A New Smoothing Model for Analyzing Array CGH Data

Nha Nguyen*, Heng Huang†, Soontorn Oraintara* and An Vo*
*Department of Electrical Engineering, University of Texas at Arlington
Email: nhn3175@exchange.uta.edu, oraintar@uta.edu, vpnan@gauss.uta.edu
†Corresponding Author
Department of Computer Science and Engineering, University of Texas at Arlington
Email: heng@uta.edu

***Abstract*—Array based Comparative Genomic Hybridization (CGH) is a molecular cytogenetic method for the detection of chromosomal imbalances and it has been extensively used for studying copy number alterations in various cancer types. Our method captures both the intrinsic spatial change of genome hybridization intensities, and the physical distance between adjacent probes along a chromosome which are not uniform. In this paper, we introduce a dual-tree complex wavelet transform method with the bivariate shrinkage estimator into array CGH data smoothing study. We tested the proposed method on both simulated data and real data, and the results demonstrated superior performance of our method in comparison with extant methods.**

I. Introduction

Array-based comparative genomic hybridization (array CGH) is a highly efficient technique, allowing the simultaneous measurement of DNA copy numbers across the whole genome at hundreds or thousands of loci and the reliable detection of local one-copy-level variations. Characterization of these DNA copy number changes is important for both the basic understanding of cancer and its diagnosis. In order to develop effective methods to identify aberration regions from array CGH data, many recent research works focus on smoothing-based data processing. For example, Eilers and De Menezes proposed a quantile regression method that employs an L1 error for both of fitness measure and roughness penalty [1]. Hsu et al. [2] used wavelet transform to fit the data. In this paper, we introduce a dual-tree complex wavelet transform method with the bivariate shrinkage estimator into array CGH data smoothing study. The unequal spacing of probes on the chromosome is taken into account. Using the synthetic data, our experimental results demonstrate our method overperforms the previous methods. In terms of the root mean squared error measurement at different noise levels, our method improves about $17.8\% - 43\%$ than other methods. Furthermore, we also use the real array CGH data to validate the efficiency of our method.

II. Wavelet Methods

In this section, we provide a brief review of wavelet transforms which were used for array CGH data smoothing and is used by this paper.

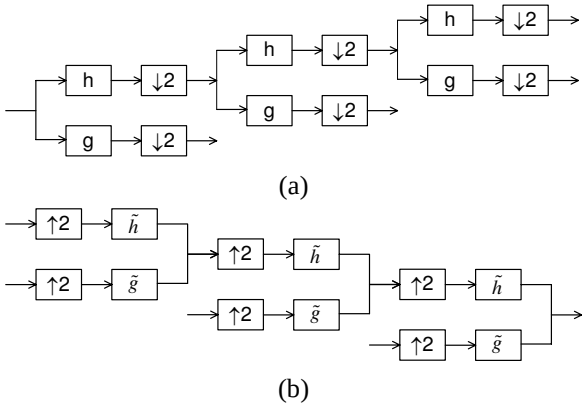


Fig. 1. A 3 level DWT. (a) Analysis FB, (b) Synthesis FB .

A. Discrete Wavelet Transform

The discrete wavelet transform (DWT), based on the octave band tree structure, can be viewed as the multiresolution decomposition of a signal. Fig. 1 shows 3 level DWT analysis and synthesis filter banks (FBs). It takes a length N sequence, and generates an output sequence of length N using a set of lowpass and highpass fiters followed by a decimator. It has $N/2$ values at the highest resolution, $N/4$ values at the next resolution, and $N/2^L$ at the level L. Because of decimation, the DWT is a critically sampled decomposition. However, the drawback of DWT is the shift variant property. In signal denoising, the DWT creates artifacts around the discontinuities of the input signal [3]. These artifacts degrade the performance of the threshold-based denoising algorithm.

B. Stationary Wavelet Transform

The stationary wavelet transform (SWT) [3] is similar to the DWT except that it does not employ a decimator after filtering, and each level's filters are upsampled versions of the previous ones. The SWT is also known as the shift invariant DWT. The absence of a decimator leads to a full rate decomposition. Each subband contains the same number of samples as the input. So for a decomposition of L levels, there is a redundant ratio of $(L+1):1$. However, the shift invariant property of the SWT makes it preferable for the usage in various signal processing applications such as denoising and classification

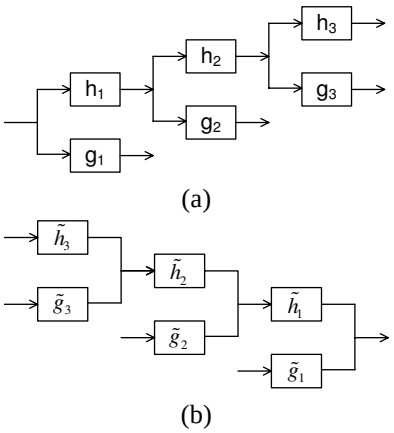


Fig. 2. A 3 level SWT. (a) Analysis FB, (b) Synthesis FB .

because it relies heavily on spatial information. It has been shown that many of the artifacts could be suppressed by a redundant representation of the signal [3]. Fig. 2 shows 3 level SWT analysis and synthesis FBs. Each level's filters are upsampled versions of the previous ones as shown in Fig. 3.

