A Blind-Deconvolution Approach for Chromatographic and Spectroscopic Peak Restoration

S. Sarkar, P. K. Dutta, and N. C. Roy

***Abstract*—In this paper a blind deconvolution method is proposed which allows the deconvolution of broad-band signals that pass through a low-pass device by combining homomorphic deconvolution and Van Cittert iterative deconvolution techniques. The method has been successfully applied on both synthetic and experimental data obtained from chromatography and deep level transient spectroscopy (DLTS).**

***Index Terms*—Chromatography, homomorphic deconvolution, iterative deconvolution, peak restoration, spectroscopy, system response.**

I. Introduction

In many experiments, the signal recorded by the instrument is modified and degraded. There can be two sources of degradation: one is physical inertia intrinsic to all measurement systems, which is due to the insufficient dynamic response of an instrument; and the other is that the phenomenon of interest may not be directly observable due to the nature of the process. Peak broadening of chromatographic signals can be classified in the first group, whereas DLTS is an example of the second. The peak-restoration process thus becomes a deconvolution procedure of the original wide-band data from the smoothed output. There are numerous deconvolution algorithms based on direct or iterative techniques [1], [4]. For all these methods, the knowledge of the system response-function is necessary. One way to estimate the system or instrument response function is to generate a nonoverlapping peak, e.g., in the case of chromatography by injecting a pure ingredient and taking the output as a standard response function [2]. However, it is difficult to choose one particular output as a standard, for chromatography or DLTS.

In this paper, a blind deconvolution technique has been proposed to avoid the problem of standardization. The first step of the algorithm is, therefore, to recover the system response function with the help of homomorphic signal processing techniques [6], [8]. The second step is to apply iterative techniques to recover the chromatographic or spectroscopic peaks. A blind deconvolution, based solely on homomorphic deconvolution, was proposed by Stockham [13] and was applied for restoring old acoustic recording and blurred images. Although theoretically homomorphic deconvolution can separate out both the signals because of their almost nonoverlapping frequency spectrums, experiments show that it is the system response function that can be recovered more reliably in the presence of noise. Furthermore, iterative techniques are generally superior to direct techniques because of their lower requirement of preconditioning and better performance [1], [2].

Manuscript received February 8, 1993; revised December 17, 1998.

The authors are with the Department of Electrical Engineering, Indian Institute of Technology, Kharagpur-721 302, India.

Publisher Item Identifier S 0018-9456(98)09921-5.

II. Theory

This section is divided into two distinct parts:

1) system response recovery;
2) iterative deconvolution with the knowledge of system response.

A. System Response Recovery

Chromatographic or spectroscopic output can be modeled by a convolution

$$X = G * H \tag{1}$$

where G is the response function of the measurement system of a slowly varying physical process and of Gaussian or semi-Gaussian nature. H is the original wide-band signal of interest, which may comprise one or several delta functions.

The blind deconvolution problem can thus be stated as to find G or H or both knowing only X, the final output from the measurement system. The problem can not be solved for general cases without some restrictions and reservations.

1) The signals should posses different frequency characteristics. The more they are separated, the better the reconstruction will be.
2) All the signal components must have finite power in the frequency band of interest.
3) The convolution operation should be a good model for the output, at least for the same run.

Homomorphic Deconvolution: The complex cepstrum $\hat{X}(w_p)$ of the signal X is defined as the inverse Fourier transform of $\ln[X(jw)]$

$$\begin{aligned}\hat{X}(w_p) &= F^{-1}[\ln[X(jw)]] \\ \hat{X}(w_p) &= F^{-1}[\ln[G(jw)]] + F^{-1}[\ln[H(jw)]]\end{aligned} \tag{2}$$

where $[w_p]$ is the virtual time domain, popularly known as the cepstrum domain and $\hat{X}(w_p)$ is the cepstrum of X. The convoluted signals of (1) are now linearly separable additive signals in the cepstrum domain. Thus, using linear filtering, various components of the signal can be separated. The filtering operation can be carried out in the frequency domain or on (2) in the cepstrum domain. The properties of

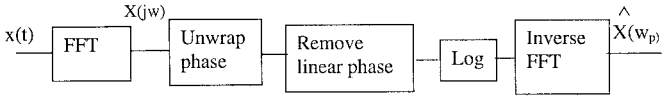


Fig. 1. Forward cepstrum operator $[D^*]$.

