Sparse Non-negative Matrix Factorization and its Application in Overlapped Chromatograms Separation

S.Anbumalar [a,*] , R.Anandanatarajan [b] , P.Rameshbabu [c] ,

[a] Department Electrical and Electronics Engineering , Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry, India 605107.

[b&c] Department of Electronics and Instrumentation Engineering, Pondicherry Engineering College, Puducherry, India 605014

ABSTRACT

A new NMF algorithm has been proposed for the deconvolution of overlapping chromatograms of chemical mixture. Most of the NMF algorithms used so far for chromatogram separation do not converge to a stable limit point. To get same results for all the runs, instead of random initialization, three different initialization methods have been used namely, ALS-NMF (robust initialization), NNDSVD based initialization and EFA based initializations. To improve the convergence, a new sNMF algorithm with modified multiplicative update (ML-sNMF) has been proposed in this work for overlapped chromatogram separation. The algorithm has been validated with the help of simulated partially, severely overlapped and embedded chromatograms. The proposed ML-sNMF algorithm has also been validated with the help of experimental overlapping chromatograms obtained using Gas Chromatography –Flame Ionization Detector (GC-FID) for the chemical mixture of acetone and acrolein.

Keywords

ML-sNMF; modified update for convergence; ALS-NMF (Robust), EFA and NNDSVD based initializations; Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS); Resolution; overlapped and embedded chromatograms; acetone and acrolein mixture

1. INTRODUCTION

The resolution techniques so far applied to many fields to extract information from multivariate data are Evolving Factor Analysis (EFA) [1-2], Fixed Size Moving Window Evolving Factor Analysis (FSMW-EFA) [3], SIMPLe-to-use Interactive Self-modeling Mixture Analysis (SIMPLISMA) [4], Multivariate Curve Resolution (MCR) [5-7], Iterative Target Transformation Factor Analysis (ITTFA) [8-9] and Heuristic Evolving Latent Projections (HELP) [10-11]. All the above methods have their limitations in separation of overlapped chromatograms when they are severely overlapped [12]. Many methods have been developed further to improve the resolution. However, it is still difficult to resolve severely overlapping peaks accurately.

Lee and Seung Suggested Non-negative Matrix Factorization (NMF) in 1999 [13] and provided more algorithms in 2001 [14]. It has been generally applied to image and text data mining. Several works have been carried out in NMF for image analysis [15-18], audio signal separation [19], spectral resolution [20] and signal separation in bio-medical application [21]. A chemometric application of the NMF method is proposed by Li et al. [22] to detect chemical compounds from a chemical substance represented through Raman spectroscopy. Hong-Tao Gao et al., (2005) applied NMF for two components overlapped spectrum resolution [20] in which NMF was applied after imposing constraints based on unimodality, smoothness and sparseness of the data. This inspiring and interesting application of NMF has motivated to test sNMF, which was not used so far for the deconvolution of overlapping chromatograms.

As a first step, overlapped chromatograms were obtained for acetone and acrolein mixture to test the algorithm to be proposed. These are the chemicals identified from the literature that produces overlapping chromatogram due to closer retention time (i.e., 1.127 min for acetone and 1.187 min for acrolein). Acetone is released from automobiles exhaust, tobacco smoke and certain kinds of burning waste materials. Acrolein can enter the environment as a result of burning wood, tobacco, vehicle fuels, overheating of cooking oils and industrial release from waste. Public will be exposed to acetone and acrolein from the environment. If the exposure limit exceeds the allowable value, it may produce ill effects [23]. Hence the existence of these chemicals has to be checked and quantified. Materials containing acetone and acrolein can be analyzed using high-performance liquid chromatography (HPLC) or Gas chromatography (GC) [23]. Due to their similar retention time (RT), this mixture produces overlapping chromatograms under certain instrumental condition and it will be difficult for the chemist to analyze the presence of those components. If some curve resolution techniques are used to separate overlapped chromatograms, then it can be easily resolved and analyzed. It also avoids the lengthier experimental separation procedure of acetone and acrolein [23]. Hence, NMF based curve resolution algorithms have been thought for separating overlapped chromatograms.

Initially, NMF algorithms have been used by the authors [24-26] for the separation of overlapped chromatogram of acetone and acrolein mixture. But, there is no convergence of the algorithm to a stable limit point [27-29]. The algorithm results with different resolved chromatograms (differs in shape or in eluting time) in different runs and same results in some runs due to random initial matrices. It makes the qualitative analysis difficult. Hence, instead of random initial matrices, fixed matrices using some initialization techniques have been proposed in this work. Three such initialization strategies namely ALS-NMF (robust initialization) [30], NNDSVD based initialization [31] and EFA based initialization are used, compared and the best initialization method for each cases taken is identified. To improve the convergence of sNMF algorithm, the multiplicative update has been modified [27] and unimodality and selectivity constraints have been imposed in this proposed modified sNMF (ML-sNMF) algorithm. The proposed ML-sNMF algorithm has been applied for the separation of simulated overlapped chromatograms of acetone and acrolein mixture. The results are encouraging. Then the proposed ML-sNMF algorithm has also been applied on the experimental overlapping

chromatograms obtained for acetone and acrolein mixture using Gas chromatography –Flame Ionization Detector (GC-FID). The results of proposed   ML-sNMF are compared with   that of the existing Multivariate Curve Resolution-Alternating Least Square (MCR-ALS) method. The results show that the proposed   ML-sNMF is suitable for severely overlapped and embedded chromatograms' resolution.

2. MATERIALS AND METHODS

2.1 Experiments

Initially the chromatograms for the acetone and acrolein standards were obtained using GC-FID by injecting the standards separately. Then the acetone and acrolein were mixed in 4 different concentrations. The mixture was vigorously stirred separately under ice cold condition for half an hour and 0.2μL of these mixtures were injected and analyzed using GC-FID.