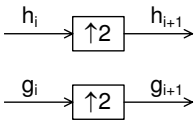


Fig. 3. SWT filters.

C. Dual-Tree Complex Wavelet Transform

A dual-tree structure that produces a dyadic complex DWT is proposed by Kingsbury [4], [5]. Since array CGH data are one dimensional signals, in this paper we only talk about the 1-D case of dual-tree CWT. The DTCWT filter bank structure is shown in Fig. 4. The analysis FB for the DTCWT is an iterative multiscale FB. Each resolution level consists of a pair of two-channel FBs. The input signal is passed through the first level of a multiresolution FB. The low frequency component, after decimation by 2, is fed into the second level decomposition for the second resolution. The outputs of the two trees are the real and imaginary parts of complex-valued subbands. To reconstruct the signal, the real part and imaginary part are inverted to obtain two real signals, respectively. These two real signals are then averaged to obtained the final output. For more details of the construction of the dual-tree, the reader is referred to [6].

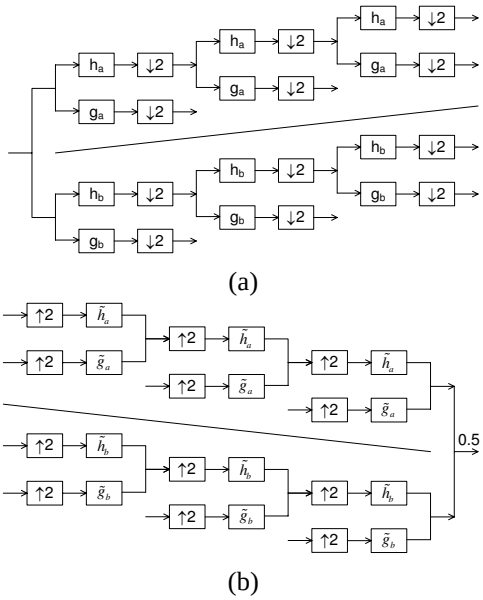


Fig. 4. The 3 level DTCWT filter bank structure. (a) Analysis FB, (b) Synthesis FB .

The most important property of the DTCWT is that all complex subbands are shift invariant in the sense that there is no significant aliasing in the decimated complex subbands. Therefore, each complex subband provides a shiftable description of signal in a specific scale. By construction of the dual-tree CWT, each pair of corresponding filters has the Hilbert transform relation [6]. It is therefore an overcomplete representation with a redundant ratio of $2:1$. In the two trees, the filters are designed in such a way that the aliasing in one branch in the first tree is approximately canceled by the corresponding branch in the second tree. The relation between the wavelet filters of the two trees yields shift invariant property [4]. The equivalent complex filter for each subband has one-sided frequency support. The real part of the complex filter is symmetric while the imaginary part is anti-symmetric.

III. Bivariate Shrinkage Function for DTCWT-Based Denoising.

We assume that we get the DNA copy number data Y which includes the deterministic signal D and the independent and identically distributed (IID) Gaussian noise n. This Gaussian noise has zero mean and variance σ_n^2.

$$Y = D + n. \tag{1}$$

After decomposing the data Y by the DTCWT, we get the complex coefficients $\mathbf{y}_k$. In the wavelet domain, the problem can be formulated as

$$\mathbf{y}_k = \mathbf{w}_k + \mathbf{n}_k, \tag{2}$$

where $\mathbf{y}_k$ are noisy wavelet coefficients, $\mathbf{w}_k$ are true coefficients, and $\mathbf{n}_k$ are independent Gaussian noise coefficients.

A simple denoising algorithm via wavelet transform consists of three steps: decompose the noisy signal by wavelet transform, denoise the noisy wavelet coefficients according to some rules and take the inverse wavelet transform from the denoised coefficients. To estimate wavelet coefficients, some of the most well-known rules are universal thresholding, soft thresholding [7], [8], [9] and BayesShrink [10]. In these algorithms, the authors assume that wavelet coefficients are independent. However, recently, algorithms utilize the dependency between coefficients can give better results compared with the ones developed using an independency assumption [11]. Sendur [11] has exploited this dependency

between coefficients and proposed a non-Gaussian bivariate pdf for the child coefficient w_1 and its parent w_2 as follows

$$p_{\mathbf{w}}(\mathbf{w}) = \frac{3}{2\pi\sigma^2}\exp(-\frac{\sqrt{3}}{\sigma}\sqrt{|w_1|^2+|w_2|^2}). \quad (3)$$

