are unobservable stochastic sequences satisfying a (uniform) weak invariance principle. Namely, we assume that, for every $n \ge 1$, there exist $l+1$ *independent* Wiener processes $\{W_n^{(j)}(t) : 0 \le t \le n_j - n_{j-1}\}$, $j = 1, \dots, l+1$, such that, if necessary after re-definition on a suitable probability space,

$$\max_{1\le i\le n_j-n_{j-1}} \left|Y_n^{(j)}(i) - W_n^{(j)}(i)\right| = \mathcal{O}_P(n^\alpha) \quad (n\to\infty) \tag{1.2}$$

with some $\alpha < 1/2$.

That means we have a piecewise linear function with an additive measurement noise. Various statistical models satisfy condition (1.2) above as will be demonstrated by a few examples:

Example 1.1 *(Partial sums) For every $n \ge 1$, let $\{X_{i,n}^{(j)} : i = 1, 2, \ldots\}$, $j = 1, \ldots, l+1$, be $l+1$ independent sequences of i.i.d. (independent, identically distributed) random variables with $P\{X_{i,n}^{(j)} \le x\} = P\{X_{1,1}^{(j)} \le x\}$ for all $x \in I\!R$, $n \ge 1$, $EX_{1,1}^{(j)} = \mu_j$, $Var(X_{1,1}^{(j)}) = \sigma_j^2 > 0$, being fixed. For some $0 = \theta_0 < \theta_1 < \ldots < \theta_{l+1} = 1$, consider $\{Z_n(i) : 1 \le i \le n\}$ as in (1.1) with $a_j = \mu_j$, $b_j = \sigma_j$, and*

$$Y_n^{(j)}(i) := \ \textstyle\sum_{k=1}^{i}(X_{k,n}^{(j)} - \mu_j)/\sigma_j \quad : \quad 1 \le i \le n_j - n_{j-1},\ j = 1, \ldots, l+1 .$$

If, in addition, $E|X_{1,1}^{(j)}|^{2+\Delta} < \infty$, for some $\Delta > 0$, then the Komlós, Major and Tusnády (1975) strong approximation implies (1.2) for $\{Y_n^{(j)}(i) : 1 \le i \le n_j - n_{j-1}\}$, with $\alpha = 1/(2+\Delta)$.

Example 1.2 *(Renewal Processes) Let the random variables $X_{i,n}^{(j)}$ be as in Example 1.1, but with $P(X_{1,1}^{(j)} \ge 0) = 1$. Set $a_j = 1/\mu_j$, $b_j = \sigma_j\ /\mu_j^{3/2}$,*

$$N_n^{(j)}(t) := \max\left\{ i \ge 0 : \sum_{1\le k\le i} X_{k,n}^{(j)} \le t \right\} , \quad 0 \le t < \infty ,$$

$$Y_n^{(j)}(t) := \frac{N_n^{(j)}(t) - t/\mu_j}{\sigma_j/\mu_j^{3/2}}, \quad 1 \le i \le n_j - n_{j-1},\ j = 1, \ldots, l+1 .$$

Then (1.2), with $\alpha = 1/(2+\Delta)$, follows directly from Example 1.1 in combination with Theorem 3.1 of Csörgő, Horváth and Steinebach (1987).

Motivated by a change-point analysis of time series, the following generalization of Example 1.1 is of particular interest. It also demonstrates that, although the increments of the observed process $\{Z_n(i) : 1 \le i \le n\}$ may have a special dependency structure before and after the change-points n_j, the approximating Wiener processes $\{W_n^{(j)}(t) : \ 0 \le t \le n_j - n_{j-1}\}$, j=1,...,l+1, can be chosen to be independent.