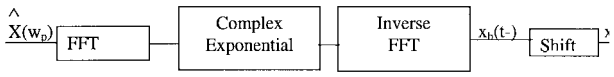


Fig. 2. Backward cepstrum operator $[D^*]^{-1}$.

complex cepstrum that are utilized profitably to implement cepstrum gating are listed below.

1) A low-frequency smoothly varying signal occupies the low-time portion of the cepstrum, smoothly decaying with time.
2) A delta function in time domain has repeated impulses in the cepstrum domain corresponding to its position in the time-domain.
3) Each signal contributes to zeroth component in the cepstrum domain, thus the information about the absolute magnitudes of the signals is lost.
4) The complex cepstrum of a real signal is real.

According to the physical model of the signal presented in this paper, the instrument or physical system response G will occupy the low time portion of the cepstrum and H will be almost separated from G apart from the zeroth component. To recover the machine response G, a low-time cepstrum gating can be applied. A choice of suitable gating function can be made from four different types of gating functions, namely 1) brickwall shape; 2) Gaussian shape; 3) exponential shape; or 4) triangular shape.

The brickwall shaped gating function, while providing the optimum cut off criterion, gives rise to unwanted ripples in the recovered impulse response function; although because of the low-pass nature of the gating function, it is less severe. The other types of gating functions reduce the ripples at the cost of suboptimal cutoff criterion [6].

Once the cepstral gating is over, the filtered signal has to be inverse mapped from the cepstrum domain back to the time domain. The forward cepstrum operator $[D^*]$ is given in Fig. 1, and the inverse mapping operator $[D^*]^{-1}$ in Fig. 2.

For computation of complex cepstrum, the phase has to be continuous in the frequency domain. The authors have used the phase unwrapping algorithm proposed by Tribolet [7], [9]. The linear phase introduced in the unwrapping process has to be removed. The contribution of the linear phase component in the cepstrum domain is a strongly decaying component which subdues the cepstrum of other signals.

Contribution of noise in the cepstrum is greatly complicated because of the nonlinear log operator.

To explain the situation, let the signal be represented as

$$X = G * H + N. \tag{3}$$

The Fourier transform will give us

$$X(jw) = G(jw) \cdot H(jw) + N(jw). \tag{4}$$

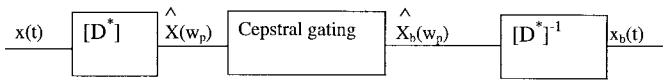


Fig. 3. Block diagram for homomorphic deconvolution.

Now

$$\ln[X(jw)] = \ln[G(jw) \cdot H(jw) + N(jw)]. \tag{5}$$

It has been shown that effect of additive noise is quite far reaching, if not taken care of before the Log operation [6], [8], [10]. Therefore, in cepstrum analysis, use of smoothing windows is almost mandatory. Though windowing does not improve the signal-to-noise ratio, it reduces the spiky nature of the additive noise which produces false peaks. An exponential weighting function has been used here because it commutes through log operation. Furthermore, exponential weighting shifts the unstable poles on the unit circle to the inside [8].

The system response recovery has been preferred over the peak recovery because the signal energy is mainly contained in G and H and can be conceived of as a series of delta functions. Thus in the cepstrum domain, the low-time portion contains most of the energy of the signals, and is better conditioned than the high-time portion with respect to noise power.

Steps: The steps of system response recovery are given in Section III.

B. Iterative Deconvolution for Peak Restoration

Four useful iterative methods elaborately discussed by Crilly [1] are recalled here. They are 1) Van Cittert's method; 2) relaxation-based method; 3) Janssons's method; and 4) Gold's ratio method.

Except for the Gold's ratio method, the others are based on some modification of the Van Cittert's method.

Van Cittert's method: In this method, the iteration equation is given by

$$\hat{h}^{k+1}(i) = \hat{h}^k(i) + b[X - G * \hat{h}^k(i)] \tag{6}$$

with $\hat{h}^0 = bX$, the initial estimate.

The value of b can be chosen by the convergence criterion

$$0 < b < \frac{2}{\max_w [g(e^{w\tau}]} \text{ [1].}$$

For deconvolving real spectroscopic data, constraints are a must, and Van Cittert's method as such can not be used to perform a reliable deconvolution. One important constraint is positivity. In chromatographic or spectroscopic data, the signal is always positive and a positivity constraint P can be incorporated as

$$P[h(n)] = h(n), \qquad \text{if } h(n) \geq 0$$
$$P[h(n)] = 0, \qquad \text{if } h(n) < 0.$$

The other two are the band limited constraint and the finite support constraint. The band limited constraint is defined by an operator

$$B[H(f)] = H(f), \qquad \text{for} - f_c < f < + f_c$$

where f_c is highest frequency the signal contains.