The marginal variance σ^2 is dependent on the coefficients index k. Using this bivariate pdf and the Bayesian estimation theory, the MAP estimator of w_1 [11] is derived to be

$$\hat{w_1} = \frac{(\sqrt{|y_1|^2+|y_2|^2} - \frac{\sqrt{3}\sigma_n{}^2}{\sigma})_+}{\sqrt{|y_1|^2+|y_2|^2}}.y_1, \quad (4)$$

where $(u)_+$ is defined by

$$(u)_+ = \begin{cases} 0 & \text{if } u < 0, \\ u & \text{otherwise.} \end{cases} \quad (5)$$

This estimator can be called as a bivariate shrinkage function. In (4), σ can be estimated by

$$\hat{\sigma} = \sqrt{(\hat{\sigma_y}{}^2 - \hat{\sigma_n}{}^2)_+}, \quad (6)$$

where $\hat{\sigma_n}$ is the noise deviation which is estimated from the finest scale wavelet coefficients by using a robust median estimator [8] as follows

$$\hat{\sigma_n}{}^2 = \frac{\text{median}(|y_i|)}{0.6745}. \quad (7)$$

$\hat{\sigma_y}$ is the deviation of observation signal estimated by

$$\hat{\sigma_y}{}^2 = \frac{1}{M}\sum_{y_i \in N(k)} |y_i|^2, \quad (8)$$

where M is the size of the neighborhood $N(k)$.

The DTCWT-based denoising algorithm using bivariate shrinkage function is summarized as follows

Step 1 Decompose the noisy signal Y by the DTCWT.
Step 2 Calculate the noise variance $\hat{\sigma_n}{}^2$ and the marginal variance $\hat{\sigma}^2$ by using (7), (8) and (6).
Step 3 Estimate the coefficients $\hat{w_k}$ as in (4).
Step 4 Take the inverse DTCWT from the denoised coefficients.

IV. Experiments and Discussions

In our experiments, the artificial chromosomes are generated as in [12] and [13]. Since they are unequally spaced data, we apply the interpolating method with the Pseudo-markers [13]. In order to guarantee the number of data points to be a power of two, the zero-padding is implemented.The DTCWT and the bivariate shrinkage function are proposed for denoising these chromosomes. The denoising results of different methods are compared.

A. Artificial Chromosome Generation

Willenbrock and Fridlyand [12] proposed a simulation model to create the synthetic array CGH data with equally spaced along the chromosome. More recently Y. Wang and S. Wang [13] extended this model by placing unequally spaced probes along chromosome. As suggested in [12] and [13], the chromosomal segments with DNA copy number $c = 1, 2, 3, 4$ and 5 are generated with probability $0.01, 0.08, 0.81, 0.07, 0.02$ and 0.01. The lengths for segments are picked up randomly from the corresponding empirical length distribution given in [12]. Each sample is a mixture of tumor cells and normal cells. A proportion of tumor cells is P_t, whose value is from a uniform distribution between 0.3 and 0.7. As in paper [12], the $log2ratio$ is calculated by

$$log2ratio = \log_2\left(\frac{cP_t + 2(1-P_t)}{2}\right), \quad (9)$$

where c is the assigned copy number. The expected $log2ratio$ value is then the latent true signal.

Gaussian noises with zero mean and variance σ_n^2 are added to the latent true signal. Till now, we get the equally spaced CGH signal. Because the distances between two probes are randomly, the best way to get these distances is from the UCSF HumArray2 BAC array. Thus, we create a real CGH signal from the equally spaced CGH signal when the unequally spaced probes are placed on the chromosome. Now, we have many artificial chromosomes of length 200 $Mbase$ which are created by many noise levels $\sigma_n = 0.125, 0.15, 0.2, 0.25, 0.275$ and 0.3.

B. DNA Copy Number Data Interpolation

The distances between two samples in DNA copy number data vary greatly. Kovac [14] proposed a new method to change this kind of data to equally spaced data and got good performance in denoising application. However, in his method, the numbers of new samples were created densely. In [13] DNA copy number data interpolation with pseudo-markers was proposed. Their method has some advantages: the number of new samples is not dense and the content of data does not change much. Suppose that the observed DNA copy number data Y at M probe locations on the chromosome of length L is

$$Y(x_i) = D(x_i) + n(x_i),$$

where $i = 1, 2, ..., M$, x_i are the probe locations with $0 \le x_1 \le x_2 \le .... \le x_M \le L$, and $D(x_i)$ is the latent true signal at location x_i. We know that x_i are not equally spaced. Therefore, the equally spaced pseudo-markers between sparse probes are inserted and this interpolating method can be summarized as the following steps.