2.1.1 Instrumental condition

Gas chromatography with Flame Ionization Detector was used with the detector temperature of 150º C. A capillary column (30 m × 0.25 mm ID, BP5) was used with a temperature of 100 º C. The temperature was programmed as 40 º C for 2 min, 5º C/min, 68 º C for 2min, 10 º C/min, 100 º C for 2min. Injector temperature was kept at 150 º c. Nitrogen was used as a carrier gas at a flow rate of 90 ml/min. The sample of 0.2μL volume was injected into the column for analysis. All the graphs were obtained using the software Iris 32 provided by chemito technologies pvt ltd., running on PC with Intel (R) Pentium4 CPU 2.00GHz and 512MB RAM.   The chromatographic data (i.e., retention time t and detector output) has been exported to an ASCII file through Iris 32 software and acquired through MATLAB R2008a software. The above instrumental condition produces overlapping chromatogram for the acetone and acrolein mixture.

2.2 Method

2.2.1 Non-negative Matrix Factorization

 NMF is a technique of decomposing a non-negative matrix A into two non-negative matrices W and H as shown in equation 1.

$$A_{mn} \approx (WH)_{mn} = \sum W_{mk} H_{kn}$$

----------- *(1)*

where k=1 to r < min (m,n).

Usually, r is the number of principal components.

The decomposition is performed so that the product WH should compress and approximates the original data matrix A. W is called basis matrix whereas H is called encoding matrix, of which each column is in one-to-one correspondence with a column in A. Thus, the original data are represented as linear combinations of these basis vectors. Usually r is chosen to be smaller than n or m, so that W and H are smaller than the original matrix A. This results in a compressed version of the original data matrix.

NMF allows only additive operations to provide meaningful decomposition. While PCA is used for decomposition, the two factorized matrices will contain positive and negative entries simultaneously, and these negative components make the result often unacceptable in chemical meanings. Instead, NMF does not allow negative entries in the factorized matrices W

and H, permitting each column of basis matrix H to represent chromatogram in this work. NMF extensions also exist which includes, projective NMF [32], shifted NMF [33], incremental NMF [34] and sparse higher order NMF [35].

2.2.2 sparse Non-negative Matrix Factorization(sNMF)

Although NMF is successful in Matrix Factorization, the NMF model does not impose the sparse constraints. Therefore, it can hardly yield a factorization, which reveals local sparse features in the data A. Related sparse coding is proposed in the work of [36] for matrix factorization. Inspired by the original NMF and sparse coding, the aim of this work is to propose sparse Non-negative Matrix Factorization which imposes the sparse and nonnegative constraint, for chromatograms resolution.Since NMF is an approximation factorization, we need to define the cost function to qualify this approximation. One natural way is to use the divergence function between A and WH. Hence, sNMF algorithm which is given in equation 2 has been applied to find the factors.

$$\sum_{i=1}^{m} \sum_{j=1}^{n} [A_{ij} \log \left(\frac{A_{ij}}{(WH)_{ij}} \right) - A + WH] + \alpha \sum_{q=1}^{k} \sum_{j=1}^{n} H_{qj}$$

----------- (2)

where   α obtained by experience was assumed a positive constant.

A sparse solution to the above constrained minimization can be found by the update rules of W and H given in [24]. Several runs of the above said sNMF algorithm with random initialization will give different answers in each run. Hence, it is important to have efficient and consistent initial matrices W and/or H because the solution and convergence provided by NMF algorithms highly depends on initial conditions.The multiplicative update of sNMF [24] do not assure convergence to a stationary point [27-29] .There are difficulties which exist in the multiplicative algorithm due to zero values in the matrices [27].

2.2.3 Proposed ML-sNMF algorithm

In this proposed ML-sNMF algorithm, the update rules of sNMF [24] have been replaced by the modified update to improve the convergence to a stable limit point. To speed up the convergence and to get same results in all runs, ALS-NMF (Robust initialization) [30], NNDSVD [31] and EFA based initialization strategies have been tested in this proposed algorithm for the separation of overlapped chromatograms.

A   sparse solution to the above constrained minimization (2) can be found by the following update rules of W and H:

i) Assign $\varepsilon > 0$ and $\delta > 0$

ii) Initialize $W_{iq}^{1} \geq O$ and $H_{qj}^{1} \geq O, \forall i, q, j$.   using Robust initialization or   NNDSVD or EFA.

iii) Let p be the number of iterations.

For $k = 1, 2…p$

a) If (W^{k}, H^{k}) is stationary, stop.

Else

$$H_{qj}^{k,n} = H_{qj}^{k} - \frac{\bar{H}_{qj}^{k}}{((W^{k})^{T} W^{k} H^{k})_{qj} + \varepsilon} \nabla_{H} f(W^{k}, H^{k})_{qj}, \forall q, j$$

$$W_{iq}^{k,n} = W_{iq}^{k} - \frac{\bar{W}_{iq}^{k}}{(\bar{W}^{k} H^{k,n} (H^{k,n})^{T})_{iq} + \varepsilon} \nabla_{W} f(W^{k}, H^{k,n})_{qj}, \forall i, q$$

Where

$$\nabla_W f(W,H) = (WH - A)H^T \text{ and } \nabla_H f(W,H) = W^T(WH - A)$$

$$\frac{\bar{H}^k_{qj}}{((W^k)^T W^k \bar{H}^k)_{qj} + \varepsilon} \quad \text{is the step}$$

size
Where

$$\bar{H}^k_{qj} \equiv \begin{cases} H^k_{qj} & \text{if } \nabla_H f(W^k, H^k)_{qj} \geq 0, \\ max(H^k_{qj}, \delta) & \text{if } \nabla_H f(W^k, H^k)_{qj} < 0 \end{cases}$$

ε [37,38] and δ are the small positive numbers assumed to avoid division by zero.

b) Normalize $W^{k,n}$ and $H^{k,n}$ to H^{k+1} and W^{k+1} respectively so that W^{k+1}'s column sum is one.

c) Unimodality and selectivity constraints are additionally imposed into the algorithm so as to use the algorithm for chromatogram separation.

d) The iterative algorithm used has been stopped when the number of iterations p achieves or exceed a predefined maximum number of iteration. It is assumed that there is no need for peak shifting, base line correction and noise filtering.

iv) Calculate the objective function using new updated W and H.

v) Repeat from (iii) to (iv) until convergence is achieved.