The finite support constraint represents the fact that the signal vanishes at least at $-\infty$ or $+\infty$, i.e.,

$$T[h(n)] = h(n), \qquad \text{for } -\infty < n_a \le n \le n_b < \infty.$$

The incorporation of these constraints in the iterative technique results in constraint-iterative techniques like the Relaxation-based method and Jansson's method [5], which are being utilized by the authors.

Another method which is different from Van Cittert's is Gold's ratio method [1], [3]. The iteration equation is given by

$$\hat{h}^{k+1} = \frac{\hat{h}^k \cdot X}{G * \hat{h}^k}.$$

The primary difference between this method and other iterative methods is that successive estimates are adjusted using ratios instead of sums. The constraint here is implicit i.e., if the estimate h is reasonably close to its true value, then nonphysical components such as negative artifacts could cancel each other out. However, Jansson's method produces better and more stable results than Gold's ratio method.

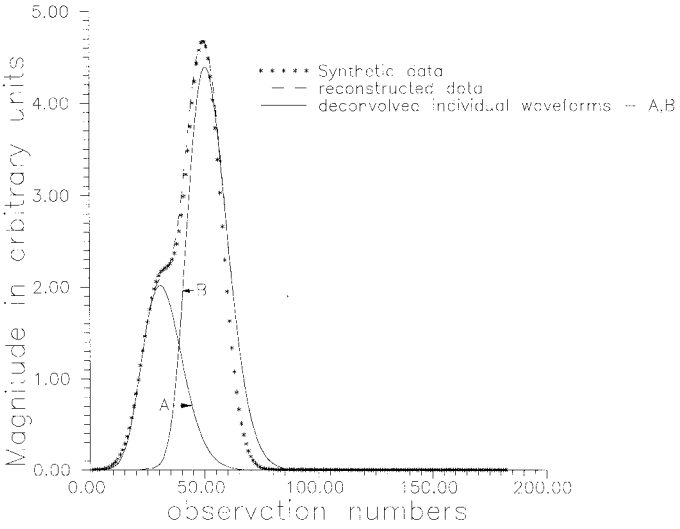
III. Application of Blind Deconvolution in HPLC and DLTS Signals

Before proceeding to the deconvolution of real data obtained from high performance liquid chromatography (HPLC) and deep level transient spectroscopy (DLTS), the algorithm will be summed up here based on the last two sections. The performance of the algorithm will also be tested with synthetic data generated by the convolution of a basic Gaussian wave with a series of impulses.

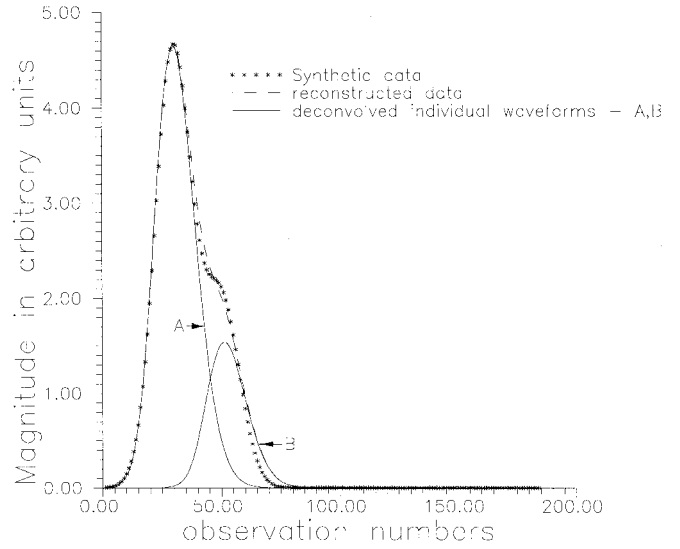
The full deconvolution algorithm with different important parameters is summarized below.