First, a set P of locations along the chromosome at the uniform spacing of q is created by

$$P = \{p_j | p_j = kq, k = 0, 1, 2, ..., \lfloor L/q \rfloor\},$$

where q is the average distance between adjacent probes.

Next, we form a set P' of locations which will be inserted to DNA copy number data $Y(x_i)$. A pseudo-marker can be inserted at each location p_j. However, to avoid the worst cases when the original marker and a pseudo-marker overlap or their distance is arbitrarily small, a subset P' of P is formed by

$$P' = \{p'_j | p'_j \in P, |p'_j - x_i| \geq q/2 \text{ for all } i = 1, 2, ..., M\}.$$

P' includes the points in the set P and satisfies a condition that the distances from that point to the original locations must be larger than $q/2$ but less than $3q/2$. The nearest neighbor interpolation is applied to obtain the interpolated values for $Y(p'_j)$.

Finally, the original signal $Y(x_i)$ and the interpolated signal $Y(p'_j)$ are merged by

$$Y' = \{Y(x_i) : i = 1, 2, ..., M\} \cup \{Y(p'_j) : p'_j \in P'\},$$

and x_i and p'_j are made sure in ascending order. The new CGH data $Y'(x_i)$ instead of the original $Y(x_i)$ will be denoised.

C. Zero-padding

In order to get the best performance in the wavelet denoising algorithm, the length of the input signal is required to be a power of two [15]. After interpolating, the length of the CGH signal is N. If N is not a power of two, we can apply the zero-padding method to make sure $N = 2^j$. For example, we get $N = 120$ after interpolating, we must insert more eight zeros and then we get $N = 128$. The numbers of decomposition levels can be computed by

$$L = \log_2(N) - 4. \tag{10}$$

In the above example, we get $L = 3$. This is a perfect number of levels which yields the best denoising result.

D. Proposed Method and Experiment Design

The DWT with the redundant ratio of $1:1$ is efficient for the denoising applications. However, the DWT creates artifacts around the discontinuities of the input signal [3] because it is shift-variant. To overcome this problem, SWT with translation invariant property was proposed for signal denoising. It has been shown that many of the artifacts could be suppressed by a redundant representation of the signal [3]. However, if a signal is decomposed into L levels, the redundant ratio is $(L+1):1$. It makes denoising algorithm more computationally intensive. Several methods were proposed for selecting thresholding values such as hard universal [8], [7] and un-universal thresholding [9]. However, the dependency between wavelet coefficients are not exploited in these methods. Thus, we propose the usage of shift invariant DTCWT with the redundant ratio is $2:1$ and bivariate shrinkage estimator which takes advantage of the dependency between wavelet coefficient and its parent for array-based DNA copy number data denoising.

As discussed in section IV-B, DNA copy number data has the unequal distances between two samples. We use interpolating method in section IV-B to reduce the difference of those distances. Data are interpolated before decomposing

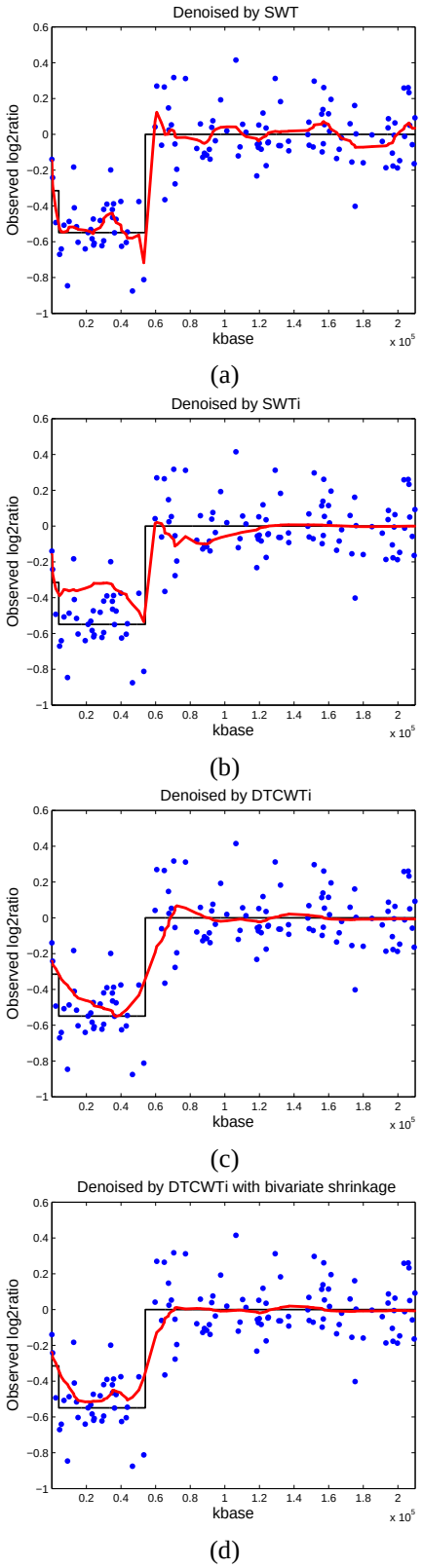
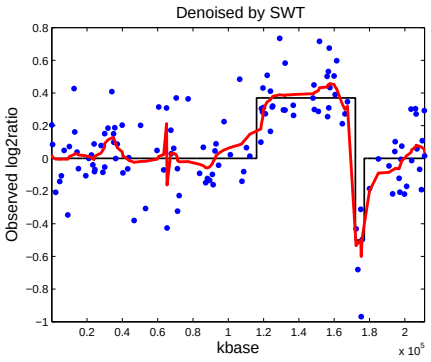
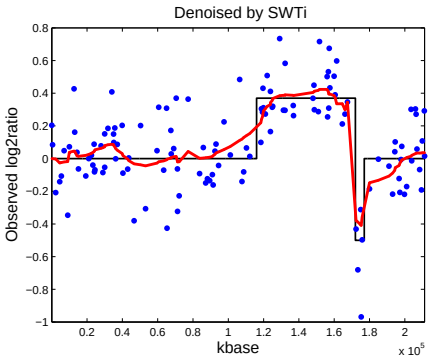


