IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 24, 2017, accepted November 30, 2017, date of publication December 14, 2017,
date of current version February 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2782158

Accurate Jitter Computation in CNA Breakpoints Using Hybrid Confidence Masks With Applications to SNP Array Probing

JORGE MUNOZ-MINJARES[1], (Member, IEEE), YURIY S. SHMALIY[1], (Fellow, IEEE), LUIS J. MORALES-MENDOZA[2], (Member, IEEE), MIGUEL VAZQUEZ-OLGUIN[1], (Member, IEEE), AND CARLOS LASTRE-DOMINGUEZ[1], (Member, IEEE)

[1]Department of Electronics Engineering, Universidad de Guanajuato, Salamanca 36855, Mexico
[2]Electronics Department, Universidad Veracruzana, Xalapa 93270, Mexico

Corresponding author: Yuriy S. Shmaliy (shmaliy@ugto.mx)

ABSTRACT Chromosomal structural changes known as copy number alterations–aberrations (CNAs) result in gains or losses in copies of deoxyribonucleic acid sections, which are typically associated with different types of cancer. An intensive noise inherent to modern technologies of CNAs probing often causes inconsistency between the estimates provided by different methods. Therefore, testing estimates by the confidence masks is recommended to guarantee an existence of genomic changes within certain regions. In known masks, jitter in the CNA's breakpoints is expected to be distributed with the skew Laplace law, which is sufficiently accurate when the segmental signal-to-noise ratio (SNR) exceeds unity. In this paper, we extend the confidence masks to low and very low SNRs often observed in subtle chromosomal changes. The modified masks employ several proposed approximations of the segmental noise variance as a function of the departure step from the candidate breakpoint. Because approximations are accurate in jitter computation only for specified SNR regions, we suggest using hybrid masks to achieve the maximum available accuracy. Confidence masks are tested experimentally by genome CNA profile data obtained using the single nucleotide polymorphism array.

INDEX TERMS Genome, copy number alterations, breakpoints, jitter distribution, confidence masks.

I. INTRODUCTION

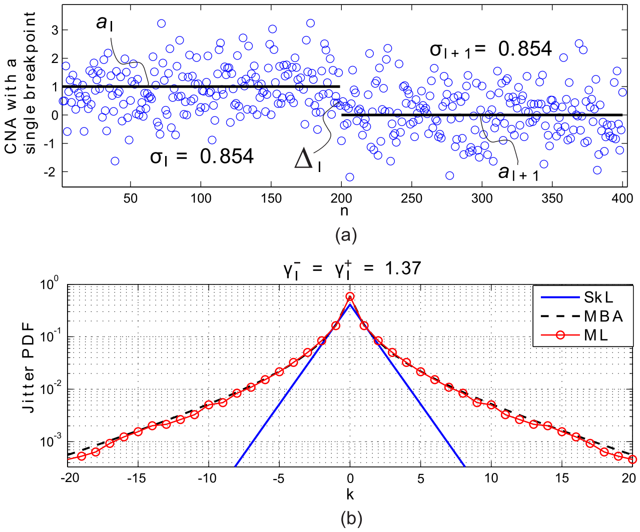
It is known that detection of structural aberrations called *copy number alterations* (CNAs) may be used to help diagnose a genetic disorder [1], [2] in the deoxyribonucleic acid (DNA) of a genome present in all forms of life [3], [4]. High-resolution techniques called next generation sequencing (NGS) have been developed to obtain the profiles of genetic structures. The NGS approach has generated an extensive development of the CNAs detection methods [7]–[9] at a resolution of 0.8–6 kb, which were recently reviewed in [10]. The single nucleotide polymorphism (SNP) arrays are nowadays one of the most efficient technologies for the CNAs identification [11]. Nevertheless, the CNAs data obtained using the SNP array and other technologies are still affected by several factors: 1) nature of biological material (tumor is contaminated by normal tissue, relative values and unknown baseline for copy number estimation), 2) technological biases (quality of material and hybridization/sequencing), and 3) intensive random noise [5], [6]. Moreover, modern technologies do not allow for multiple probing of the same chromosome that makes statistical simulation the only tool to determine confidence limits for the CNAs estimates.

A simulated measurement of the CNAs with one breakpoint and two constant segments is exampled in Fig. 1a. Here, the a_l and a_{l+1} segmental levels are contaminated with zero mean white Gaussian noise (WGN) [14], [15] having the variances σ_l^2 and σ_{l+1}^2. The segmental signal-to-noise ratios (SNRs) in the lth and $(l+1)$th segments are specified as in [13], respectively,

$$\gamma_l^- = \frac{\Delta_l^2}{\sigma_l^2}, \quad \gamma_l^+ = \frac{\Delta_l^2}{\sigma_{l+1}^2}, \tag{1}$$

where $\Delta_l = a_{l+1} - a_l$ is the segmental difference, which corresponds to the breakpoint i_l at $n = 200$. The factors described above do not grant a precise detection of i_l.

Such a phenomenon is called "jitter" [16], which denotes a deviation from the true breakpoint location and causes the detection uncertainty. In fact, the segmental levels are most accurately estimated by simple averaging between the breakpoints, which reduces the variance of the segmental noise by the number of the segmental probes that can be small for short segments. On the other hand, in spite of a certain progress in developing methods to refine the breakpoints [17]–[19], detecting the true breakpoint location is often difficult due to high segmental variances. In view of that, the problem of denoising while preserving edges in stepwise signals and thereby estimate the CNAs with highest precision has been extensively studied during decades [20]–[23].



(a)

(b)

FIGURE 1. Simulated CNAs with a single breakpoint at $n = 200$ and segmental noise standard deviations $\sigma_l = \sigma_{l+1} = 0.854$ corresponding to segmental SNRs of $\gamma_l^- \approx \gamma_l^+ = 1.37$: (a) CNA profile and (b) jitter distributions. Jitter histogram (circled) is obtained experimentally in the ML sense over 50 × 10^3 runs, SkL (solid) is the skew Laplace density (2), and MBA (dashed) is the modified Bessel function-based approximation proposed in [6].

It has been shown in [25] that the skew Laplace distribution depicted in Fig. 1b as SkL can approximate the jitter distribution in the breakpoints [24]. It has also been shown [25] that the SkL allows for a sufficient accuracy only when the SNR values exceed unity. Otherwise, if the SNR is low, γ_l^-, $\gamma_l^+ < 1$, and when it is extremely low, $\gamma_l^-, \gamma_l^+ \ll 1$, the SkL is inaccurate and improvements are required. That can be seen in Fig. 1b, where the experimentally measured distribution (circles) goes away from the SkL as the departure of $|k|$ from the breakpoint location ($k = 0$) increases. That means that for low SNRs the SkL-based confidence masks will be insufficiently accurate and the actual breakpoint may appear beyond the region predicted for the given probability. In Fig. 1b, we show an approximation MBA derived in [6] based on the modified Bessel function as a preliminary improvement to the confidence masks. Note that data in Fig. 1a were simulated with no relation to the actual genomic position, i.e., the step k in not equivalent to the number of the basepairs.

In this paper, we propose and investigate several other approximations of the jitter distribution in the CNA's breakpoints for low and extra low SNRs. We use these approximations to specify the lower bound (LB) and upper bound (UB) confidence masks and suggest using the hybrid masks for predicting possible breakpoint locations and segmental levels with a highest precision. All confidence masks are experimentally tested by the SNP array data. The rest of the paper is organized as follows. In Section II, we discuss the jitter in the CNA breakpoints and errors in the jitter confidence limits caused by the SkL. A parametrization of the SkL distribution by the k-varying segmental noise variances is provided in Section III using several approximations. Experimental testing of the modified confidence UB and LB masks by the SNP array probing is provided in Section IV and conclusions can be found in Section V.

II. JITTER REPRESENTATION IN THE BREAKPOINTS

In view of large probe noise, jitter typically exists in all CNA breakpoints. When $(\gamma_l^-, \gamma_l^+) > 1$, the jitter can be small and its distribution approximated with the SkL [30]. If $(\gamma_l^-, \gamma_l^+) < 1$, an actual breakpoint may appear several points apart from the candidate one detected by an estimator. Subtle chromosomal changes are often observed with $(\gamma_l^-, \gamma_l^+) \ll 1$ and, for the required high confidence probability, the actual breakpoint can be found tens of points apart to the left or to the right from the candidate one. In the latter case, the SkL becomes highly inefficient.