System Response Recovery:

1) Pass the original data through an exponential weighting, the weighting factor α lying between 0.96 and 0.99, the default choice is 0.98.
2) Padd zeros after the signal, the number of zeros being more or less equal to the number of data points. This will help to have a better estimate of unwrapped phase and will reduce the aliasing problem.
3) Compute fast Fourier transform (FFT), unwrap the phase, and compute the shift corresponding to the linear phase component.
4) Before computing the cepstrum, pass the log-magnitude and unwrapped phase through a five-point Gaussian smoother; by default this step is skipped. Step 1) will reduce the effect of additive random noise unless the signal strength is too poor, e.g., below the signal-to-noise ratio of 12 db.
5) If Tribolet's algorithm fails, reexecute from Step 1) by reducing the weighting factor in steps of 0.005.
6) Remove linear phase component and compute inverse Fourier transform.
7) Corresponding to low-time cepstral gating, an energy criterion is chosen to obtain data-adaptive setting. The magnitude of the complex cepstrum is summed up cumulatively starting from the zero position. Whenever it exceeds 50% of the total energy (up to $n/2$ points, n sample points) the point is selected as the default cutoff setting, n_c.
8) Choose the optional cutoff points and the type of the cepstral gating. The default setting is Gaussian side lobe and and the default sigma is $n/2$.
9) Compute fast Fourier transform of the gated output, take complex exponential and compute inverse Fourier transform.
10) After all the inverse operations are over and the shift to get back the reconstructed wave is applied, the

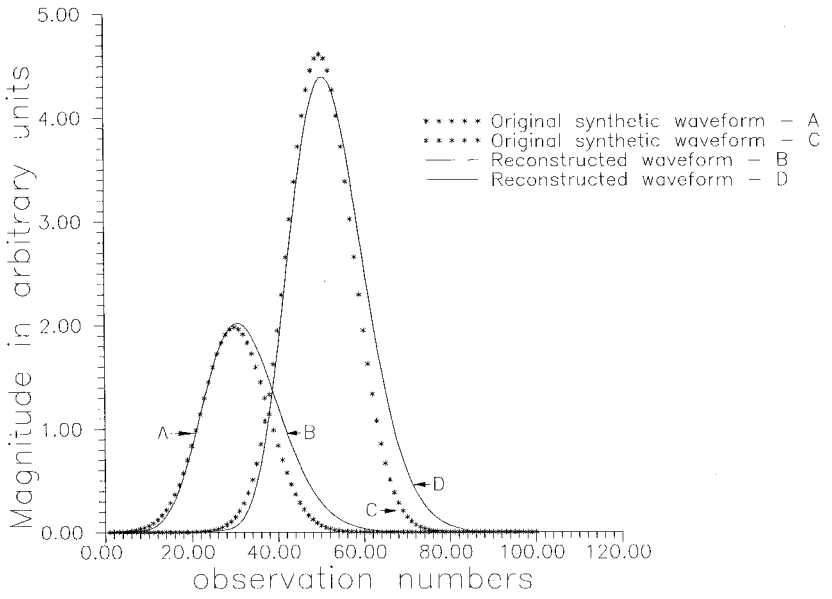


(a)

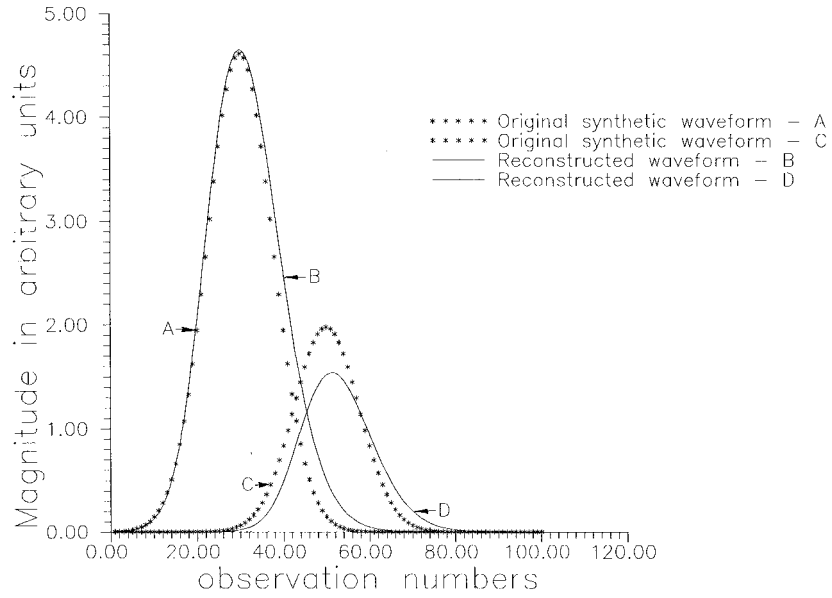


(b)