The modified multiplicative update overcomes the difficulties that exist in the convergence of sNMF algorithm. Each column of W matrix gives the chromatogram of individual chemical present in the mixture. The algorithm can handle sparseness in the data. The maximum number of iterations has to be specified.

Procedure

Step1: The detector output of each experiment has been taken as an individual column of matrix A. The experimental data were taken for mixtures of different concentration ratios to get a pseudo second order data. (Only one-way data can be taken with the help of detector available with us.) Hence, the shape and area of the overlapped chromatogram is based on its chemical concentration. Here, the shift in position or shape of the chromatogram couldn't be differentiated. Hence, preprocessing was not needed. The main focus of this work is resolving the overlapping components. It is proposed to use ML-sNMF algorithm to perform the deconvolution of a data matrix A.

Step2: The size of matrix W is based on the number of components k in a mixture. The unknown number of components k, in a mixture taken for analysis, was determined by principal component analysis (PCA).

Step3: The W and H matrices have been initialized with Robust initialization (or) NNDSVD based initialization (or) EFA based initialization. The zero elements, if any, in the matrices (A, W or H) are replaced by a small positive number (ε). Hence, the iteration never breaks and the algorithm can converge to minima.

Step 4: The proposed ML-sNMF algorithm is used to decompose the matrix A into W and H matrices.

Step 5: After decomposition, the columns of W matrix are the individual chemical's chromatogram of the mixture taken. The resolved chromatograms are compared with standard's chromatogram and the correlation coefficients are obtained. The signal recovery is calculated using the formula given by

$$Pr_i(dB) = 10 \log_{10} \frac{\sum_{t=1}^{p} [(S_i(t)/std(S_i)]^2}{\sum_{t=1}^{p} [(S_i(t)/std(S_i)) - (R_i(t)/std(R_i))]^2}$$

where S_i - chromatogram obtained for standard chemical (i.e., for acetone, acrolein in the experimental work; the chromatogram of individual component in the simulated mixtures)

R_i - resolved signal;

Std - standard deviation;

$Pr_i(dB)$ -- closeness of ith resolved signal power with that of ith standard signal;

i -- 1,2...n ; where 'n' is the number of components in the mixture.

t -- 1, 2...p; where 'p' is the end point of the chromatogram.

$S_i(1)$ = detector output value at the starting point (t=1) of the chromatogram for the ith standard chemical

$S_i(p)$ = detector output value at the ending point (t= p) of the chromatogram for the ith standard chemical

$R_i(1)$ = detector output value at the starting point (t=1) of the chromatogram for the ith resolved chemical

$R_i(p)$ = detector output value at the ending point (t= p) of the chromatogram for i^{th} resolved chemical

The signal recovery says about how far the resolved signal deviates from its standard with respect to peak retention time point as well as in shape. All the above procedure was implemented using Matlab software.

Initially, the algorithm is tested on the simulated chromatograms of three cases, i.e., partially overlapped, severely overlapped and embedded peaks. One dimensional, GC-FID overlapped chromatograms containing tailing peaks (so that, it can be similar to the experimental chromatograms) of acetone and acrolein were simulated using cross product multiplication of Gaussian functions. The proposed ML-sNMF algorithm has been performed to deconvolute the data matrix A of simulated overlapping and embedded chromatograms.

Finally, the proposed ML_sNMF algorithm is applied on the experimental severely overlapping chromatograms of acetone and acrolein mixtures. All the results are compared with that of MCR-ALS method.

3. RESULTS AND DISCUSSION

The chromatograms obtained for standards, acetone and acrolein, are shown in Fig.1 and Fig.2 respectively. The mixed solution of acetone and acrolein at four different concentrations were prepared.
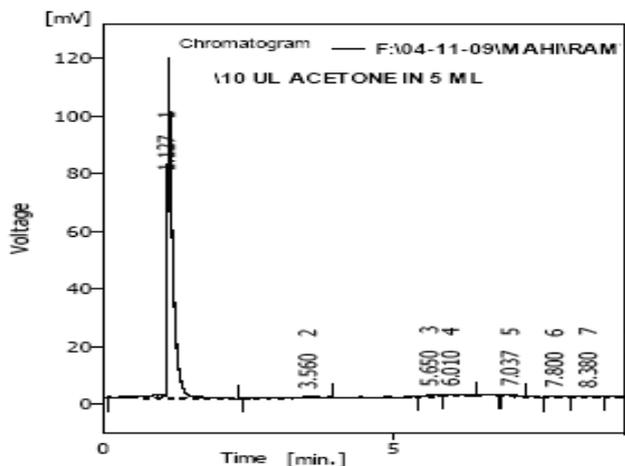
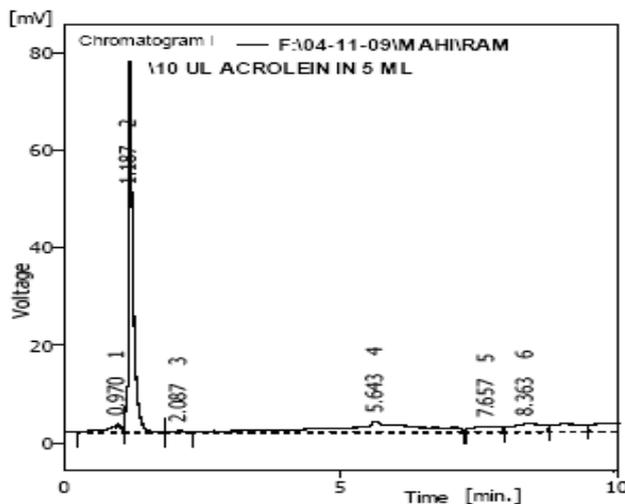**Fig.1. Real experimental chromatogram of acetone standard. Retention time of acetone 1.128min.**

Fig.2. Real experimental chromatogram of acrolein standard. Retention time of acrolein 1.188 min.