The following conjectures were made in [26] to arrive at the jitter distribution. Assume that a set of probes to the left of the breakpoint i_l in Fig. 1a belongs to segment a_l (event A_l) and a set of probes to the right of the breakpoint belongs to segment a_{l+1} (event B_l). The *jitter probability* is defined as the probability that one or more probes belong to another segment or event. It has been demonstrated in [26] that, under such a supposition, the jitter probability in the CNA breakpoints measured in WGN can be approximated with the discrete SkL probability density function (pdf) [27],

$$p(k|d_l, q_l) = \frac{(1-d_l)(1-q_l)}{1-d_l q_l} \begin{cases} d_l^k, & k \geqslant 0, \\ q_l^{|k|}, & k \leqslant 0, \end{cases} \tag{2}$$

where $0 < d_l = e^{-\frac{\kappa_l}{\nu_l}} = P(B_l)^{-1} - 1 < 1, 0 < q_l = e^{-\frac{1}{\kappa_l \nu_l}} = P(A_l)^{-1} - 1 < 1$, $\kappa_l = \sqrt{\frac{\ln x_l}{\ln(x_l/\mu_l)}}$, $\nu_l = -\frac{\kappa_l}{\ln x_l}$, and

$$x_l = \frac{\phi_l(1+\mu_l)}{2(1+\phi_l)}\left(1 - \sqrt{1 + \frac{4\mu_l(1-\phi_l^2)}{\phi_l^2(1+\mu_l)^2}}\right), \tag{3}$$

$$\mu_l = \frac{P(A_l)[1-P(B_l)]}{P(B_l)[1-P(A_l)]}, \tag{4}$$

$$\phi_l = \frac{P(A_l)+P(B_l)-1}{[1-2P(A_l)][1-2P(B_l)]}, \tag{5}$$

where $P(A_l)$ is the probability of event A_l and $P(B_l)$ is the probability of event B_l.

For our purpose, we now discuss Fig. 1 in more detail. The breakpoint i_l and segments a_l and a_{l-1} are detected here

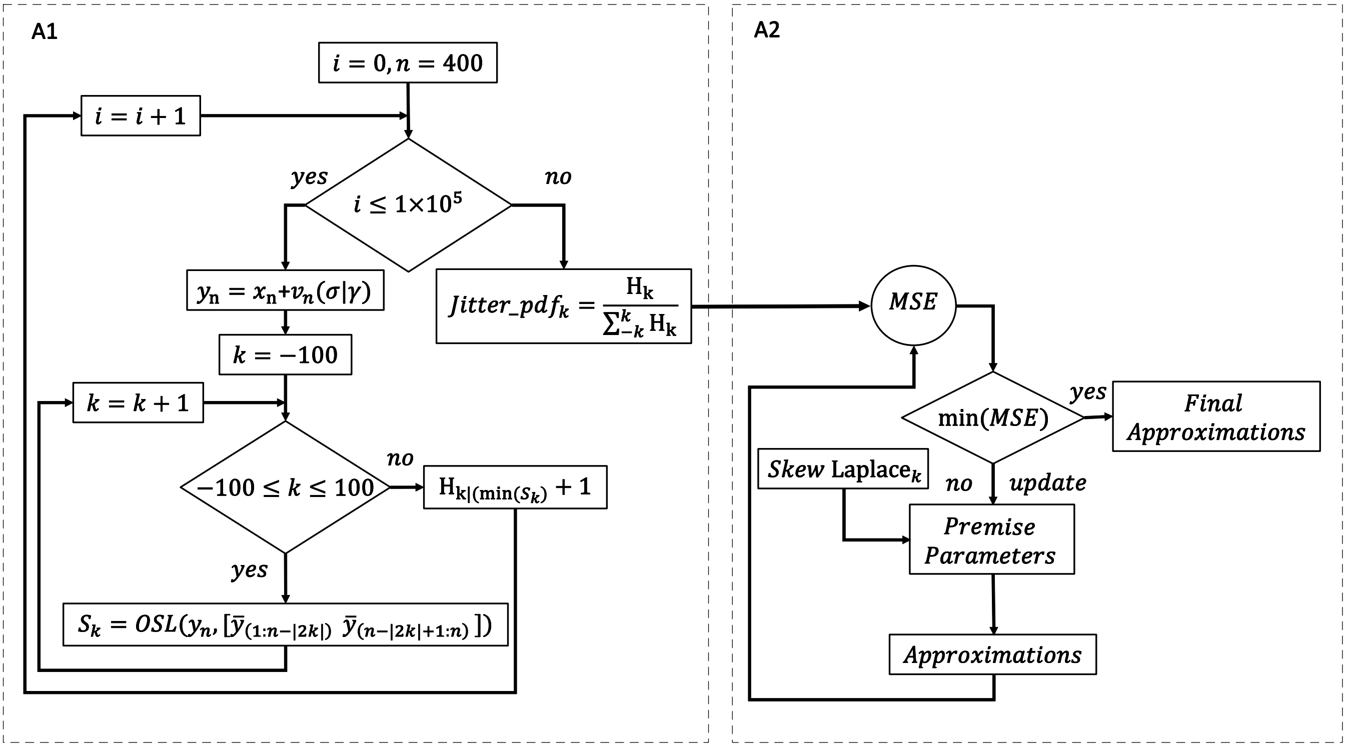


FIGURE 2. A flowchart to approximate the jitter distribution in the CNA breakpoints by simulating a stepwise signal in the presence of WGN with different segmentsl SNRs: A1 provides the jitter histogram and A2 provides the jitter distribution approximation by minimizing the MSE. An example of signal y_n is given in Fig. 1.

using the maximum likelihood (ML) estimator based on the ordinary least squares (OSL). The mean square error (MSE) produced by the ML estimator is minimized for the stepwise signal, in which a_l, a_{l-1}, and i_l are used as variables. The breakpoint location is detected when the MSE in the ML estimate reaches a minimum. In our simulation, the ML estimates were repeated 50×10^3 times for different noise realizations with constant SNR and then averaged. For each SNR value, a histogram was plotted as a number of the events in the k scale. To smooth ripples, such a procedure was repeated 9 times and the estimates were averaged. The histogram obtained in such a way was further normalized and accepted as an experimental *jitter pdf*, which is depicted in Fig. 1b with circles.

A. ERRORS OF SkL-BASED APPROXIMATION

An extensive analysis of the SkL pdf (2) in applications to jitter in the CNA-like signals measured in WGN has allowed making the following statements [26]. The SkL-based approximation (2) is:

- Acceptable when $\gamma_l^-, \gamma_l^+ > 1$ and very accurate if $\gamma_l^-, \gamma_l^+ \gg 1$;
- Also acceptable if at least one of the SNRs exceeds unity, $\gamma_l^- > 1$ or $\gamma_l^+ > 1$, and very accurate if $\gamma_l^- \gg 1$ or $\gamma_l^- \gg 1$;
- Inaccurate when $\gamma_l^-, \gamma_l^+ < 1$ and unacceptable if $\gamma_l^-, \gamma_l^+ \ll 1$.

An overall conclusion that can be made following [13], [25] is that the SkL-based approximation (2) fits only easily seen breakpoints. If chromosomal changes are not brightly pronounced, the SkL should not be used to make decisions about the CNAs structures [28], [30]. Therefore more accurate approximations are required, which will be discuss next.

III. PARAMETRIZATION OF LAPLACE DENSITY

The SkL pdf (2) still can be applied in a parameterized form as follows. An increase in the discrete-step index k diminishes the effect of the segmental noise on jitter in the breakpoint. For example, noise at $l - 10$ has a smaller effect on i_l that noise at $l - 1$. To provide the same effect of noise at any point $l \pm k$ on i_l as required by the derivation of the SkL-based approximation [26], the noise variances must be increased with k. That makes the variances, $\sigma_l^2(k)$ and $\sigma_{l+1}^2(k)$, k-variant and the SkL pdf (2) parameterized with k. Because exact analytical functions are unavailable for $\sigma_l^2(k)$ and $\sigma_{l+1}^2(k)$, in this paper we investigate them numerically and find reasonable approximations in the minimum MSE sense based on simulations.