Example 1.3 *(Linear processes) Consider* $\{Z_n(i) : 1 \le i \le n\}$ *as in (1.1), but with* $b_j = \sigma_j A$, *and*

$$Y_n^{(j)}(i) := \ \sum_{k=n_{j-1}+1}^{n_{j-1}+i} e_{k,n}/(\sigma_j A) \quad : \quad 1 \le i \le n_j - n_{j-1} \ ,$$

where $\{a_u : u = 0, 1, \ldots\}$ *is a real sequence such that*

$$A = \sum_{u=0}^{\infty} a_u \neq 0 \ .$$

Let

$$e_{k,n} = \sum_{u=0}^{\infty} a_u \varepsilon_{k-u,n} \ ,$$

where $\{\varepsilon_{k,n} : k = 0, \pm 1, \pm 2, \ldots ; k \le n_1\}$ *and* $\{\varepsilon_{k,n} : k = 0, \ldots, n; n_{j-1} < k \le n_j\}$, *j=2,...,l+1 are* $l+1$ *independent sequences, each i.i.d. (white noise), with* $E\varepsilon_{k,n} = 0$,

$$Var(\varepsilon_{k,n}) = \begin{cases} \sigma_1^2 & : \quad k \le n_1, \\ \sigma_j^2 & : \quad n_{j-1} < k \le n_j, \ j = 2, \ldots, l+1, \end{cases}$$

$\sigma_j^2 > 0$, *and* $E|\varepsilon_{k,n}|^{2+\Delta} < \infty$ *(for some* $\Delta > 0$*) . We further assume that there exists an exponent* $\beta > 3/2$ *so that*

$$a_u = \mathcal{O}\left(u^{-\beta}\right) \quad (u \to \infty) \ . \tag{1.3}$$

If, in addition, the $\varepsilon_{k,n}$'s have smooth densities and $\{a_u : u = 0, 1, \ldots\}$ satisfies some regularity conditions, then the invariance principle (1.2), with $\alpha = 1/(2 + \Delta)$ follows from Horváth (1997). For details and exact conditions confer Lemmas 2.1 and 2.2 of Horváth (1997).

We need an additional technical condition: For all $j = 1, \ldots, l+1$ and $\varepsilon > 0$ there exist $\tilde{n}, C_0, C_1 > 0$ such that for all $n \ge \tilde{n}$:

$$P\left(C_0 n \le \sum_{i=1}^{n_j - n_{j-1}} \left(Y_n^{(j)}(i) - Y_n^{(j)}(i-1)\right)^2 \le C_1 n\right) \ge 1 - \varepsilon \tag{1.4}$$

This is obviously fulfilled for Example 1.1. For a verification concerning Example 1.2 and Example 1.3 confer Kühn (1999), pp. 52-55.

For any k, let $\hat{j}(k) = (\hat{j}_1, \ldots, \hat{j}_k)$ be the vector which minimizes

$$S\left(j_1, \ldots, j_k\right) = \sum_{i=0}^{k} \sum_{j=j_i+1}^{j_{i+1}} \left\{Z_n(j) - Z_n(j-1) - \bar{Z}_n(j_i, j_{i+1})\right\}^2 \tag{1.5}$$

subject to $j(k) := (j_1, \ldots, j_k),\ j_i \in \mathbb{N}_0,\ 0 = j_0 < j_1 < \ldots < j_k < j_{k+1} = n$, where

$$\bar{Z}_n(a,b) := \frac{1}{b-a}\left\{Z_n(b) - Z_n(a)\right\}.$$

We define

$$\hat{\sigma}_k^2 := \frac{S(\hat{j}(k))}{n} := \min_{j(k)} \frac{S(j(k))}{n}.$$

Using a criterion similar to that of Schwarz (1978), we estimate l, the unknown number of change points, by $\hat{l}$ which maximizes the function

$$SC(k) := -\frac{n}{2}\log\hat{\sigma}_k^2 - kd_n$$

subject to $k \leq l_u$. Here kd_n is a penalty term in order to avoid over-fitting.