The mixer concentrations were i) 2 µL acrolein and 8 µL acetone, ii) 4µL a acrolein and 6µL acetone , iii) 10µL acrolein and 2µL acetone, and iv) 6µL acrolein and 4µL acetone respectively. The chromatograms obtained are as shown in Fig.3 to Fig.6**.** It is found that the chromatograms are overlapped due to close retention time of acetone and acrolein. The chromatogram has some unknown peaks due to some impurities in the chemical mixture. But, focus is only on the overlapping region (from 1 min to 2 min), which has to be resolved to separate severely overlapped acetone and acrolein, The remaining parts of the chromatogram are not considered in this work for analysis.
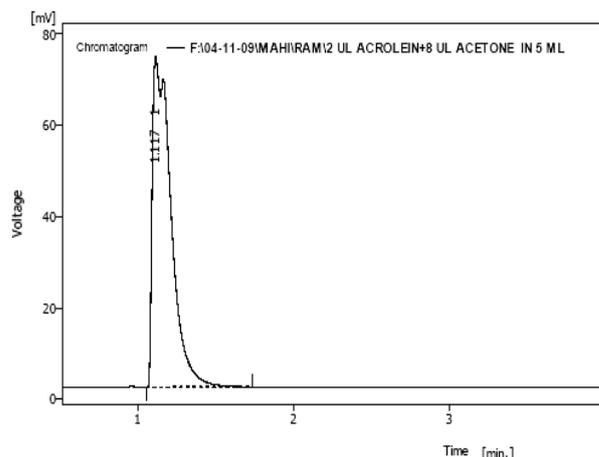
Fig.3. Real experimental severely overlapped chromatogram of acetone and acrolein mixture (2 µl acrolein and 8 µl acetone) .
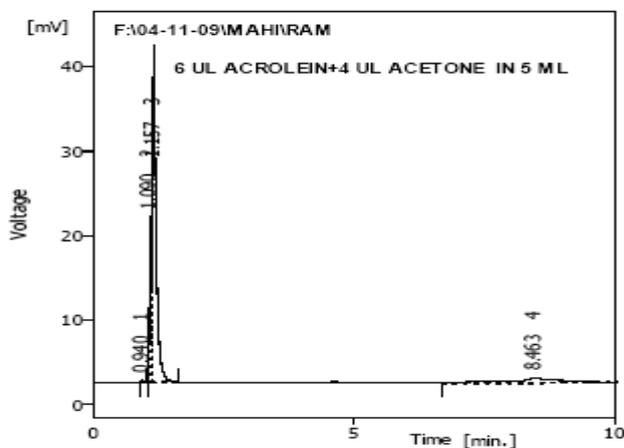
Fig.4. Real experimental severely overlapped chromatogram of acetone and acrolein mixture (6 µl acrolein and 4 µl acetone).
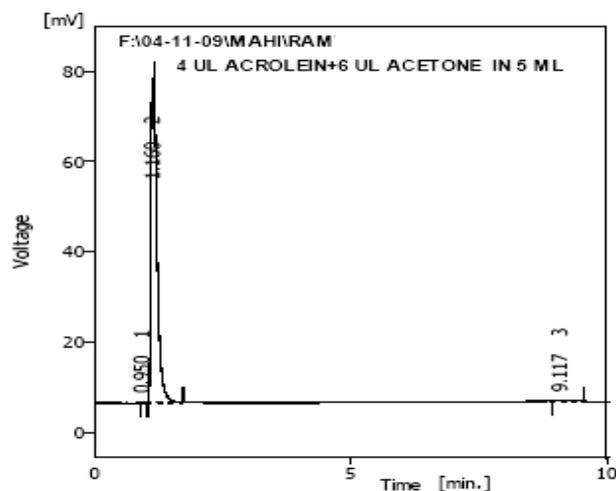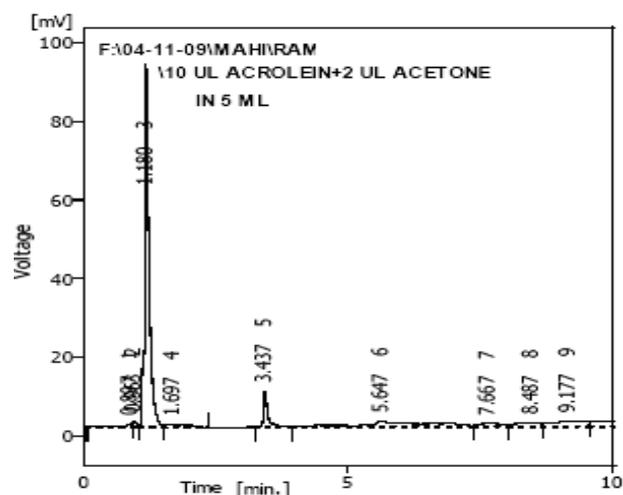

Fig.5. Real experimental severely overlapped chromatogram of acetone and acrolein mixture (4 µl acrolein and 6 µl acetone).

Three different strategies namely EFA, NNDSVD and Robust initialization are used to estimate the initial matrices of the proposed ML-sNMF algorithm. Then the algorithm was used to resolve the overlapped chromatograms.

3.1 Simulation study
3.1.1 Simulation study using proposed ML-sNMF algorithm with EFA based initialization

Initially, the EFA based initialization has been used to initialize the matrices of the proposed ML-sNMF algorithm and used to resolve partially overlapped, severely overlapped and embedded chromatograms. The results are shown in Fig.7- Fig.9. The resolved chromatograms are compared with standard's chromatogram and its correlation coefficient and signal recovery are given in Tables 1, 2 and 3.

Fig.6. Real experimental severely overlapped chromatogram of acetone and acrolein mixture (10 µl acrolein and 2 µl acetone).