To this end, we redefine the k-varying segmental SNRs as

$$\gamma_l^-(k) = \frac{\Delta^2}{\sigma_l^2(k)}, \quad \gamma_l^+(k) = \frac{\Delta^2}{\sigma_{l+1}^2(k)}, \tag{6}$$

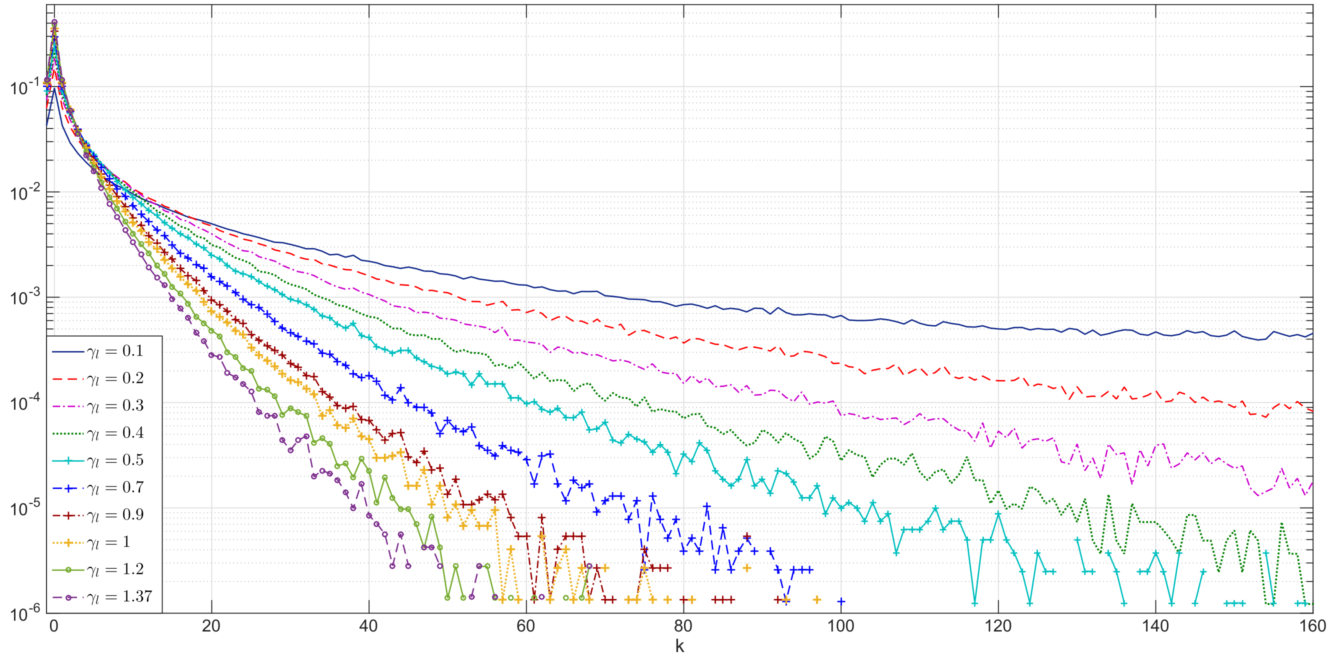


FIGURE 3. **The one-sided jitter probability densities obtained by simulations using sub-diagram A1 in Fig. 3 for identical segmental SNRs in the region of 0.1 $\leqslant \gamma_l^- = \gamma_l^+ \leqslant$ 1.37. A chromosomal segment is generated at $M = 400$ points with a single breakpoint at $n = 200$. The simulated density functions are found using a ML estimator. The jitter histogram is build over 1 × 10^5 runs repeated 10 times and then averaged.**

where $\sigma_l^2 \triangleq \sigma_l^2(0)$, $\sigma_{l+1}^2 \triangleq \sigma_{l+1}^2(0)$, $\gamma_l^- \triangleq \gamma_l^-(0)$, and $\gamma_l^+ \triangleq \gamma_l^+(0)$. Otherwise, when $k \neq 0$, we assign

$$\sigma_l^2(k) = \sigma_l^2\,[1 + f_l(k)],$$

where $f_l(k)$ is a function to be specified later.

A. SIMULATION OF JITTER DISTRIBUTION

Modern technologies do not allow for multiple probing of the same chromosome and simulation remains the only way to investigate jitter in the CNA breakpoints. The first relevant results were presented in [6]. Here, the MATLAB-based algorithm [6] was run using a computer based on Intel Core i5, 2.5 GHz. The computation time required to produce a histogram was about 12.7 hours. To make it possible to operate faster, in this paper we removed ‘‘for’’ cycles and did not save variables in RAM memory. Thereby, the computation time was significantly reduced and the jitter histogram computed with a higher accuracy in a wide range of k.

The modified algorithm is shown in Fig. 2. Its left part (Fig. 2,A1) allows getting the jitter histogram. Here, x_n is an idealized CNA signal with a single breakpoint at $n = 200$, v_n is a vector of WGN with the variance σ^2 corresponding to the given γ, and y_n is the CNA probe. To find the ML estimate using OLS, the breakpoint location has been changed with respect to its actual position at $n = 200$ on $-100 \leqslant k \leqslant 100$ points. The output is taken when the likelihood reaches a maximum. Then the jitter histogram H_k is computed and normalized to produce the discrete jitter pdf depicted as *jitter_pdf*$_k$. To reduce errors, H_k was created over 10^5 times repeated measurements. The simulated one-sided jitter distributions provided by the sub-algorithm A1 for equal segmental values of SNR are shown in Fig. 3.

TABLE 1. **Computation Time, in sec, Consumed by Algorithm [6] (A0) and Proposed SkL-Based Algorithm (A1) Parameterized with (8), (9), and (10).**

pdf	A0	A1	A0/A1
(2)	45605	1629	27.99
(2) with (8)	45600	1632	27.94
(2) with (9)	45601	1625	28.06
(2) with (10)	45609	1630	27.98
Average time	45604	1629	27.99

Referring to the necessity of estimating the CNAs with low segmental SNRs [5] and taking into account that the Laplace distribution (2) is sufficiently accurate when the SNR values exceed unity [6], we next investigate jitter in the region of $0.1 \leqslant \gamma_l^- = \gamma_l^+ \leqslant 1.37$. As can be seen in Fig. 4, a decrease in the SNRs makes the actual jitter distribution less straight in the logarithmic scale and the SkL has thus limited applications for low segmental SNRs.

1) COMPUTATIONAL COMPLEXITY

In view of massive data and the necessity to average the estimates over a big number of runs, the algorithm A0 designed in [6] consumes large time. The modified algorithm A1 developed in this paper and shown in Fig. 2 is computationally

TABLE 2. Typical MSEs produced by the algorithm designed in [6] and proposed algorithm given in Fig. 3. Both algorithms are exploited using the SkL density (2) and (2) parameterized with (8), (9) and (10).