Our main result is as follows:

Theorem 1.4 *Consider Z_n as in (1.1) satisfying (1.2), $0 < \alpha < 1/2$, and (1.4). Let l_u be a known upper bound for l, and assume the penalty factor d_n to increase such that*

$$\frac{d_n}{n^{2\alpha}} \to \infty \quad (n \to \infty), \tag{1.6}$$

but

$$\frac{d_n}{n} \to 0 \quad (n \to \infty). \tag{1.7}$$

Then we have

$$P\left\{\arg\max_{k=0,\ldots,l_u} SC(k) = l\right\} \to 1 \quad (n \to \infty).$$

Remark. In the case of i.i.d.-increments (Example 1.1) Theorem 1.4 improves Theorem 3 in Yao and Au (1989) in two ways. First, we only require a $(1/\alpha)$-th moment (α is an arbitrary real number $< 1/2$) for the increments (instead of the sixth moment). Second, the penalty factor d_n only has to increase faster than $n^{2\alpha}$ instead of $n^{4\alpha}$ which was the corresponding condition in Yao and Au (1989) translated into our notation.

To prove Theorem 1.4 we need three lemmas.

Lemma 1.5 *Let $d_0(n) := \min_{j=1,\ldots,l+1} |n_j - n_{j-1}|$ and $a < n_{j_1} < \ldots < n_{j_2} < b$ with $\min\{|a - n_j|, |b - n_j|\} \geq \frac{d_0(n)}{2}$ for j with $j_1 \leq j \leq j_2$, then there exist $c \in [n_{j_1}, n_{j_2}]$ and $\delta > 0$, only depending on the jumb hights, but not on n, such that*

$$|\mu_{a,b}(c)| \geq \frac{b-a}{(c-a)(b-c)} d_0(n)\delta,$$

where

$$\mu_{a,b}(c) := \frac{\zeta_n(c) - \zeta_n(a)}{c-a} - \frac{\zeta_n(b) - \zeta_n(c)}{b-c}$$

and, for $i = 1, \ldots, n$, $\zeta_n(i) := \sum_{j=1}^{l+1} a_j((i \wedge n_j) - n_{j-1})_+$ represents the "systematic part" of $Z_n(i)$.

The proof is exactly the same as for Lemma 2 in Lee (1995) because we only consider the "systematic part" ζ_n of Z_n (which we get by setting $Y_n^{(j)} \equiv 0$ in (1.1)).

Lemma 1.6 *Suppose that (1.2) with $\alpha \in (0, 1/2)$ and (1.4) hold. Then*

$$a) \quad \max_{0 \leq k_1 < k_2 \leq n,\ k_1, k_2 \in I\!N} \frac{\{Z_n(k_2) - \zeta_n(k_2) - (Z_n(k_1) - \zeta_n(k_1))\}^2}{k_2 - k_1} = \mathcal{O}_P\left(n^{2\alpha}\right). \tag{1.8}$$

b) It holds, uniformly over all $c_r, c_{r+1} \in \{0, \ldots, n\}$ with no change-point n_j in between, and all $c \in (c_r, c_{r+1})$:

$$S(c_1, \ldots, c_k) - S(c_1, \ldots, c_r, c, c_{r+1}, \ldots, c_k) = \mathcal{O}_P\left(n^{2\alpha}\right).$$

PROOF OF LEMMA 1.6: a): We want to derive the assertion from Lemma 1 in Yao (1988) concerning i.i.d.-*normal* r.v.'s by taking (1.2) into account. We have