3.1.2 Simulation study using proposed ML-sNMF algorithm with NNDSVD based initialization

Secondly, the proposed ML-sNMF algorithm with NNDSVD based initialization has been used for overlapped chromatogram separation.The proposed ML-sNMF algorithm (NNDSVD based initialization) was applied on the simulated partially overlapped, severely overlapped and embedded chromatograms for separation. The unresolved and the resolved

chromatograms are shown in Fig.10- Fig.12. The resolved chromatograms are compared with standard's chromatogram
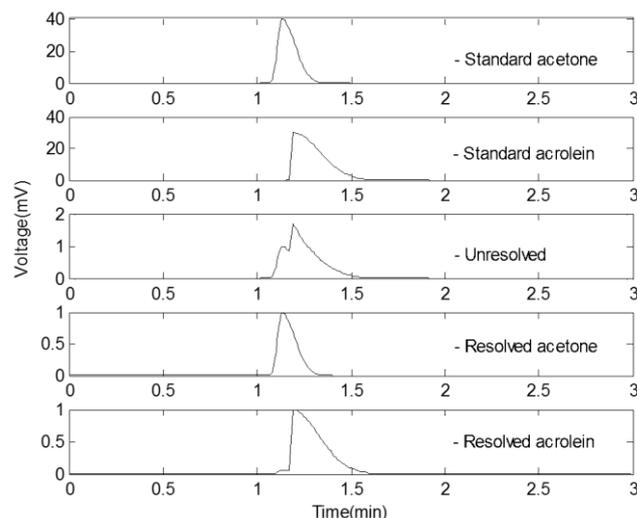

Fig.7 Unresolved (partially overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with EFA based initialization and its correlation coefficient and signal recovery are given in Tables 1, 2 and 3.

3.1.3 Simulation study using proposed ML-sNMF algorithm with Robust initialization

In the third case, the proposed ML-sNMF algorithm has been initialized with Robust initialization method and applied on overlapped chromatograms for separation.

The proposed ML-sNMF algorithm (Robust initialization) was used to separate individual chromatograms from the simulated partially overlapped, severely overlapped and embedded chromatograms. The unresolved and the resolved
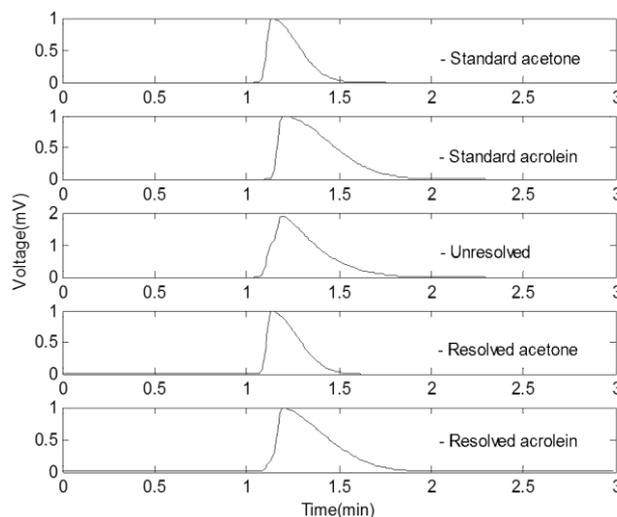

Fig.8 Unresolved (severely overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with EFA based initialization
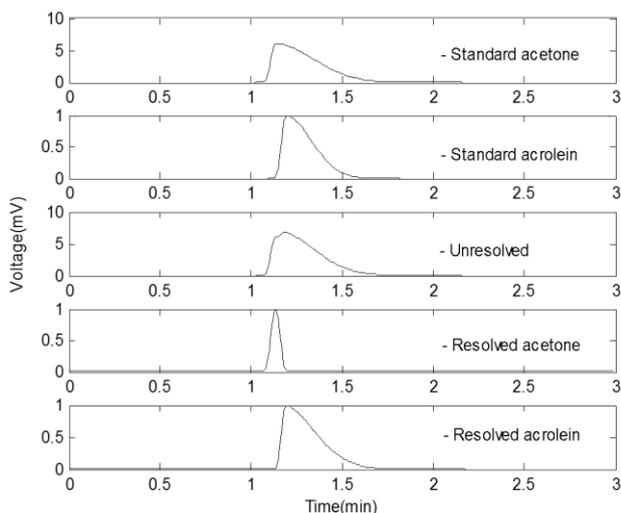
Fig.9 Unresolved (embedded) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with EFA based initialization
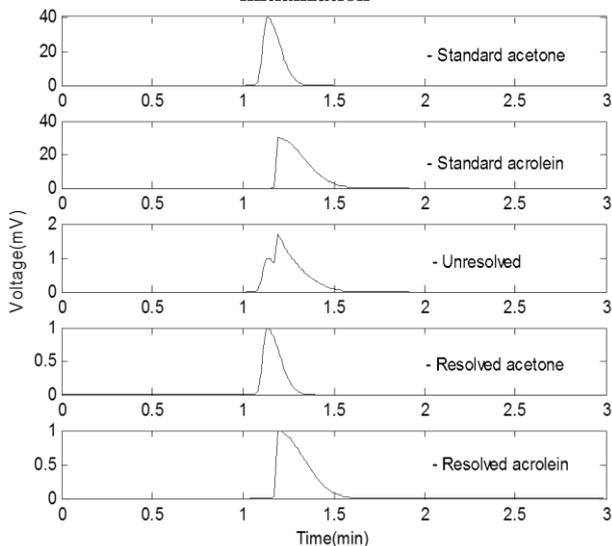

Fig.10 Unresolved (partially overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with NNDSVD based initialization
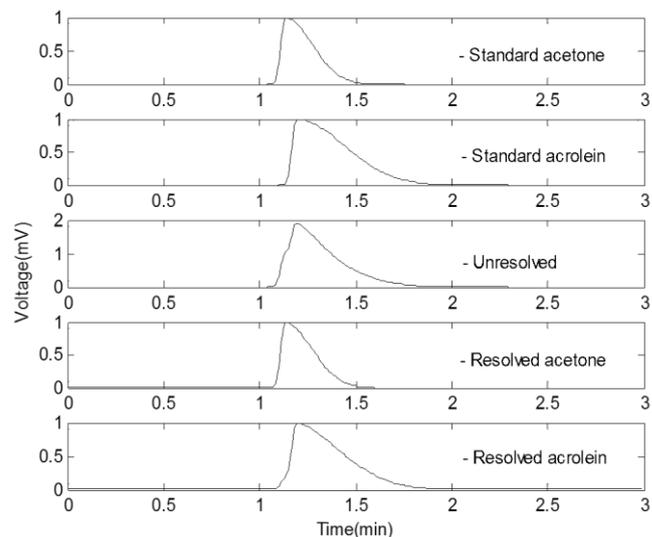

Fig.11 Unresolved (severely overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with NNDSVD based initialization

chromatograms are shown in Fig.13- Fig.15. The resolved chromatograms are compared with standard's chromatogram and its correlation coefficient and signal recovery are given in Tables 1, 2 and 3.