Fig. 4 (a) Synthetic data (i) and deconvoluted waveforms obtained from the synthetic data. Exponential weighting—0.98. Cepstral gating—5, 4 with Gaussian sidelobe of sigma 4. Iterative technique—Jansson's method (with cross-correlation prefiltering). no. of iterations—2500. (b) Synthetic data (ii) and deconvoluted waveforms obtained from the synthetic data. Exponential weighting—0.97. Cepstral gating—5, 4 with Gaussian sidelobe of sigma 4. Iterative technique—Jansson's method (with cross-correlation prefiltering). no. of iterations—2500.



(a)



(b)

Fig. 5 (a) Original waveforms and reconstructed waveforms of Fig. 4(a) shown together for comparison. (b) Original waveforms and reconstructed waveforms of Fig. 4(b) shown together for comparison.

reconstructed data is passed through an inverse window function, which by default is the exponential weighting function. Thus the basic wavelet or the instrument response is recovered.

Iterative Deconvolution Utilizing the System Response:

1) Choose the length of the basic wavelet, by default it is same as the length of the signal.
2) Choose the type of iterative algorithm. The default choice is Gold's ratio method because of its faster convergence rate.
3) Choose the convergence factor and the maximum number of iterations. The convergence factor varies between 0 and 1. The maximum number of iterations is inversely proportional to the convergence rate. The default settings correspond to Gold's ratio method set at 0.05 and 500, respectively.
4) Give the choice to apply a cross-correlation prefilter [1] on the data; the user has to make the decision by observing the strength of the random noise. The number of iterations required will be more if prefiltering is applied on the data. The default choice is NO. If the choice of cross-correlation prefiltering is YES, the number of iterations will be four times more than that of the default choice.
5) After the peak restoration is over, each of the recovered peaks is convolved with the basic wavelet which will

TABLE I
CHOICE OF ITERATIVE TECHNIQUE AFTER SYSTEM RESPONSE RECOVERY BY HOMOMORPHIC DECONVOLUTION WAS JANSSON'S METHOD WITH CROSS-CORRELATION PREFILTERING

Case No.(SNR 19db)	Error in peak height	r.m.s. error in retaining area	Error in peak position
Fig. 4.2(a) (First peak)	2%	0.19	31-30=1
Fig. 4.2(a) (Second peak)	4.5%	0.16	51-50=1
Fig. 4.2(b) (First Peak)	0.86%	1.01	30-30=0
Fig. 4.2(b) (Second peak)	21%	2.42	51-50=1

resemble the output of a chromatograph or spectrogram when each of the inputs is injected separately.

IV. EXAMPLES

A. Example 1

Two synthetic waveforms are generated by convolving a basic Gaussian wave with an impulsive sequence having two impulses. An additive random noise of Gaussian distribution and zero mean value was mixed with the signal. The SNR was 19 dB for each of the cases and all of them were passed through cross-correlation prefiltering. The convoluted Gaussian waveform and the resultant deconvolved waveforms for the two cases are shown in Fig. 4(a) and (b). For the purpose of comparison, the original and reconstructed waveforms for the two cases are shown in Fig. 5(a) and (b), respectively. Table I gives the normalized rms error and the other errors for the two cases. The poor performance of the algorithm in the last case can be explained by the presence of random noise and the precedence of a stronger peak to a weaker peak. The application of cross-correlation prefiltering further increases the effect of the stronger peak. In each figure, the deconvoluted waveform shown is the result of the convolution of an individual deconvoluted peak with the system response. This way the quality of the reconstruction of both the system response function and the peaks is verified.