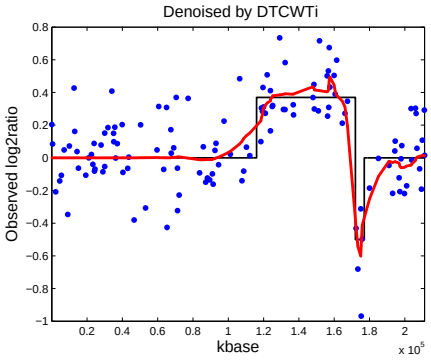
Fig. 5. Example of wavelet denoising results at the noise level of $\sigma = 0.15$ using (a) SWT, (b) SWTi, (c) DTCWTi and (d) DTCWTi-bi.



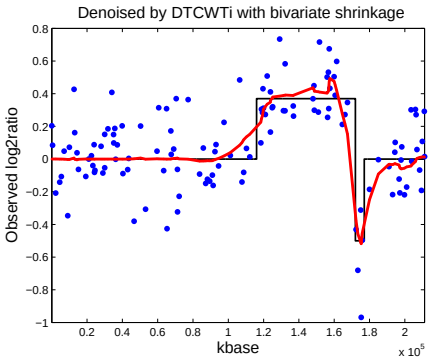
(a)



(b)

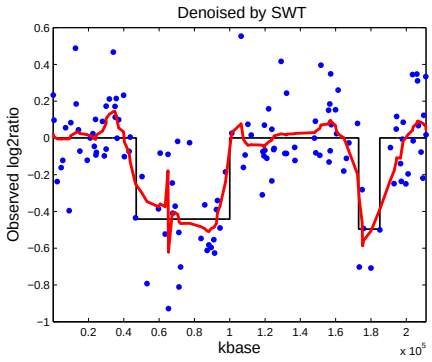


(c)

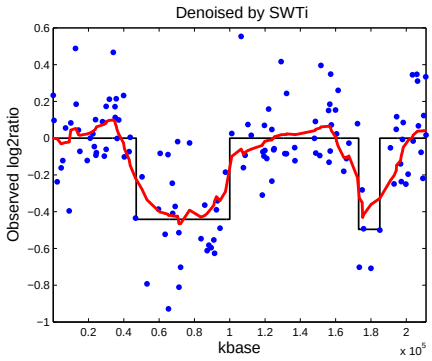


(d)

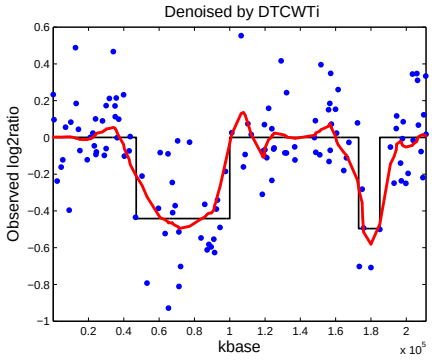
Fig. 6. Example of wavelet denoising results at the noise level of $\sigma = 0.2$ using (a) SWT, (b) SWTi, (c) DTCWTi and (d) DTCWTi-bi.