The application of the ML-sNMF algorithm on the simulated overlapped chromatograms of acetone and acrolein mixture show that the proposed method fails to perform better than MCR-ALS method for a partially (slightly) overlapped chromatogram; but performs better than MCR-ALS method for severely overlapped and embedded chromatograms. For the resolution of severely overlapped chromatogram, the ML-sNMF algorithm along with EFA method of initialization outperforms all other methods as shown in Tables 2. In the case of embedded chromatogram, the ML-sNMF algorithm along with Robust method of initialization performs better than all other methods as shown in Tables 3.

3.2 Experimental study using proposed ML-sNMF algorithm

The proposed ML-sNMF algorithm is applied on the experimental chromatogram to resolve acetone and acrolein. Initially, the W and H matrices are initialized with robust
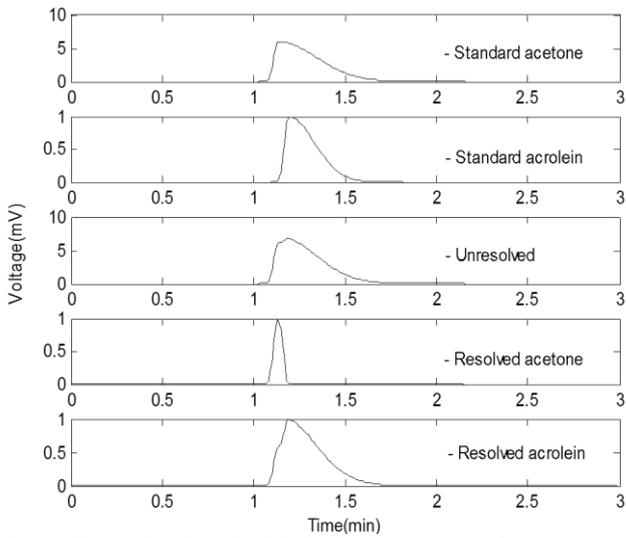
Fig.12 Unresolved (embedded) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with NNDSVD based initialization
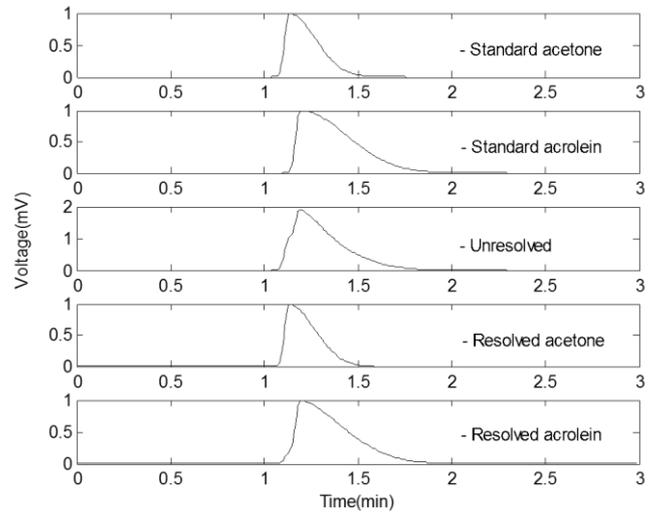

Fig.14 Unresolved (severely overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with Robust initialization
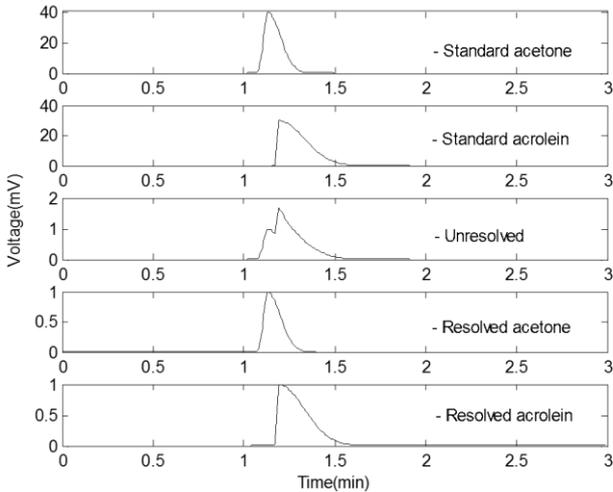

Fig.13 Unresolved (partially overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with Robust initialization

The results are unique for all the runs. The overlapped chromatogram of acetone and acrolein mixture and the resolved chromatograms are shown in Fig.16.
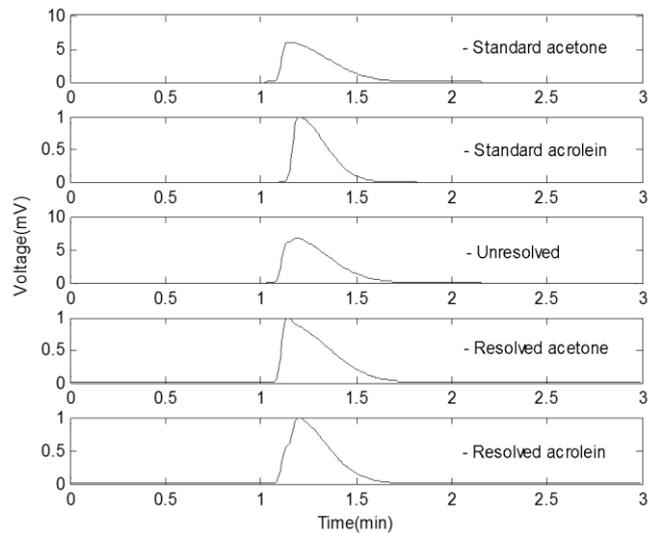

Fig.15 Unresolved (embedded) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF algorithm with Robust initialization

The same dataset are again tested using proposed ML-sNMF algorithm by taking NNDSVD estimate as initial matrices and the resolved results are shown in Fig.17. The results are unique for all the runs. Fig.18 shows the convergence of NNDSVD based initialization for the proposed algorithm. Then the same chromatograms are resolved using proposed ML-sNMF algorithm by taking EFA based estimate as initial matrices and the results are shown in Fig.19 and are unique for all the runs. The correlation coefficients and the signal recoveries for the acetone and acrolein are calculated and given in Tables 4.