B. Example 2: Chromatographic Data

Chromatography uses a separating column to separate different gases or liquids. The time taken to pass through the column is a characteristic of the particular fluid which is known as the retention time of that element. The quality of the resolution in separating the different components is dependent on column length, injection rate, fluid pressure, etc. It is possible to optimize parameters to obtain an optimum performance [11]. The final output of the chromatography will depend on the retention time of the constituents and overall response function of the system

$$X = G_{\text{overall}} * H.$$

There are various system components that will contribute to G_{overall}, e.g,. injection process, column, detector, recorder etc.

$$G_{\text{overall}} = G_{\text{inj}} * G_{\text{col}} * G_{\text{det}} * G_{\text{rec}}.$$

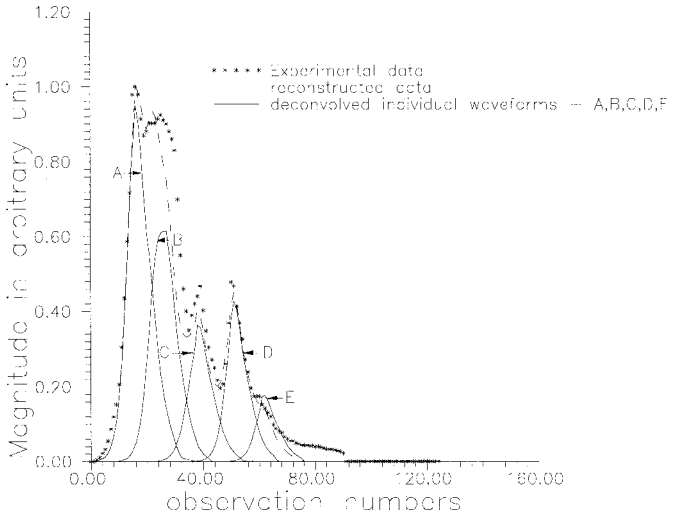


Fig. 6. Experimental data obtained through HPLC and the deconvoluted waveforms shown together. Exponential weighting—0.985. Cepstral gating—8, 7 with Gaussian sidelobe of sigma 6. Iterative technique—Jansson's method (with corss-correlation prefiltering). no. of iterations—2500.

The basic wavelet recovered through homomorphic deconvolution is G_{overall}.

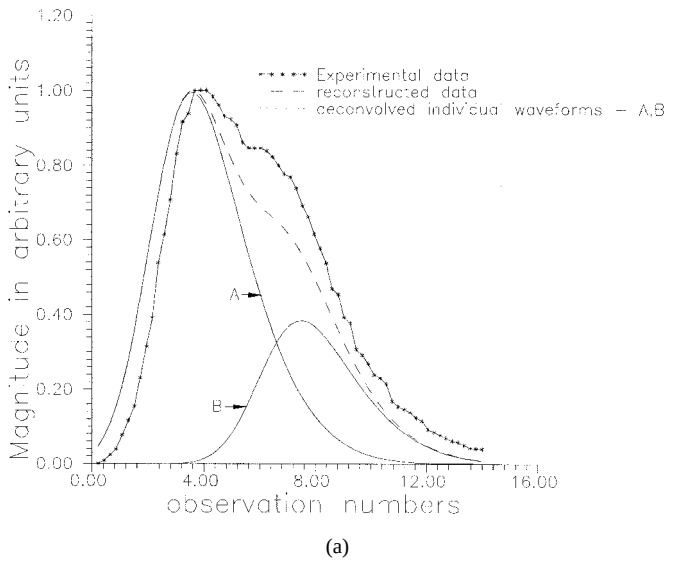
Fig. 6 is the chromatograph of one biotechnology sample, which is formed through a mixture of five or more components. For output like that of Fig. 6, it is very difficult to estimate the actual amount of the constituents by simple methods like exponential curve fitting, although the presence of different components may be detected. The deconvolved individual outputs as well as the reconstructed output from the deconvolved ones are also shown in Fig. 6.

C. Example 3 (Chromatographic Data Taken from [2])

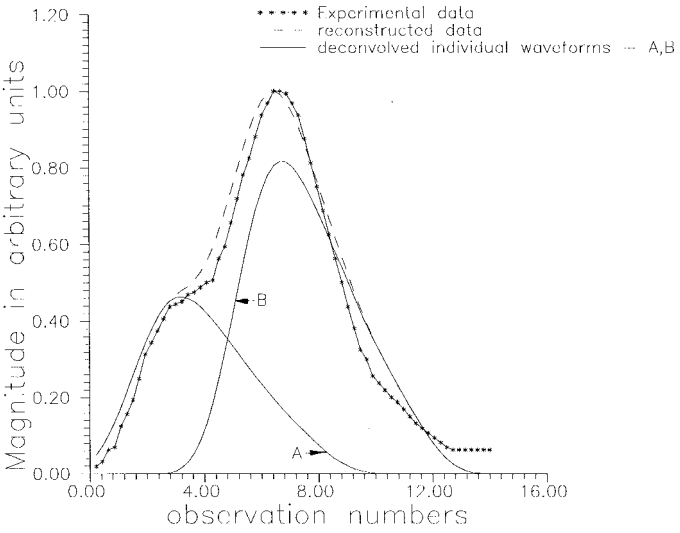
Example 2 amply demonstrates the capability of the deconvolution technique in resolving the peaks. But the result could not be verified directly because the concentration of the individual components was not known exactly. To demonstrate the strength of our algorithm, the data in Fig. 7(a) and (b) has been taken from a paper by Crilly [2] (Figs. 4, 5, and [2, Table III]). The data was extracted by him from a set of controlled experiments. Thus, in these cases, the deconvolved results can be compared to the known concentration of the individual components. Deconvolved peak height is taken to be proportional to the concentration. The peak positions obtained by the authors, without the prior knowledge of system response function, give the same order of accuracy as obtained by Crilly [2] with the knowledge of system response function.