γ	Algorithm [6]				Modified Algorithm			
	pdf (2)	(2) with (8)	(2) with (9)	(2) with (10)	pdf (2)	(2) with (8)	(2) with (9)	(2) with (10)
0.1	$7.6e^{-5}$	$2.7e^{-6}$	$2.5e^{-6}$	$3.6e^{-7}$	$8.6e^{-5}$	$1.4e^{-6}$	$1.5e^{-6}$	$1.8e^{-7}$
0.2	$7.7e^{-5}$	$6.7e^{-6}$	$6.9e^{-6}$	$1.7e^{-7}$	$7.8e^{-5}$	$4.1e^{-6}$	$3.8e^{-6}$	$1.1e^{-7}$
0.3	$7.5e^{-5}$	$9.7e^{-6}$	$1.0e^{-5}$	$1.6e^{-7}$	$7.4e^{-5}$	$5.6e^{-6}$	$5.3e^{-6}$	$7.9e^{-8}$
0.4	$7.3e^{-5}$	$1.1e^{-5}$	$1.3e^{-5}$	$1.6e^{-5}$	$7.0e^{-5}$	$6.9e^{-6}$	$6.3e^{-6}$	$8.5e^{-8}$
0.5	$6.6e^{-5}$	$1.4e^{-5}$	$1.6e^{-5}$	$1.7e^{-7}$	$6.6e^{-5}$	$8.1e^{-6}$	$7.3e^{-6}$	$1.1e^{-7}$
0.7	$5.9e^{-5}$	$1.6e^{-5}$	$1.9e^{-5}$	$2.3e^{-7}$	$5.9e^{-5}$	$1.0e^{-5}$	$9.0e^{-6}$	$1.2e^{-7}$
0.9	$5.3e^{-5}$	$1.9e^{-5}$	$2.3\ e^{-5}$	$2.0e^{-7}$	$5.3e^{-5}$	$1.2e^{-5}$	$1.0e^{-5}$	$9.2e^{-8}$
1.0	$5.1e^{-5}$	$2.1e^{-5}$	$2.5e^{-5}$	$2.2e^{-7}$	$5.0e^{-5}$	$2.2e^{-5}$	$2.7e^{-5}$	$1.7e^{-7}$
1.2	$4.8e^{-5}$	$2.3e^{-5}$	$2.7e^{-5}$	$2.5e^{-7}$	$4.4e^{-5}$	$2.5e^{-5}$	$3.2e^{-5}$	$1.7e^{-7}$
1.37	$4.1e^{-5}$	$2.7e^{-5}$	$3.4e^{-5}$	$3.4e^{-7}$	$4.0e^{-5}$	$2.7e^{-5}$	$3.4e^{-5}$	$2.6e^{-7}$

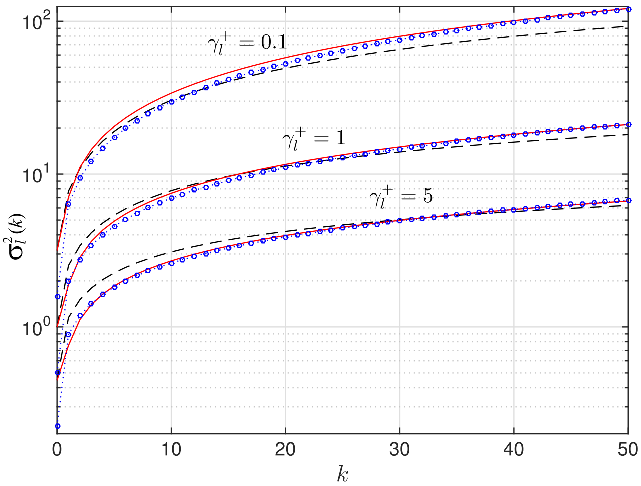


FIGURE 4. The proposed k-varying variance functions $\sigma_l^2(k)$ used to parameterize the SkL pdf (2) for equal low ($\gamma = 0.1$), normal ($\gamma = 1$), and large ($\gamma = 5$) SNRs: (8) is dashed, (9) is solid, and (10) is circled and dotted.

much more efficient that is illustrated in Table 1. In average, the algorithm A1 operates about 28 times faster than A0 and requires about 27 min to produce one histogram in Fig. 4.

B. JITTER DISTRIBUTION APPROXIMATION

In this section, we provide three efficient approximations of the jitter distribution in the CNA breakpoints based on the modified Bessel function and a power function.

1) FIRST BESSEL-BASED APPROXIMATION

Testing several non–conventional functions has revealed that the modified Bessel equation $K_\nu(x)$ of the second kind and fractional order $\nu = 0.5$ is a good candidate to approximate the measured jitter histogram, because it is positive-valued for $x(k) > 0$, smooth, and decreases with x to zero. We use

TABLE 3. SNR regions for MBA [6], Laplace pdf (2), and (2) parameterized with (8), (9), and (10) to detect the right jitter k^- and the left jitter k^+ in the ML sense.

γ_l^-, γ_l^+	k^-, k^+	pdf	SNR region		
			0.1...0.9	0.9...1.37	> 1.37
$=$	Any	(2) with (8)	–	X	–
		(2) with (9)	X	–	–
	> 1	(2) with (10)	X	X	–
$\neq$	Any	(MBA) [6]	X	X	–
	Any	(2)	–	–	X

the following representation of $K_\nu(x)$,

$$K_\nu[x(k)] = \int_0^\infty \cos[x(k)\sinh t]\,dt = \int_0^\infty \frac{\cos[x(k)t]}{\sqrt{t^2+1}}\,dt > 0, \quad x(k) > 0, \tag{7}$$

where $x(k)$ is k-varying.

Based on simulations, it has been found that the following parameterizing function makes the SkL pdf (2) accurate in fitting the jitter histogram for any k,

$$\sigma_l^2(k) = \sigma_l^2\left[1 + K_{1/2}^{-1}\left(\log_{k+1}^{a(\gamma_l^-)^b}\right)\right], \tag{8}$$

if to assign $a = 0.6951$ and $b = -0.1296$. In fact, $k = 0$ turns the parameterized SkL to (2) and, by $\gamma_l^-, \gamma_l^+ \gg 1$, it also converges to (2). An important property of the SkL parameterized with (8) is that it shows that when $\gamma_l^+ \to 0$ and $\gamma_l^- \to 0$ then $\sigma_l^2(k) \longrightarrow \infty$ and $\sigma_{l+1}^2(k) \longrightarrow \infty$ and i_l thus cannot be localized or, most likely, does not exists.

TABLE 4. Left jitter k_l^- and right jitter k_l^+ detected by different masks in the CNA breakpoints of the 13th chromosomal sample "BLC_B1_T37.txt" in the 3σ sense with the confidence probability of $P = 99.73\%$. The chromosome is associated with breast cancer and all values are given for Log$_2$ Ratio.