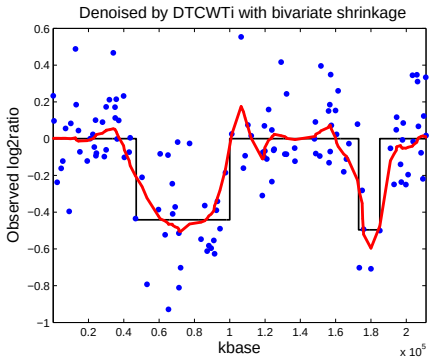
(a)



(b)



(c)



(d)

Fig. 7. Example of wavelet denoising results at the noise level of $\sigma = 0.25$ using (a) SWT, (b) SWTi, (c) DTCWTi and (d) DTCWTi-bi.

by the DTCWT. Our purpose is to find $\hat{D}$ from Y so that the root mean squared error (RMSE) (11) is the smallest.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(\hat{D}_i - D_i)^2}, \quad (11)$$

and N is the number of input samples.

The proposed method as called the DTCWTi-bi can be summarized as follows:

Step 1 : *Interpolate the DNA copy number data Y and get the interpolated DNA copy number data $Y^{'}$.*

Step 2 : *Insert zeros into $Y^{'}$ and decompose new data $Y^{''}$ by the DTCWT to L levels as (10).*

Step 3 : *Calculate the noise variance $\hat{\sigma_n}^2$ and the marginal variance $\hat{\sigma}^2$ for wavelet coefficient y_k by using (7), (8) and (6).*

Step 4 : *Estimate the coefficients $\hat{w_k}$ as in (4).*

Step 5 : *Reconstruct data $\hat{D}$ from the denoised coefficients $\hat{w_k}$ by taking inverse DTCWT.*

To compare the performance of the DTCWTi and the DTCWTi-bi algorithm with other methods such as the SWT and the SWTi [13] in array-based DNA copy number data denoising, we also implement the DTCWTi, the SWTi and the SWT. The main points of these methods are as follows

DTCWTi : The DTCWTi method has the same step 1, step 2 and step 5 as the DTCWTi-bi. They are different in step 3 and step 4. The universal thresholding value is applied to modify the noisy DTCWT coefficients. DWT coefficients of the array CGH signal at level 1 are used to estimate noise by (7).

SWT : The SWT method comes from paper [13]. Compared with the DTCWTi, the SWT method has two different steps: 1) the array CGH signal is decomposed by the SWT; 2) the real coefficients of SWT subbands are denoised instead of the complex coefficients. The universal thresholding is also used in this method.

SWTi : It extends the SWT method by interpolating array CGH signal before decomposition. The term by term thresholding is applied to denoise the SWT coefficients [13].

For the SWT and SWTi, only the scaling coefficients are denoised. However, for the DTCWTi and DTCWTi-bi, all subband coefficients are denoised.

E. Empirical Results

In this section, we will present the results when applying four methods in section IV-D. One thousand artificial chromosomes with six different noise levels $\sigma_n = 0.125, 0.15, 0.2, 0.25, 0.275$ and 0.3 are denoised.

The denoising results of all methods are shown in the table I. We can see that the proposed DTCWTi-bi method yields the better performance than the others. The DTCWTi-bi outperforms the SWT by $24.6\% - 43\%$, the SWTi [13] by $17.8\% - 32.5\%$ and the DTCWTi by $0.9\% - 1.5\%$ in terms of the RMSEs. Moreover, the DTCWTi-bi is more efficient and has less computation than the SWTi because the redundancy ratio of the DTCWT $2:1$ is much less than that of the SWT $4:1$ (if the number of level decomposition $L = 3$). For all noise levels, the DTCWTi-bi consistently achieves much better results than the SWT and SWTi.

TABLE I
COMPARISON OF AVERAGE RMSES OBTAINED FROM THE 1,000 ARTIFICIAL CHROMOSOMES WITH EACH OF THE 6 NOISE LEVELS USING THE SWT, THE SWTI, THE DTCWTI AND THE DTCWTI-BI.

σ	SWT	SWTi	DTCWTi	DTCWTi-bi
0.125	0.0460	0.0422	0.0350	0.0347
0.15	0.0548	0.0497	0.0393	0.0387
0.2	0.0715	0.0631	0.0469	0.0463
0.25	0.0874	0.0751	0.0530	0.0525
0.275	0.0952	0.0810	0.0558	0.0555
0.3	0.1027	0.0867	0.0587	0.0585

From the table I, we also found the evidence to prove that the bivariate shrinkage function should be applied to the CGH data denoising instead of the universal thresholding or the term by term thresholding. For example, at the noise level $\sigma_n = 0.15$, the RMSE of the DTCWTi-bi method is 0.0387, but that of the DTCWTi method is 0.0393. In this case, the DTCWTi-bi outperforms DTCWTi by 1.5%.