D. Example 4: Data Obtained from Deep Level Transient Spectroscopy (DLTS)

In DLTS the peak broadening effect is caused by the physical process itself. A peak in a typical DLTS output corresponds to a particular energy level. However, due to the excited state of charges, the energy level is no longer a sharp line but a dispersed one. This dispersion causes the peak broadening in the DLTS output. If there are closely spaced energy levels due to trapped charges or impurities, they



(a)



(b)

Fig. 7 (a) Experimental data taken from [2] and the deconvoluted waveforms shown together. Exponential weighting—0.9. Cepstral gating—8, 7 with Gaussian sodelobe of sigma 5. Iterative technique—Jansson's method (with cross-correlation prefiltering). (b) Experimental data taken from [2] and the deconvoluted waveforms shown together. Exponential weighting—0.9. Cepstral gating—8, 7 with Gaussian sidelobe of sigma 5. Iterative technique—Jansson's method (with corss-correlation prefiltering). no. of iterations—200.

TABLE II
COMPARISON OF DECONVOLVED PEAK HEIGHTS WITH KNOWN CONCENTRATION. DATA TAKEN FROM [2]

Case No.	Peak height(%) computed by authors	known concentration	peak height(%) computed by Crilly	peak position in sec.
Fig. 5.1(a) (First peak)	72.3	72.29	74.85	3.6
Fig. 5.1(a) (Second peak)	27.7	27.71	25.15	7.5
Fig. 5.1(b) (First peak)	36.4	38.08	37.98	3.0
Fig. 5.1(b) (Second peak)	63.6	61.92	62.02	6.5

cannot be detected without doing some kind of curve fitting or wave-shape fitting [12]. In material science literature, some

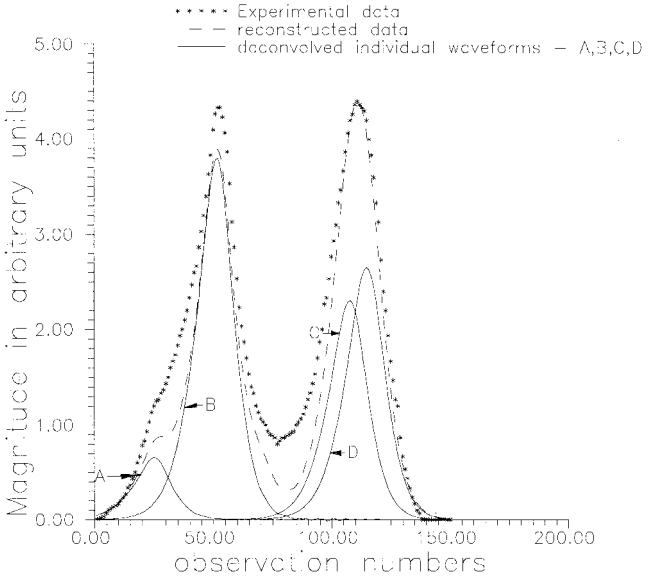


Fig. 8. Experimental data obtained through DLTS and the deconvoluted waveforms shown together. Exponential weighting—0.975. Cepstral gating—8, 7 with Gaussian sidelobe of sigma 6. Iterative technique—Gold's ratio method (with cross-correlation prefiltering). no. of iterations—2500.

methods are available to decompose such signals based on exponential and multiexponential curve fitting. In real cases, a priori physical insight to the device is necessary to resolve the DLTS. The blind deconvolution approach adopted here not only gives accurate peak positions but also produces a reliable estimate of the power of each component. Fig. 8 shows one output from DLTS. It is known from the fabrication process and impurity levels that at least four or five different energy levels exist, but it is hardly noticeable in the graphs. The deconvolved outputs are shown in the same figure. All the hidden peaks have been resolved quite clearly. Fig. 8 also gives the outputs when the peaks are individually convolved again with the basic system response function.