The proposed ML-sNMF algorithm is effective for chromatographic resolution. The correlation coefficient is greater than 0.9 and signal recovery is greater than 7 dB always. From the results of different initialization strategies given in Tables 4, it is clear that the NNDSVD based initialization is found to be better for the proposed ML-sNMF algorithm on experimental chromatograms separation.
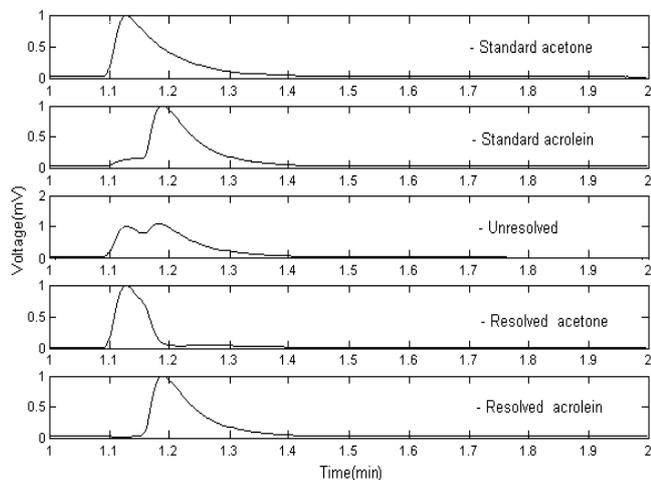

Fig.16 Experimental (severely overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF (Robust initialization)
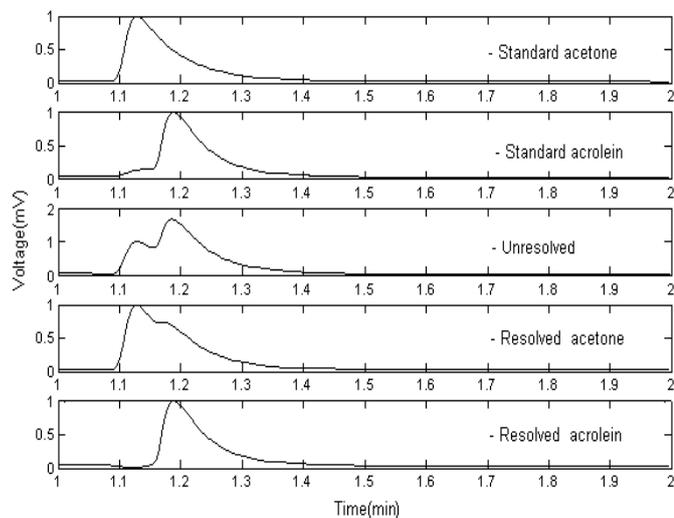

Fig.17 Experimental (severely overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed ML-sNMF (NNDSVD based initialization)

The resolved components are indicating the retention time of chemicals only but the magnitude of the resolved chromatograms has scaling ambiguity**.** Hence the scale of the resolved chromatogram has been normalized and presented in all results. Even though the correlation coefficients are same for the resolved acetone and acrolein in some cases, their signal recoveries differ due to the shift in peak position and shape of the resolved chromatogram from its standard. Hence, the signal recovery seems to be a useful performance measure in chromatogram resolution.

4. CONCLUSION

In this paper, existing MCR-ALS and proposed ML-sNMF algorithms were applied on the experimental overlapped chromatograms of acetone and acrolein mixture. The algorithms were also tested on the simulated overlapped and embedded peaks.

From the results, it can be concluded that the proposed ML-sNMF algorithm converge to a stable limit point and the results are found to be unique, which are the advantages of the proposed algorithm. The proposed ML-sNMF algorithm is better than the existing MCR-ALS method with higher mean correlation coefficient value of 0.9901 as well as with higher mean signal recovery value of 17.8561dB under experimental condition. Under simulated conditions, the mean correlation coefficient value and mean signal recovery value are much higher than the experimental results and thus validating the proposed algorithm. It does not require any peak parameters as an input and hence no prior information about the chemical mixture is needed. Even though the overlapped chromatograms have tailing peaks, the proposed ML-sNMF algorithm resolve it without any difficulty. The proposed ML-sNMF gives better performance than the existing MCR-ALS algorithm for the severely overlapped and embedded chromatograms separation.

5.REFERENCES

[1] H. Gampp, M. Maeder, C.J. Meyer, A.D. Zuberbuhler, Calculation of equilibrium constants from multiwavelength spectroscopic data. III. Model-free analysis of spectrophotometric and ESR titrations, Talanta, 32 (1985)1133-1139.

[2] M. Maeder, Evolving Factor Analysis for the resolution of overlapping chromatographic peaks, Anal. Chem., 59 (1987) 527-530.

[3] H.R. Keller, D.L.Massart, Peak purity control in liquid chromatography with photodiode array detection by fixed size moving window evolving factor analysis, Anal. Chim. Acta, 246(1991) 379-390.

[4] W. Windig, J.Guilment, Interactive self-modeling mixture analysis, Anal. Chem., 63(1991) 1425-1432.

[5] R. Tauler, A.K. Smilde, B.R Kowalski, Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution, J. Chemometr. 9(1995) 31-58.