Some examples of wavelet denoising results by using SWT, SWTi, DTCWTi and DTCWTi-bi methods are shown in Fig. 5 at the noise level of $\sigma = 0.15$, in Fig. 6 at the noise level of $\sigma = 0.2$ and in Fig. 7 at noise level $\sigma = 0.25$. In those figures, the black solid lines represent the latent true signals, the blue points stand for the noisy DNA copy data log_2ratio at the probe loci and the red lines correspond to the denoised data. We should note that the line connecting the denoised data points is only for visualization purpose.

At the copy one (from 1 $kbase$ to 4.001 $kbase$) as shown in Fig. 5, the log_2ratio value of the latent true signal is -0.2975, but these values of the SWT-based data in Fig. 5(a) are from -0.1454 to -0.4156. These values can cause a mistake when we segment the DNA copy number data. However, the denoised data using the DTCWTi and the DTCWTi-bi will be segmented correctly as the copy one (from -0.4 to -0.2) because the log_2ratio values are from -0.2371 to -0.2957 in Fig. 5 (d) and from -0.2403 to -0.3236 in Fig. 5(c). At the copy zero $c = 0$ (from 4.001 $kbase$ to 0.54x10^5 $kbase$), the denoised data in Fig. 5(b) has an amplitude of 0.3181 which will make an error in segmentation process, while the denoised data in Fig. 5 (c) and (d) will give a correct segmentation. Fig 5(a) has a negative peak of -0.7178 which corresponds to an error rate of 31% at 0.53x10^5 $kbase$. Fig 5(b) gives the larger error rate of 42%, while it is only 16% in Fig 5(d). At the copy two $c = 2$ (form 0.54x10^5 $kbase$ to the end of the chromosome), Fig 5(c) and (d) look smoother than (a) and (b). When compared with Fig 5(c), the denoised data in Fig 5(d) tracks the latent true signal more closely than in Fig 5(c) because the bivariate shrinkage function is applied to recover the latent true signal in Fig 5(d).

From 1 to 1.16x10^5 $kbase$, the denoised signals using

DTCWTi and DTCWTi-bi are approximately the latent true signals, while the denoised data using the SWT and the SWTi have many ripples as shown in Fig. 6. From $1.72\text{x}10^5$ $kbase$ to $1.77\text{x}10^5$ $kbase$, $log_2 ratio$ of the latent true signal is -0.5001. When the DTCWTi-bi method is applied, this value is -0.5185 with the error rate of 3.7%. However, when the SWT, the SWTi and the DTCWTi are used, this value is -0.6002, -0.409 and -0.6011 with the error rate of 20%, 18.2% and 20.2%, respectively. It means that the denoised data using DTCWTi-bi tracks the latent true signal very well. In Fig. 6, the DTCWTi-bi achieves better accuracy than the SWTi by 9.4% and the SWT by 15.38% in terms of the RMSEs.

In Fig. 7, the RMSE values of the SWT, the SWTi, the DTCWTi and the DTCWTi-bi are $0.0794, 0.0774, 0.0723$ and 0.0716, respectively. If compared to SWTi method, the DTCWTi-bi outperforms by 7.5%.

From above results, we can see that our proposed DTCWTi-bi method with the bivariate shrinkage estimator is better than the others in terms of the RMSEs and the efficiency in the DNA copy number data denoising application.

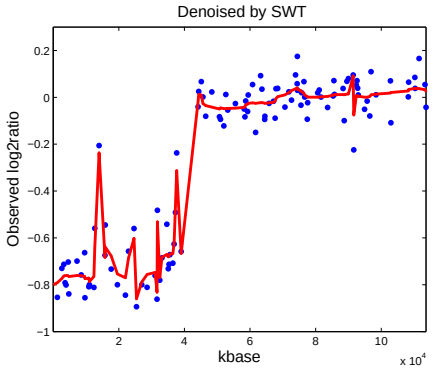
F. Real Data Examples

In this paper, the BAC array data on 15 fibroblast cell lines [16], [2] has been used to show that denoising by the DTCWTi-bi is better than by the SWTi. This data set is from Stanford University, which can be freely downloaded at http://www.nature.com/ng/journal/v29/n3/suppinfo/ng754 S1.html. Because the true copy number changes are known for these cell lines, we choose these data as a proof of principles. We pick up the chromosome 9 of MPE600 from these data and apply the SWT, the SWTi, the DTCWTi and the DTCWTi-bi algorithm for denoising.

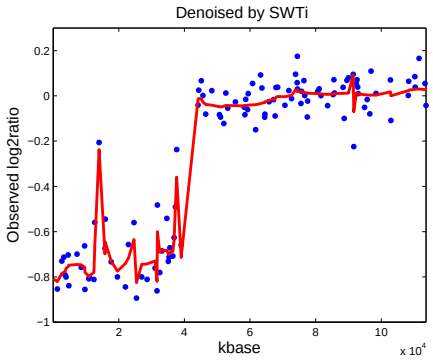
In Fig. 8, the number copy is from zero to two. With the copy two, from $4.43\text{x}10^4$ $kbase$ to the end of this chromosome, Fig. 8(c) and (d) show that the DTCWTi and DTCWTi-bi give smoother denoised signal than the SWT and the SWTi. With the copy zero and the copy one, the performance of the DTCWTi-bi denoising method in Fig. 8(d) is the best when compared with Fig. 8(a), (b) and (c). From the above figures, it is easy to see that DTCWTi and, specially, DTCWTi-bi perform better than SWT and SWTi in denoising of real CGH data.