V. CONCLUSION

A blind deconvolution algorithm suitable for chromatographic and spectroscopic peak restoration has been developed. The algorithm will be helpful for data interpretation, particularly for unmanned and fully automated systems. Proper parameter selection is vital for the success of the algorithm, e.g., cutoff points in the cepstral gating and convergence rate in iterative deconvolution. The default choice will give an average performance but not the best for a particular set of data. An experienced user can vary the parameters to get an optimal result.

ACKNOWLEDGMENT

The authors sincerely acknowledge P. Bhattacharya of Biotechnology Centre and D. Pal of Material Science Centre, I.I.T., Kharagpur for providing HPLC and DLTS data. The authors are also thankful to the reviewers for their meticulous review of the paper at different stages and for providing many useful suggestions, which are responsible for the present version of the paper.

REFERENCES

[1] P. B. Crilly, "A quantitative evaluation of various iterative deconvolution algorithms," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 558–562, 1991.

[2] ———, "Numerical deconvolution of gas chromatograph peaks using Jansson's method," *J. Chemometrics*, vol. 1, pp. 79–90, 1987.

[3] ———, "The use of a ratio method as an alternative to constrained deconvolution," *J. Chemometrics*, vol. 5, pp. 85–95, 1991.

[4] A. Bennia and S. M. Riad, "Filtering capabilities and convergence of the Van-Cittert deconvolution technique," *IEEE Trans. Instrum. Meas.*, vol. 41, pp. 246–250, Apr. 1992.

[5] P. A. Jansson, "Deconvolution with application in spectroscopy." New York: Academic, 1984.

[6] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[7] J. M. Tribolet, *Seismic Applications of Homomorphic Signal Processing*, A. V. Oppenheim, Ed., (Signal Processing Series). Englewood Cliffs, NJ: Prentice-Hall, 1979.

[8] D. G. Childers, D. P. Skinner, and R. C. Kemerait, "The cepstrum: A guide to processing," *Proc. IEEE*, vol. 65, pp. 1428–1443, Oct. 1977.

[9] J. M. Tribolet, "A new phase unwrapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing.*, vol. ASSP-25, pp. 170–177, Apr. 1977.

[10] A. P. Petropulu and C. L. Nikias, "The complex cepstrum and bicepstrum, analytic performance evaluation in the presence of noise," *IEEE Tran. Acoust., Speech, Signal Processing.*, vol. 38, pp. 1246–1256, July 1990.

[11] J. R. Conder and C. L. Yong, *Physicochemical Measurement By Gas Chromatography*. New York: Wiley, 1979.

[12] C. W. Wang and C. H. Mu, "A novel technique of analyzing multiexponential transients for DLTS Spectra," *Solid-State Electron.*, vol. 35, no. 12, pp. 1771–1777, 1992.

[13] T. G. Stockham, Jr., T. M. Cannon, and R. B. Ingebretsen, "Blind deconvolution through digital signal processing," *Proc. IEEE*, vol. 63, Apr. 1975.

S. Sarkar received the B.E. degree in electrical engineering from Regional Engineering College, Durgapur, India, in 1991 and the M.Tech. degree in electrical engineering from IIT, Kharagpur, India, in 1992.

Presently, he is with TSIL, USA as a Senior Software Engineer. His current interests are network management and communication protocol.

P. K. Dutta received the B.E. and M.E. degrees in electrical engineering from Bengal Engineering College, Sibpur, Howrah, India, and the Ph.D. degree from IIT, Kharagpur, India, in 1992.

He joined the Department of Electrical Engineering, IIT, Kharagpur, in 1993, and is continuing as an Assistant Professor in the same department. His research interests are signal processing, image processing, and optoelectronic sensors.

N. C. Roy received the B.E. and M.E. degrees in electrical engineering from Jadavpur University, Calcutta, India, and the Ph.D. degree from IIT, Kharapur, India.

He has been with the Electrical Engineering Department, IIT, Kharagpur, since 1968. Presently, he is a Professor in Electrical Engineering. His research interest includes signal processing, imaging, and indentification.