$$\begin{aligned}
&\max_{0 \leq k_1 < k_2 \leq n,\ k_1, k_2 \in I\!N} \frac{\{Z_n(k_2) - \zeta_n(k_2) - (Z_n(k_1) - \zeta_n(k_1))\}^2}{k_2 - k_1} \\
&= \max_{0 \leq k_1 < k_2 \leq n,\ k_1, k_2 \in I\!N} \frac{\left\{\sum_{j=1}^{l+1} b_j \left\{Y_n^{(j)}((k_2 \wedge n_j - n_{j-1})_+) - Y_n^{(j)}((k_1 \wedge n_j - n_{j-1})_+)\right\}\right\}^2}{k_2 - k_1} \\
&= \max_{0 \leq k_1 < k_2 \leq n,\ k_1, k_2 \in I\!N} \frac{\left\{\sum_{j=1}^{l+1} b_j \left\{W_n^{(j)}((k_2 \wedge n_j - n_{j-1})_+) - W_n^{(j)}((k_1 \wedge n_j - n_{j-1})_+)\right\} + \mathcal{O}_P(n^\alpha)\right\}^2}{k_2 - k_1} \\
&= \max_{0 \leq k_1 < k_2 \leq n,\ k_1, k_2 \in I\!N} \frac{\left\{\sum_{j=1}^{l+1} b_j \left\{W_n^{(j)}((k_2 \wedge n_j - n_{j-1})_+) - W_n^{(j)}((k_1 \wedge n_j - n_{j-1})_+)\right\}\right\}^2}{k_2 - k_1} \\
&+ \max_{0 \leq k_1 < k_2 \leq n,\ k_1, k_2 \in I\!N} \frac{\left|\sum_{j=1}^{l+1} b_j \left\{W_n^{(j)}((k_2 \wedge n_j - n_{j-1})_+) - W_n^{(j)}((k_1 \wedge n_j - n_{j-1})_+)\right\}\right|}{k_2 - k_1} \times \mathcal{O}_P\left(n^\alpha\right) + \mathcal{O}_P\left(n^{2\alpha}\right) \\
&= \mathcal{O}_P\left(\log n\right) + \mathcal{O}_P\left(\sqrt{\log n}\ n^\alpha\right) + \mathcal{O}_P\left(n^{2\alpha}\right) \\
&= \mathcal{O}_P\left(n^{2\alpha}\right)
\end{aligned}$$

b) : We have

$$
\begin{aligned}
& S((c_1,\ldots,c_k)) - S((c_1,\ldots,c_r,c,c_{r+1},\ldots,c_k)) \\
&= \frac{(c-c_r)(c_{r+1}-c)}{c_{r+1}-c_r}\left(\bar{Z}_{c_r,c}-\bar{Z}_{c,c_{r+1}}\right)^2 \\
&= \frac{(c-c_r)(c_{r+1}-c)}{c_{r+1}-c_r}\left(\frac{Z_n(c)-\zeta_n(c)-(Z_n(c_r)-\zeta_n(c_r))}{c-c_r}\right. \\
&\qquad \left. - \frac{Z_n(c_{r+1})-\zeta_n(c_{r+1})-(Z_n(c)-\zeta_n(c))}{c_{r+1}-c}\right)^2 \\
&= \underbrace{\frac{(c-c_r)(c_{r+1}-c)}{c_{r+1}-c_r}\left(\frac{1}{c-c_r}+2\frac{1}{\sqrt{(c-c_r)(c_{r+1}-c)}}+\frac{1}{c_{r+1}-c}\right)}_{\leq 4}\times\mathcal{O}_P\left(n^{2\alpha}\right) \\
&= \mathcal{O}_P(n^{2\alpha})
\end{aligned}
$$

Lemma 1.7 *For all $k \in I\!N$, $\varepsilon > 0$, there exist $\tilde{n}$, C_0, $C_1 > 0$, such that*

$$
P\left\{C_0 \leq \frac{S(j(k))}{n} \leq C_1 \quad \forall j(k)\right\} \geq 1-\varepsilon \quad \forall n \geq \tilde{n}. \tag{1.9}
$$

PROOF OF LEMMA 1.7: Here, we need the additional assumption (1.4).

STEP 1 : („ $\leq$ ") :