l	SNR			MBA [6]		Laplace (2)		(2) with (8)		(2) with (9)		(2) with (10)		**Hybrid**	
	γ_l^-	γ_l^+	$\lvert\gamma_l^- - \gamma_l^+\rvert$	k_l^-	k_l^+	k_l^-	k_l^+	k_l^-	k_l^+	k_l^-	k_l^+	k_l^-	k_l^+	k_l^-	k_l^+
1	0.2101	0.3052	0.0950	10	12	–	–	–	–	–	–	–	–	**10**	**12**
2	0.4722	0.5412	0.0690	7	8	10	6	–	–	–	–	–	–	**7**	**8**
3	3.7920	3.9122	0.1201	2	2	3	3	3	3	3	3	2	2	**3**	**3**
4	0.0574	0.0536	0.0037	–	–	–	–	–	–	–	–	–	–	–	–
5	0.0221	0.0232	0.0011	–	–	–	–	–	–	–	–	–	–	–	–
6	13.6833	13.4657	0.2175	–	–	1	1	1	1	1	1	1	1	**1**	**1**
7	0.8156	0.8993	0.0837	6	6	7	5	10	8	10	8	6	5	**10**	**8**
8	8.7577	8.9276	0.1698	–	–	2	2	2	2	2	2	1	1	**2**	**2**
9	7.1710	5.0402	2.1308	–	–	2	2	2	3	2	3	1	1	**2**	**2**
10	1.1958	2.3267	1.1309	4	4	7	3	–	–	–	–	–	–	**7**	**3**
11	1.6459	1.1516	0.4942	4	4	4	7	–	–	–	–	–	–	**4**	**7**
12	0.3676	0.4120	0.0444	9	9	12	6	–	–	–	–	–	–	**9**	**9**
13	5.9443	4.8291	1.1151	2	2	2	3	3	3	3	3	2	2	**3**	**3**
14	5.6081	6.5830	0.9749	–	–	2	2	2	2	2	2	1	1	**2**	**2**
15	5.8848	3.9626	1.9222	2	2	2	3	3	3	3	3	2	2	**2**	**3**
16	3.5660	4.9407	1.3746	2	2	3	2	3	3	3	3	2	2	**3**	**2**
17	4.5575	4.2765	0.2810	2	2	3	3	3	3	3	3	2	2	**3**	**3**
18	7.5862	29.3433	21.7570	–	–	1	1	1	1	1	1	1	1	**1**	**1**
19	32.2283	8.9617	23.2666	–	–	1	1	1	1	1	1	0	1	**1**	**1**
20	5.4114	5.0298	0.3816	–	–	2	2	3	3	3	3	2	2	**2**	**2**
21	1.5649	1.4423	0.12253	4	4	4	5	6	6	6	7	4	4	**4**	**5**
22	1.5162	1.6035	0.0873	4	4	5	4	6	6	6	6	4	4	**5**	**4**
23	1.3162	1.3841	0.0678	4	4	5	5	7	7	7	6	4	4	**7**	**7**
24	1.7423	1.9598	0.2175	4	4	5	4	6	5	6	5	4	3	**5**	**4**
25	1.5106	1.9640	0.4534	4	4	5	4	6	5	7	5	4	3	**5**	**4**
26	7.1659	5.6248	1.5410	–	–	2	2	2	2	2	2	1	1	**2**	**2**
27	0.5958	0.6090	0.0132	7	7	7	7	11	11	11	11	7	7	**11**	**11**
28	1.2939	1.2883	0.0056	4	4	5	5	7	7	7	7	4	4	**7**	**7**
29	2.7600	2.0822	0.6777	3	3	3	4	4	5	4	5	3	3	**3**	**4**
30	2.9136	3.2142	0.3006	3	3	3	3	4	4	4	4	2	2	**3**	**3**
31	0.8566	0.7435	0.1131	6	6	5	8	8	11	8	11	6	7	**8**	**11**
32	0.7246	1.1731	0.4484	5	5	11	3	–	–	–	–	–	–	**5**	**5**
33	5.6282	3.5115	2.1166	2	2	2	3	3	3	3	3	2	2	**2**	**3**
34	3.0722	4.1954	1.1232	2	2	3	3	4	3	4	3	2	2	**3**	**3**
35	0.4690	0.3940	0.0749	9	8	5	13	–	–	–	–	–	–	**9**	**8**
36	0.5150	0.5073	0.0077	8	8	7	8	12	12	12	12	8	8	**12**	**12**
37	2.8922	2.5560	0.3362	3	3	3	4	4	4	4	4	3	3	**3**	**4**
38	0.1076	0.1119	0.0042	17	18	21	10	–	–	–	–	–	–	**17**	**18**
39	0.9593	0.9784	0.0190	5	5	6	6	8	8	8	8	5	5	**8**	**8**
40	0.1841	0.1513	0.0327	15	13	–	–	–	–	–	–	–	–	**15**	**13**
41	0.0828	0.0777	0.0051	–	–	–	–	–	–	–	–	–	–	–	–
42	0.8173	0.6956	0.1217	6	6	5	8	–	–	–	–	–	–	**6**	**6**
43	0.0286	0.0425	0.0138	24	39	–	–	–	–	–	–	–	–	**24**	**39**
44	0.0597	0.0697	0.0100	21	26	–	–	–	–	–	–	–	–	**21**	**26**
45	1.6864	0.9808	0.7056	4	4	3	8	–	–	–	–	–	–	**3**	**8**
46	0.9639	1.7280	0.7641	4	4	8	3	–	–	–	–	–	–	**8**	**3**
47	0.0837	0.0797	0.0039	–	–	9	35	–	–	–	–	–	–	**9**	**35**
48	8.3634	9.7859	1.4225	–	–	2	2	2	2	2	2	1	1	**2**	**2**
49	9.9867	9.0587	0.9279	–	–	2	2	2	2	2	2	1	1	**2**	**2**

TABLE 4. ***(Continued.)*** **Left jitter k_l^- and right jitter k_l^+ detected by different masks in the CNA breakpoints of the 13th chromosomal sample "BLC_B1_T37.txt" in the 3σ sense with the confidence probability of $P = 99.73\%$. The chromosome is associated with breast cancer and all values are given for Log_2 Ratio.**

50	10.3886	16.6565	6.2678	–	–	1	1	1	1	1	1	1	1	**1**	**1**
51	17.3197	10.4969	6.8227	–	–	1	1	1	1	1	1	1	1	**1**	**1**
52	0.6826	0.6563	0.0262	7	6	6	7	10	11	10	11	7	7	**10**	**11**
53	0.2242	0.2322	0.0080	12	12	13	9	22	18	22	18	14	12	**22**	**18**
54	0.2228	0.2449	0.0220	11	12	17	7	–	–	–	–	–	–	**11**	**12**
55	0.9405	0.5706	0.3698	6	6	3	14	–	–	–	–	–	–	**6**	**6**
56	0.3219	0.3171	0.0048	10	10	9	10	15	16	15	16	10	11	**15**	**16**
57	12.7085	16.4145	3.7059	–	–	1	1	1	1	1	1	1	1	**1**	**1**
58	14.1556	10.3335	3.8221	–	–	1	1	1	1	1	1	1	1	**1**	**1**

2) SECOND BESSEL-BASED APPROXIMATION

The second approximation was obtained employing the same Bessel function (7), but with another variable,

$$\sigma_l^2(k) = \sigma_l^2 \left[1 + K_{\frac{1}{2}} \left(\frac{1}{(k+1)^{\beta(\gamma_l^-)} - 1} \right) \right], \qquad (9)$$

where $\beta(\gamma_l^-) = \sqrt{2}/\gamma_l^{0.1734}$. Testing (9) by simulations has shown that this function can produce more accuracy for certain values of SNR and that (8) can be more accurate otherwise, although both (8) and (9) can be applied to any k.

3) FUNCTIONAL APPROXIMATION

A simple approximation has appeared by using a power function of

$$\sigma_l^2(k) = \sigma_l^2 \frac{1}{2} \left[1 + k^{a(\gamma_l^-)^b} \right]^2, \qquad (10)$$

where $a = 0.436$ and $b = -0.1575$. An analysis has shown that (10) is about 10 times more accurate than (8) and (9) when $k > 1$, but cannot be applied to $k = 0$ or $k = 1$.

Functions (8) (dashed), (9) (solid), and (10) (circled and dotted) are sketched in Fig. 4 for $\gamma = 0.1, 1.0, 5.0$ in the range of $0 \leqslant k \leqslant 50$. As can be seen, the proposed k–varying variances are consistent, but produce different errors in the k-domain. Note that an exact function $\sigma_l^2(k)$ is still unavailable.

4) APPROXIMATION ERRORS

Based on the results illustrated in Fig. 3, the jitter distribution can now be approximated by the SkL pdf (2) parameterized by (8)–(10) using the sub-diagram A2 in Fig. 2. The approximation of the histogram *jitter_pdf$_k$* obtained by sub-diagram A1 in Fig. 2 has been provided iteratively by decreasing the approximation MSE produced by the parameterized SkL pdf in each cycle as represented in sub-diagram A2 in Fig. 2. Thereby, several approximations were obtained for functions (8)–(10).

Table 2 summarizes typical approximation MSEs produced by the SkL pdf (2) and parameterized SkL pdf using k-varying noise variances (8), (9) and (10). As can be seen, the parametrization by (8), (9), and (10) allow for much smaller errors than the SkL pdf (2) for low and extra low SNRs. One can see the goodness of fit in Fig. 5a and Fig. 5b obtained for $\gamma_l^- = \gamma_l^+ = 0.1$, which suggest that an essential difference between the SkL law (2) and (2) parameterized with (8)–(10) exists for all k.

In turn, Fig. 5c and Fig. 5d sketch the approximations for a normal SNR of $\gamma_l^- = \gamma_l^+ = 1.37$. Analysing this figure, one may conclude that the SkL pdf (2) is reasonably accurate when $|k| < 5$ and that the parameterized (2) should be used otherwise. An overall conclusion that comes up from this analysis is that the parameterized SkL pdf represents jitter in the CNA breakpoints with much more accuracy and should thus be used in the design of the confidence masks that we will discuss next.

C. IMPROVING CONFIDENCE UB AND LB MASKS

It follows from an analysis of errors produced by the proposed approximations that the most reliable results can be achieved if to develop the confidence masks worked out in [28] to be hybrid by using different approximations in diverse regions of the segmental SNRs.