V. Conclusions

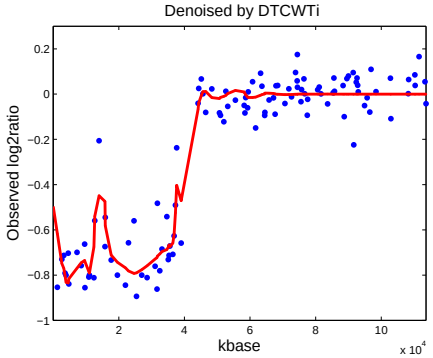
In this paper, we explored the dual-tree complex wavelet transform method with the bivariate shrinkage estimator in array CGH data denoising study. The unequal spacing of probes on the chromosome is taken into account. In the simulation situations, the denoising results from DTCWTi-bi are much better (improve $17.8\% - 43\%$) than the previous methods in terms of the root mean squared error measurement at different noise levels. Furthermore, we also demonstrate our method by using the real array CGH data. In our future work, we will develop a smoothing and segmentation combinatorial algorithm to improve the aberration regions identification from DNA copy number data.



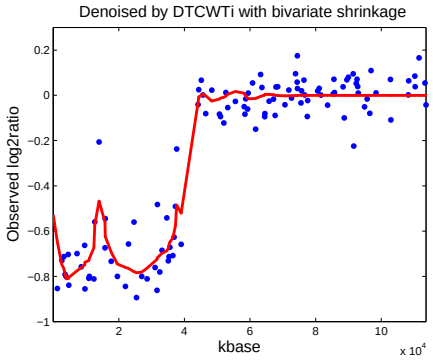
(a)



(b)



(c)



(d)

Fig. 8. The wavelet denoising results of array CGH data on chromosome 9 in the real signal MPE600 using (a) SWTi, (b) SWTi, (c) DTCWTi and (d) DTCWTi-bi.

REFERENCES

[1] P. Eilers and R. de Menezes, "Quantile smoothing of array cgh data," *Bioinformatics*, vol. 21, p. 1146C1153, 2005.

[2] L.Hsu, SG.Self, D.Grove, T.Randolph, K.Wang, JJ.Delrow, L.Loo, and P.Porter, "Denoising array-based comparative genomic hybridization data using wavelets," *Biostatistics(Oxford,England)*, vol. 6, no. 2, pp. 211–226, 2005.

[3] R. Coifman and D. Donoho, "Translation-invariant de-noising," *Wavelets and Statistics*, vol. 103 of Lecture Notes in Statistics, pp. 125–150, 1995.

[4] N. G. Kingsbury, "Image processing with complex wavelets," *Phil. Trans. Royal Society London A*, vol. 357, no. 1760, pp. 2543–2560, Sept 1999.

[5] ——, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, May 2001.

[6] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov 2005.

[7] D. Donoho and I. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

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

[9] I. Johnstone and B. Silverman, "Wavelet threshold estimators for data with correlated noise," *Journal of the Royal Statistical Society*, no. 59, pp. 319–351, 1997.

[10] S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans.Image processing*, vol. 9, pp. 1532–1546, Sept.2000.

[11] L. Sendur and I. Selesnick, "Bivariate shinkage function for wavelet-based denoising exploiting interscale dependency," *IEEE Transaction on Signal Processing*, vol. 50, no. 11, November 2002.

[12] H. Willenbrock and J. Fridlyand, "A comparison study: applying segmentation to array cgh data for downstream analyses," *Bioinformatics*, vol. 21, no. 22, pp. 4084–4091, 2005.

[13] Y. Wang and S. Wang, "A novel stationary wavelet denoising algorithm for array-based dna copy number data," *International Journal of Bioinformatics Research and Applications*, vol. 3, no. 2, pp. 206 – 222, 2007.

[14] A. Kovac and B. W. Silverman, "Extending the scope of wavelet regression methods by coefficient-dependent thresholding," *Journal of the American Statistical Association*, vol. 95, no. 449, pp. 172–183, 2000.

[15] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2006.

[16] A.M.Snijders, N.Nowak, R.Segraves, S.Blackwood, and *et al.*, "Assembly of microarrays for genome-wide measurement of dna copy number by cgh," *Nature Genetics*, vol. 29, p. 263264, 2001.