$$
\begin{aligned}
S(j(k)) &\leq S(j(0)) \\
&= \sum_{i=1}^{n}\left\{Z_n(i)-Z_n(i-1)-\frac{Z_n(n)}{n}\right\}^2 \\
&= \sum_{j=1}^{l+1} b_j^2 \sum_{i=1}^{n_j-n_{j-1}}\left\{Y_n^{(j)}(i)-Y_n^{(j)}(i-1)-\frac{Y_n^{(j)}(n_j-n_{j-1})}{n_j-n_{j-1}}\right\}^2 \\
&\quad + \sum_{j=1}^{l+1}(n_j-n_{j-1})\left\{b_j\frac{Y_n^{(j)}(n_j-n_{j-1})}{n_j-n_{j-1}}+a_j-\frac{Z_n(n)}{n}\right\}^2 \\
&\leq \sum_{j=1}^{l+1} b_j^2 \underbrace{\sum_{i=1}^{n_j-n_{j-1}}\left\{Y_n^{(j)}(i)-Y_n^{(j)}(i-1)\right\}^2}_{=\mathcal{O}_P(n),\ \text{due to (1.4)}} \\
&\quad + \sum_{j=1}^{l+1}(T_j-T_{j-1})\underbrace{\left\{b_j\frac{Y_n^{(j)}(n_j-n_{j-1})}{n_j-n_{j-1}}+a_j-\frac{Z_n(n)}{n}\right\}^2}_{=\mathcal{O}_P(1)} \\
&= \mathcal{O}_P(n).
\end{aligned}
$$

STEP 2 : („ $\geq$ ") :

$$
\begin{aligned}
S(j(k)) &\geq S(j(k) \cup \{n_1, \ldots, n_l\}) \\
&= S((n_1, \ldots, n_l)) + \mathcal{O}_P\left(n^{2\alpha}\right) \\
&= \sum_{j=1}^{l+1} b_j^2 \sum_{i=1}^{n_j - n_{j-1}} \left\{ Y_n^{(j)}(i) - Y_n^{(j)}(i-1) - \frac{Y_n^{(j)}(n_j - n_{j-1})}{n_j - n_{j-1}} \right\}^2 + \mathcal{O}_P\left(n^{2\alpha}\right) \\
&= \sum_{j=1}^{l+1} b_j^2 \sum_{i=1}^{n_j - n_{j-1}} \left\{ Y_n^{(j)}(i) - Y_n^{(j)}(i-1) \right\}^2 + \sum_{j=1}^{l} b_i^2 \underbrace{\frac{\left\{Y_n^{(j)}\right\}^2 (n_j - n_{j-1})}{n_j - n_{j-1}}}_{=\mathcal{O}_P(1)} + \mathcal{O}_P\left(n^{2\alpha}\right) \\
&\geq C_0 n.
\end{aligned}
$$

The $\mathcal{O}$'s are uniformly valid for $j(k)$ (k fixed) and the last inequality is true with a probability smaller than a given $\varepsilon > 0$ for n big enough. The constant C_0 is derived from assumption (1.4) and depends on ε.

PROOF OF THEOREM 1.4 (Compare Lee (1995)): We want to show that $SC(k)$ strictly increases on $\{0, \ldots, l\}$ and strictly decreases on $\{l, \ldots, l_u\}$ with a probability tending to 1 if n tends to infinity. This implies that $SC(k)$ takes its maximum at l, i.e. $\hat{l} := \arg\max_{k=0,\ldots,l_u} SC(k)$ is a consistent estimator of l. We analyze the difference $SC(k+1) - SC(k)$ and get:

$$
\begin{aligned}
SC(k+1) - SC(k) &= \frac{n}{2}\left(\log \hat{\sigma}_k^2 - \log \hat{\sigma}_{k+1}^2\right) - d_n \\
&= \frac{n}{2} \log \frac{S(\hat{j}(k))}{S(\hat{j}(k+1))} - d_n \\
&= \frac{n}{2} \log \left(1 + \frac{S(\hat{j}(k)) - S(\hat{j}(k+1))}{S(\hat{j}(k+1))}\right) - d_n
\end{aligned}
$$

Due to Lemma 1.7 the denominator $S(\hat{j}(k+1))$ is of order n (uniformly in $k = 1, \ldots, l_u$). Since there are $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, such that *for all* $x \in I\!\!R_+$

$$
\min\{(1-\varepsilon_1)x, \varepsilon_2\} \leq \log(1+x) \leq x, \tag{1.10}
$$

it is sufficient to show:

1.) $$S(\hat{j}(k)) - S(\hat{j}(k+1)) \gg d_n \quad \text{for } k < l. \tag{1.11}$$

2.) $$S(\hat{j}(k)) - S(\hat{j}(k+1)) \ll d_n \quad \text{for } k \geq l. \tag{1.12}$$

where $\ll$ means that the quotient tends to zero if n tends to infinity.