Based on the MSEs produced by the approximations (Table 1), in Table 3 we select several segmental SNR regions, within which the MBA developed in [6], SkL pdf (2), and SkL pdf (2) parameterized with (8), (9), and (10) are most successful in detecting the right jitter k^- and the left jitter k^+ in the minimum MSE sense.

Table 3 suggests that for $\gamma_l^- = \gamma_l^+$ and $|k| \geqslant 0$, the SkL pdf (2) parameterized with (8) is most accurate in the SNR region of $0.9 \ldots 1.37$, while (9) gives a better accuracy in $0.1 \ldots 0.9$. The parametrization with (10) is also accurate when $0.1 <$ SNR < 1.37, but it cannot be applied to $k = 0$ and $k = 1$. When $\gamma_l^- \neq \gamma_l^+$, the MBA is preferable in the SNR region of $0.1 \ldots 1.37$ and (2) can be used otherwise for any step-index k.

It then follows that the best accuracy for the confidence UB and LB masks designed in [6] can be achieved if to make the masks hybrid. The difference between the hybrid masks and the basic ones [6] is in the parametrization of (2) and in the conditions introduced for the SNR values γ_l^- and γ_l^+. With such modifications, the basic masks can be used straightforwardly and we refer the reader to [6] for a detailed description of the basic algorithm.

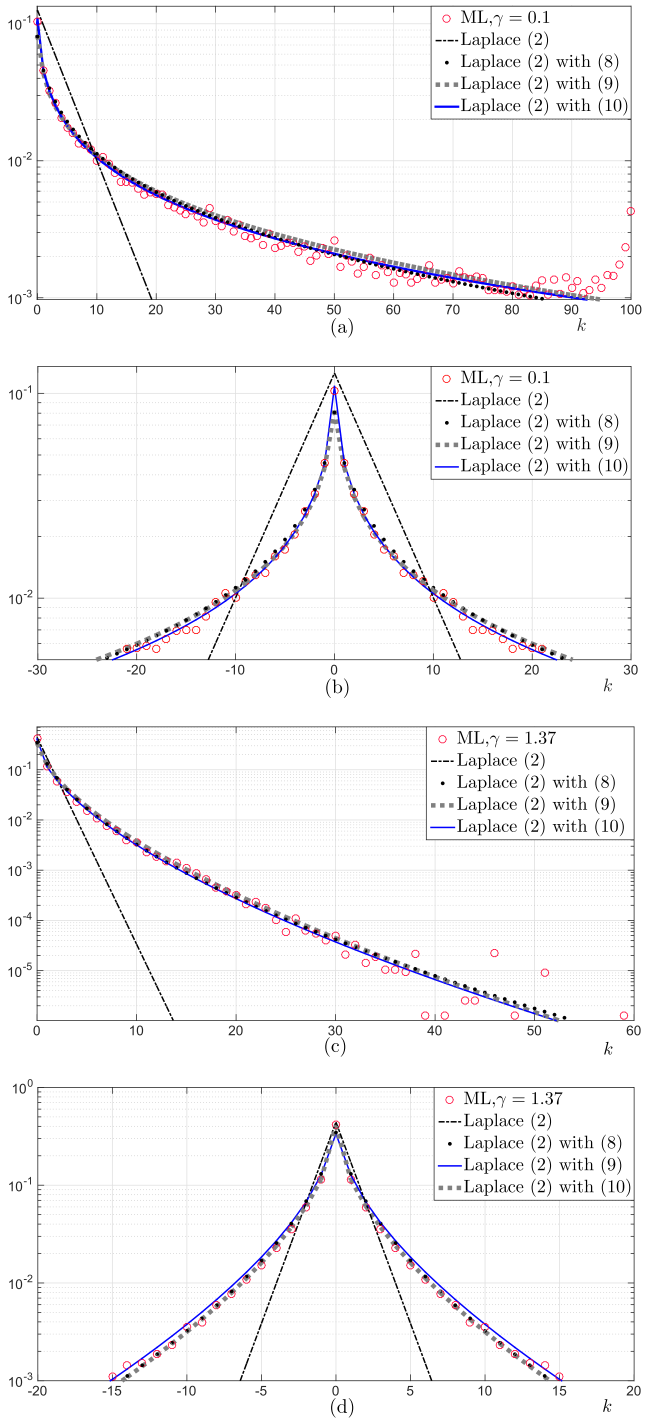


FIGURE 5. Measured jitter pdf functions (circles) and the approximations by the SkL law (2) and by the SkL law parameterized with (8), (9), and (10) for equal segmental SNRs $\gamma_l = \gamma_l^- = \gamma_l^+$: (a) low SNR $\gamma_l = 0.1$, (b) zoomed for $\gamma_l = 0.1$, (c) normal SNR $\gamma_l = 1.37$, and (d) zoomed for $\gamma_l = 1.37$. Measurement points (circles) are obtained using the ML estimator over 10^4 runs.

IV. APPLICATIONS TO SNP ARRAY PROBING

In this section, we experimentally test the parameterized SkL pdf (2) and several confidence masks by the SNP array-based CNAs probe data taken from the database BLC_BI_T31 available from the project GAP [5].

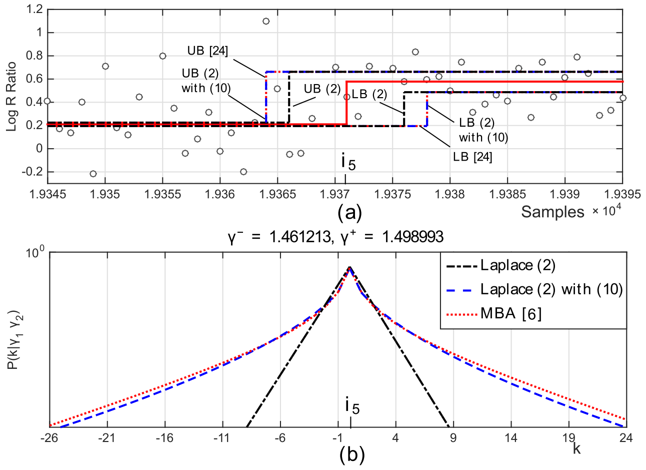


FIGURE 6. Testing the ML estimate of the breakpoint i_5 location of the sample BLC_BI_T31 in the 13th Chromosome [5] by the confidence masks: (a) ML estimate (solid) with the UB and LB masks and (b) jitter distribution approximated with the SkL law (2) and with (2) parameterized by (10) and MBA [6].

A. CONFIDENCE OF THE BREAKPOINT LOCATION

To emphasize again on a practical importance of the hybrid confidence masks, in Fig. 6 we examine a part of the 13rd chromosome [5] consisting of a single breakpoint i_5 and two segments with the segmental SNR values of $\gamma_5^- \cong 1.46$ and $\gamma_5^+ \cong 1.5$, as investigated in [5].

The candidate breakpoint was detected using the ML estimator and then the ML estimates were tested by different masks based on the SkL pdf (2), SkL pdf parameterized with (10), and the one proposed in [6] for the confidence probability of $P = 99.73\%$. The MBA and (2) parameterized with (10) demonstrate in Fig. 6b more accurate approximations. Therefore, the regions of possible breakpoint locations produced by these approximations (Fig. 6a) must be accepted as more realistic. As can be seen, these regions are wider than produced by the SkL pdf (2).

B. CHROMOSOME PROBING BY SNP ARRAY

We now apply the confidence masks to the estimates of the breakpoint locations in the complete chromosome 13th of the profile BLC_BI_T31 taken from the series of basal-like carcinomas (BLCs) available from the project GAP [5]. These series are included to the study of primary breast carcinomas (40 cases) and two cell lines measured on a 300K Illumina SNP-arrays (Human Hap300-Duo). The CNA profile is represented by the Log R ratios (LRRs) centered at zero for each sample. The estimates were obtained using the circular binary segmentation algorithm *cghcbs* [29], which suggests that the chromosome has 59 segments and 58 breakpoints as shown in Fig. 7.