CASE 1: $\mathbf{k < l}$

We start with an optimal solution $\hat{j}(k)$ in the minimization of (1.5) for k and add one more point to $\hat{j}(k)$. Due to the fact that $S(\hat{j}(k+1))$ is minimal for given $k+1$, the amount the least squares $S(\hat{j}(k))$ decrease is a lower bound for $S(\hat{j}(k)) - S(\hat{j}(k+1))$. Consequently, if it is possible to choose this additional point in such a way that the order of the decrease is bigger than d_n, then (1.11) is proved.

Let $\hat{j}(k) = (c_1, c_2, \ldots, c_k)$. Suppose one plugs in an additional point c between c_r and c_{r+1}. Then:

$$
\begin{aligned}
& S((c_1,\ldots,c_k)) - S((c_1,\ldots,c_r,c,c_{r+1},\ldots,c_k)) \\
&= \sum_{i=c_r+1}^{c_{r+1}} \left(Z_n(i) - Z_n(i-1) - \bar{Z}_{c_r,c_{r+1}}\right)^2 - \sum_{i=c_r+1}^{c} \left(Z_n(i) - Z_n(i-1) - \bar{Z}_{c_r,c}\right)^2 \\
&\quad - \sum_{i=c+1}^{c_{r+1}} \left(Z_n(i) - Z_n(i-1) - \bar{Z}_{c,c_{r+1}}\right)^2 \\
&= (c-c_r)\left(\bar{Z}_{c_r,c} - \bar{Z}_{c_r,c_{r+1}}\right)^2 + (c_{r+1}-c)\left(\bar{Z}_{c,c_{r+1}} - \bar{Z}_{c_r,c_{r+1}}\right)^2 \\
&= (c-c_r)\left(\frac{c_{r+1}-c}{c_{r+1}-c_r}\right)^2\left(\bar{Z}_{c_r,c} - \bar{Z}_{c,c_{r+1}}\right)^2 + (c_{r+1}-c)\left(\frac{c-c_r}{c_{r+1}-c_r}\right)^2\left(\bar{Z}_{c_r,c} - \bar{Z}_{c,c_{r+1}}\right)^2 \\
&= \frac{(c-c_r)(c_{r+1}-c)}{c_{r+1}-c_r}\left(\bar{Z}_{c_r,c} - \bar{Z}_{c,c_{r+1}}\right)^2 . \qquad (1.13)
\end{aligned}
$$

On account of $k < l$ at least one of the l *disjoint* intervals $\left(n_j - \frac{d_0(n)}{2}, n_j + \frac{d_0(n)}{2}\right)$, $j = 1,\ldots,l$, covers no element of $\{c_1,\ldots,c_k\}$. Consequently, there exists an $r \in \{1,\ldots,k-1\}$ such that $c_r < n_{i_1} < \ldots < n_{i_2} < c_{r+1}$ with $\min\{|c_r - n_i|, |c_{r+1} - n_i|\} \geq \frac{d_0(n)}{2}$ for i with $i_1 \leq i \leq i_2$. So, Lemma 1.5 can be applied and there is a c for which the following estimation of the decrease of S holds (here we also make use of (1.13) and Lemma 1.6(a) where the uniformity is needed because c_r and c_{r+1} are random):

$$
\begin{aligned}
& S((c_1,\ldots,c_k)) - S((c_1,\ldots,c_r,c,c_{r+1},\ldots,c_k)) \\
&= \frac{(c-c_r)(c_{r+1}-c)}{c_{r+1}-c_r}\left(\bar{Z}_{c_r,c} - \bar{Z}_{c,c_{r+1}} - \mu(c) + \mu(c)\right)^2 \\
&\geq \frac{(c-c_r)(c_{r+1}-c)}{c_{r+1}-c_r}|\mu(c)|\left(|\mu(c)| - 2\left|\bar{Z}_{c_r,c} - \bar{Z}_{c,c_{r+1}} - \mu(c)\right|\right) \\
&\geq d_0(n)\delta\left\{\left(\frac{1}{c-c_r} + \frac{1}{c_{r+1}-c}\right)d_0(n)\delta - 2\left|\bar{Z}_{c_r,c} - \bar{Z}_{c,c_{r+1}} - \mu(c)\right|\right\} \\
&= d_0(n)\delta\left\{\left(\frac{1}{c-c_r} + \frac{1}{c_{r+1}-c}\right)d_0(n)\delta - \left(\frac{1}{\sqrt{c-c_r}} + \frac{1}{\sqrt{c_{r+1}-c}}\right)\mathcal{O}_P(n^\alpha)\right\} \\
&\geq \frac{d_0(n)\delta}{\sqrt{c-c_r}}\left\{\frac{d_0(n)\delta}{\sqrt{n}} + \mathcal{O}_P(n^\alpha)\right\} + \frac{d_0(n)\delta}{\sqrt{c_{r+1}-c}}\left\{\frac{d_0(n)\delta}{\sqrt{n}} + \mathcal{O}_P(n^\alpha)\right\}
\end{aligned}
$$

$$
\begin{aligned}
&\geq 2\frac{d_0^2(n)\delta^2}{n}\left\{1+\mathcal{O}_P\left(\frac{n^{\alpha+1/2}}{d_0(n)}\right)\right\} \\
&= 2\frac{d_0^2(n)\delta^2}{n}\{1+o_P(1)\} \\
&\gg d_n.
\end{aligned}
$$

The last estimation holds because of $d_0(n) \propto n$ together with (1.7). This proves (1.11).

CASE 2: $\mathbf{k \geq l}$

Now we start with an optimal solution $\hat{j}(k+1)$ for given $k+1$ and reduce the number of points to k. It turns out that this does not change too much. First we add the l (theoretical) change points to be able to apply Lemma 1.6(b) and then we take out $l+1$ other points. So,

$$
\begin{aligned}
S(\hat{j}(k+1)) &\geq S(\hat{j}(k+1)\cup\{n_1,\ldots,n_l\}) \\
&= S(n_1,\ldots,n_l)+\mathcal{O}_P\left(n^{2\alpha}\right) \\
&\geq S(\hat{j}(k))+\mathcal{O}_P\left(n^{2\alpha}\right)
\end{aligned}
$$

Thus

$$
S(\hat{j}(k))-S(\hat{j}(k+1))=\mathcal{O}_P\left(n^{2\alpha}\right)
$$

and therefore (1.12) in view of assumption (1.6). This completes the proof of Theorem 1.4.

Acknowledgment

I want to thank Josef Steinebach, the supervisor of my diploma thesis, for many helpful comments.

References

Csörgő, M., Horváth, L., Steinebach, J. (1987) Invariance principles for renewal processes. *Ann. Probab.* **15** 1441-1460.

Horváth, L. (1997) Detection of changes in linear sequences. *Ann. Inst. Statist. Math.* **49** 271-283.

Horváth, L., Steinebach, J. (1999) Testing for changes in the mean or variance of a stochastic process under weak invariance. *J. Statist. Plann. Infer.* (to appear)

Komlós, J., Major, P., Tusnády, G. (1975) An approximation of partial sums of independent R.V.'s and the sample DF. *I.Z. Wahrsch. Verw. Gebiete* **32** 111-131.

Kühn, C. (1999) Statistische Analyse von ,,Changeparametern" auf der Basis von Invarianzprinzipien. Diplomarbeit, Philipps-Universität Marburg.

Lee, C.-B. (1995) Estimating the number of change points in a sequence of independent normal random variables. *Statist. Probab. Letters* **25** 241-248.

Schwarz, G. (1978) Estimating the dimension of a model. *Ann. Statist.* **6** 461-464.

Yao, Y.-C. (1988) Estimating the number of change-points via Schwarz' criterion. *Statist. Probab. Letters* **6** 181-189.

Yao, Y.-C., Au, S.T. (1989) Least-squares estimation of a step function. *Sankhyā* **51** 370-381.