It follows from Fig. 7 that the confidence intervals are wider for the segmental levels than for the breakpoints. Therefore, we supply this figure with Table 4, in which the left jitter k_l^- and the right jitter k_l^+ are estimated for the confidence probability $P = 99.73\%$ in the 3σ sense [28]. Here symbol "–" means that the jitter cannot be calculated by the masks.

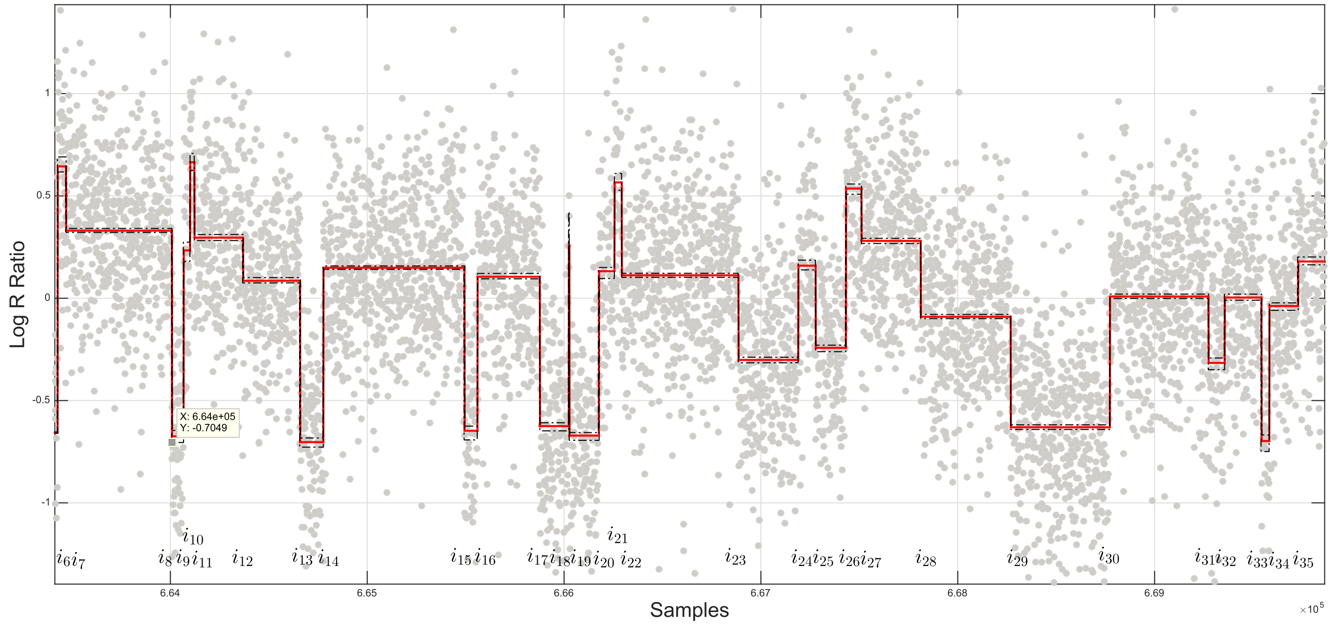


FIGURE 7. Probes (points), CNAs estimates (solid), and confidence regions (dashed) provided by the hybrid masks for the 13th chromosome taken from *BLC_BI_T*37 of GAP [5]. The breakpoint locations are detected using the algorithm *cghcbs* [29].

Table 4 suggests that the masks often produce unequal jitter estimates and that the difference between the estimates can be in several points, as in the case of $l = 7$ or $l = 27$. Large jitter in i_1, i_{40}, i_{43}, and i_{44} was detected only by the MBA. But the MBA was unsuccessful in detecting any jitter in a larger number of the breakpoints such as i_8, i_9, i_{14}, i_{18}–i_{20}, i_{26}, i_{48}–i_{51}, i_{57}, and i_{58}, while other masks provided it with near similar errors. One can also notice that the extra low SNR values make the jitter unavailable for bounding by any of the masks, as in the cases of i_4, i_5, and i_{41}.

Jitter computed by the hybrid masks is put to two last columns of Table 4. Because the hybrid masks combine the most accurate outputs of the particular masks, the left and right jitter computed by the hybrid masks can be considered as most reliable. What the hybrid masks suggest is that jitter in the breakpoints of this chromosome ranges from 1 point to tens of points and thus an actual breakpoint may be found tens points apart from the candidate one provided by an estimator.

V. CONCLUSION

The confidence masks are intended to bound regions of existence for the CNA segments and breakpoints with a given confidence probability in the presence of intensive probe noise. The parametrization of the SkL density provided by several approximations of the k-varying segmental noise variance has demonstrated much higher accuracy in bounding the breakpoints jitter for a given probability. The parametrization has appeared to be especially efficient for low and extra low segmental SNR values, when the breakpoints cannot be localized visually. The hybrid confidence masks combining best outputs of the particular masks have demonstrated an ability to bound the jitter with a high accuracy for practically all segmental SNR values observed in chromosomal probing. That was confirmed by testing the masks with a chromosome sample having 59 segments and 58 breakpoints and associated with breast cancer.

It has also been revealed that the left and right jitter in the breakpoints correlate each other. The parametrization of the SkL density can be provided with more accuracy if to account for the correlation properties of the k-varying segmental noise variances. This problem is now under investigation and we plan to report the results in near future.

Acknowledgment

This paper was presented at the 13th International Conference on Electrical Engineering, Computing Science and Automatic Control, September 26–30, 2016, Mexico City [12].

REFERENCES

[1] R. Redon *et al.*, "Global variation in copy number in the human genome," *Nature*, vol. 444, no. 7118, pp. 444–454, Nov. 2006.

[2] P. J. Hastings, J. R. Lupski, S. M. Rosenberg, and G. Ira, "Mechanisms of change in gene copy number," *Nature Rev. Genet.*, vol. 10, no. 8, pp. 551–564, Aug. 2009.

[3] A. Reymond, C. N. Henrichsen, and L. Harewood, "Side effects of genome structural changes," *Current Opinion Genet. Develop.*, vol. 17, no. 5, pp. 381–386, Oct. 2007.

[4] C. Alkan, B. P. Coe, and E. E. Eichler, "Genome structural variation discovery and genotyping," *Nature Rev. Genet.*, vol. 12, no. 5, pp. 363–376, May 2011.

[5] T. Popova, V. Boeva, E. Manié, Y. Rozenholc, E. Barillot, and M. H. Stern, "Analysis of somatic alterations in cancer genome: From SNP arrays to next generation sequencing," in *Genomics I: Humans, Animals and Plants*. Hong Kong: iConcept, 2013, pp. 133–154.

[6] J. Muñoz-Minjares and Y. S. Shmaliy, "Improving estimates of the breakpoints in genome copy number alteration profiles with confidence masks," *Biomed. Signal Process. Contr.*, vol. 10, pp. 238–248, Jan. 2017.

[7] C. Xie and M. T. Tammi, "CNV-seq, a new method to detect copy number variation using high-throughput sequencing," *BMC Bioinf.*, vol. 10, p. 80, Mar. 2009.

[8] S. Ivakhno, T. Royce, A. J. Cox, D. J. Evers, R. K. Cheetham, and S. Tavaré, "CNAseg—A novel framework for identification of copy number changes in cancer from second-generation sequencing data," *Bioinformatics*, vol. 26, no. 24, pp. 3051–3058, 2010.

[9] V. Boeva *et al.*, "Control-free calling of copy number alterations in deep-sequencing data using GC-content normalization," *Bioinformatics*, vol. 27, no. 2, pp. 268–269, 2011.

[10] J. Duan, J.-G. Zhang, Y.-P. Wang, and H. W. Deng, "Comparative studies of copy number variation detection methods for next-generation sequencing technologies," *PLoS ONE*, vol. 8, no. 3, p. e59128, 2013.

[11] L. J. Engle, C. L. Simpson, and J. E. Landers, "Using high-throughput SNP technologies to study cancer," *Oncogene*, vol. 25, pp. 1594–1601, Mar. 2006.

[12] J. Muñoz-Minjares, Y. S. Shmaliy, R. Olivera-Reyna, and O. Vite-Chavez, "Improving approximation of jitter probability in the breakpoints of simulated copy number alterations," in *Proc. 13th Int. Conf. Elect. Eng., Comput. Sci. Autom. Control (CCE)*, Mexico City, Mexico, Sep. 2016, pp. 1–5.

[13] Y. S. Shmaliy, "On the multivariate conditional probability density of a vector perturbed by Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4792–4797, Dec. 2007.

[14] J. Muñoz-Minjares, Y. S. Shmaliy, and J. Cabal-Aragon, "Noise studies in measurements and estimates of stepwise changes in genome DNA chromosomal structures," in *Proc. Int. Conf. Pure Math., Appl. Math., Comput. Methods (PMAMCM)*, Santorini Island, Greece, Jul. 2014, pp. 212–221.

[15] R. Pique-Regi, A. Ortega, A. Tewfik, and S. Asgharzadeh, "Detection changes in DNA copy number: Reviewing signal processing techniques," *IEEE Signal Process. Mag.*, vol. 29, no. 1, pp. 98–107, Dec. 2011.

[16] A. Joshi, "Speech emotion recognition using combined features of HMM & SVM algorithm," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 8, pp. 387–393, 2013.

[17] P. Hupé, N. Stransky, F. Radvanyi, E. Barillot, and J.-P. Thiery, "Analysis of array CGH data: From signal ratio to gain and loss of DNA regions," *Bioinformatics*, vol. 20, no. 18, pp. 3413–3422, 2004.

[18] C. Lemaitre, E. Tannier, M.-F. Sagot, and C. Gautier, "Precise detection of rearrangement breakpoints in mammalian chromosomes," *BMC Bioinf.*, vol. 9, no. 1, p. 286, 2008.

[19] K. Wong, T. M. Keane, J. Stalker, and D. J. Adams, "Enhanced structural variant and breakpoint detection using SVMerge by integration of multiple detection methods and local assembly," *Genome Biol.*, vol. 11, no. 12, p. R128, 2010.

[20] D. L. Donoho, "De–noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[21] O. V. Lepski, E. Mammen, and V. G. Spokoiny, "Optimal spatial adaptation to inhomogeneous smoothness: An approach based on kernel estimates with variable bandwidth selectors," *Ann. Stat.*, vol. 25, pp. 929–947, Jun. 1997.

[22] Y. Li, Y. Ding, and T. Li, "Nonlinear diffusion filtering for peak-preserving smoothing of a spectrum signal," *Chemometrics Intell. Lab. Syst.*, vol. 156, pp. 157–165, Aug. 2016.

[23] G. Deng, "Guided wavelet shrinkage for edge-aware smoothing," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 900–914, Feb. 2017.

[24] F. Picard, S. Robin, M. Lavielle, C. Vaisse, and J.-J. Daudin, "A statistical approach for array CGH data analysis," *BMC Bioinf.*, vol. 6, no. 1, pp. 27–37, Jan. 2012.

[25] J. Muñoz-Minjares and Y. S. Shmaliy, "Approximate jitter probability in the breakpoints of genome copy number variations," in *Proc. 10th Int. Conf. Elect. Eng., Comput. Sci. Autom. Control (CCE)*, Mexico City, Mexico, Sep./Oct. 2013, pp. 128–131.

[26] J. Muñoz-Minjares, J. Cabal-Aragon, and Y. S. Shmaliy, "Effect of noise on estimates of stepwise changes in genome DNA chromosomal systems," *WSEAS Trans. Biol. Biomed.*, vol. 11, pp. 52–61, Apr. 2014.

[27] T. J. Kozubowski and S. Inusah, "A skew Laplace distribution on integers," *Ann. Inst. Stat. Math.*, vol. 58, pp. 555–571, Sep. 2006.

[28] J. Muñoz-Minjares, J. Cabal, and Y. S. Shmaliy, "Confidence masks for genome DNA copy number variations in applications to HR-CGH array measurements," *Biomed. Signal Process. Control*, vol. 13, pp. 337–344, Sep. 2014.

[29] A. B. Olshen, E. S. Venkatraman, M. Wigler, and R. Lucito, "Circular binary segmentation for the analysis of array-based DNA copy number data," *Biostatistics*, vol. 5, no. 4, pp. 557–572, 2004.

[30] J. Muñoz-Minjares, Y. S. Shmaliy, and J. Cabal-Aragón, "Confidence limits for genome DNA copy number variations in HR-CGH array measurements," *Biomed. Signal Process. Control*, vol. 10, pp. 166–173, Mar. 2014.

JORGE MUNOZ-MINJARES was born in Zacatecas, Mexico, in 1987. He received the B.S. degree in communications and electronics engineering from the Universidad Autonoma de Zacatecas in 2010 and the M.S. degree in electrical engineering from DICIS, Universidad de Guanajuato, in 2012, where he is currently pursuing the Ph.D. degree. His research interests include digital signal e-image processing, optimal filtering, and probability and statistics.

YURIY S. SHMALIY (M'96–SM'00–F'11) was born in Beltsy, Moldova, in 1953. He received the B.S., M.S., and Ph.D. degrees from the Kharkiv Aviation Institute, Ukraine, in 1974, 1976, and 1982, respectively, all in electrical engineering, and the D.Sc. degree from the Kharkiv Railroad Institute in 1992. From 1985 to 1999, he was with Kharkiv Military University. He was a Full Professor with Kharkiv Military University in 1986. In 1992, he founded the Scientific Center Sichron, where he was the Director in 2002. Since 1999, he has been with the Universidad de Guanajuato of Mexico. From 2012 to 2015, he was the Head of the Department of Electronics Engineering, Universidad de Guanajuato of Mexico. He was a Visiting Professor–Researcher with City University London from 2015 to 2016. He has 399 journal and conference papers and 81 patents. He has authored the books *Continuous-Time Signals* (Springer, 2006), *Continuous-Time Systems* (Springer, 2007), *GPS-Based Optimal FIR Filtering of Clock Models* (New York: Nova Science Publishers, 2009), and *Probability: Interpretation, Theory and Applications* (New York: Nova Science Publishers, 2012) and contributed to several books with invited chapters. His current interests include optimal estimation, statistical signal processing, and stochastic system theory. He was rewarded a title, Honorary Radio Engineer of the USSR, in 1991. He was listed in Outstanding People of the 20th Century, Cambridge, U.K., in 1999. He has received the Royal Academy of Engineering Newton Research Collaboration Program Award in 2015. He is currently an associate editor and an editorial board member in several journals. He was invited many times to give tutorial, seminar, and plenary lectures.

LUIS J. MORALES-MENDOZA was born in Veracruz, Mexico, in 1974. He received the B.S. and M.S. degrees in electrical engineering from Guanajuato University, Mexico, in 2001 and 2002, respectively, and the Ph.D. degree in electrical engineering from the Research Center (Cinvestav), National Polytechnical Institute of Mexico, Guadalajara, in 2006. From 2006 to 2009, he was an Assistant Professor with the Electronics Department, Guanajuato University of Mexico. He is currently an Associate Professor with the Electronics Department, Universidad Veracruzana. He has authored or co-authored 23 journal and conference papers. His scientific interests are in the artificial neural networks applied to optimization problems, image restoration and enhancing, and ultrasound image processing.

MIGUEL VAZQUEZ-OLGUIN (M'16) was born in Mexico in 1982. He received the B.S. degree in electronics and communications from the Universidad Iberoamericana de Leon, Leon, Mexico, in 2005, and the M.S. degree in electronics and communications from the Center for Scientific Research and Higher Education of Ensenada, Ensenada, Mexico, in 2009. He is currently pursuing the Ph.D. degree with the Universidad de Guanajuato, Salamanca, Mexico. His current areas of interest are consensus filtering, wireless sensor networks, and optimal estimation.

CARLOS LASTRE-DOMINGUEZ was born in Sincé, Colombia, in 1987. He received the B.S. degree from the Universidad de Pamplona, Colombia, in 2011, and the M.I. degree from the Universidad Industrial de Santander, Santander, Colombia, in 2016. He is currently pursuing the Ph.D. degree in electrical engineering with the Universidad de Guanajuato. His scientific interests are machine learning, digital signal processing, and optimum filter applied to biomedical signals. He has also participated in various congresses about the mentioned topics.