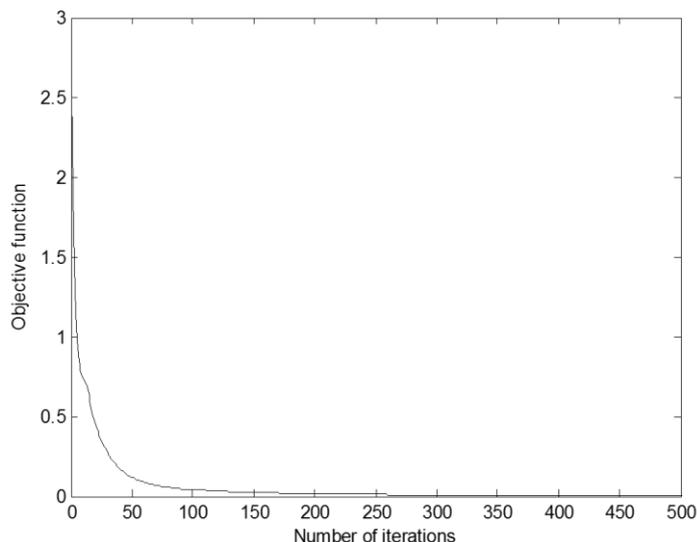[6] R. Tauler, Multivariate curve resolution applied to second order data, Chemom. Intell. Lab. Sys., 30(1995)133-146.

Fig.18 Objective function vs. Number of iterations for proposed   ML-sNMF implementation using NNDSVD based initialization
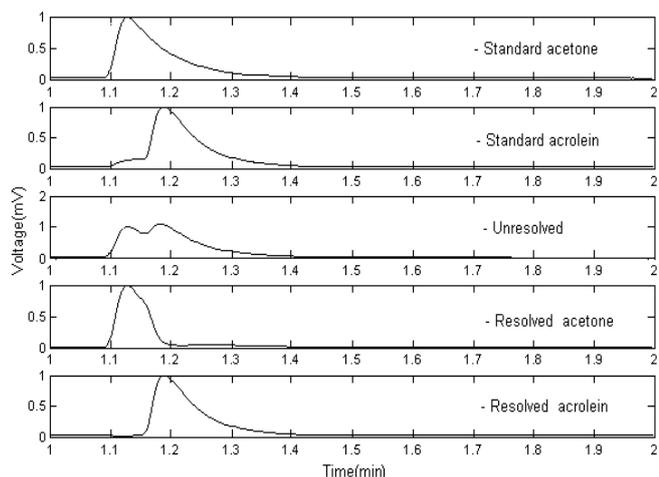

Fig.19 Experimental (severely overlapped) chromatogram of acetone and acrolein mixture and resolved chromatograms using proposed   ML-sNMF (EFA based initialization)

[7] R. Tauler, S. Lacorte, D. Barcelo, Application of multivariate curve self-modeling curve resolution for the quantitation of trace levels of organophosphorous pesticides in natural waters from interlaboratory studies, J. of Chromatogr. A, 730(1996)177-183.

[8] P.J. Gemperline, A priori estimates of the elution profiles of the pure components in overlapped liquid chromatography peaks using target factor analysis, J. Chem. Inf. Comput. Sci. 24 (1984) 206-212.

[9] B.G.M. Vandeginste, W. Derks, G. Kateman, Multicomponent self-modelling curve resolution in high-performance liquid chromatography by iterative target transformation analysis, Anal. Chim. Acta, 173 (1985) 253-264.

[10] O.M. Kvalheim, Y.Z. Liang, Heuristic evolving latent projections: resolving two-way multicomponent data. 1. Selectivity, latent-projective graph, datascope, local rank, and unique resolution, Anal. Chem., 64 (1992) 936-946.

[11] Y.Z. Liang, O.M. Kvalheim, H.R. Keller, D.L. Massart, P. Kiechle, F. Erni, Heuristic evolving latent projections: resolving two-way multicomponent data. 2.Detection and resolution of minor constituents, Anal. Chem., 64 (1992) 946-953.

[12] P. V. van Zomeren, H. Darwinkel, P. M. J. Coenegracht, G. J. de Jong, Comparison of several curve resolution methods for drug impurity profiling using high-performance liquid chromatography with diode array detection, Anal. Chim. Acta, 487(2003)155–170.

[13] DD.Lee, H.Seung, Learning the parts of objects by non-negative matrix factorization,Nature, 401(1999) 788-791.

[14] DD.Lee, H.Seung, Algorithms for non-negative matrix factorization, Adv. Neural Inf.Process. Syst., 13 (2001) 556-562.

[15] P. O. Hoyer, Non-negative Matrix Factorization with Sparseness Constraints, Journal of Machine Learning Research, 5 (2004) 1457-1469.

[16] S. Z. Li, X. W. Hou, H. J. Zhang , Q. S. Cheng, Learning Spatially Localized, Parts-based Representation, International Conference on Computer Vision and Pattern Recognition, (2001) 207- 212.

[17] S. Zafeiriou, A.Tefas, I.Buciu, I.Pitas, Exploiting Discriminant Information in Nonnegative Matrix Factorization with Application to Frontal Face Verification, IEEE Trans. on Neural Networks, 17(3) (2006) 683-695.

[18] I.Buciu, I.Pitas, A New Sparse Image Representation Algorithm Applied to Facial Expression Recognition, IEEE Workshop on Machine Learning for Signal Processing, (2004) 539- 548.

[19] M. N. Schmidt, M. Morup, Nonnegative Matrix Factor 2-D deconvolution for blind single channel source separation, Independent Component Analysis and Blind Signal Separation Lecture Notes in Computer Science, 3889/2006 (2006) 700-707.

[20] Hong-Tao Gao, Tong-Hua Li, Kai Chen, Wei-Guang Li, Xian Bi ,Overlapping spectra resolution using non-negative matrix factorization, Talanta , 66 (2005) 65–73.

[21] Liu Mingyu , Ji Hongbing, Zhao Chunhong, Non negative Matrix Factorization and Its Application in EEG Signal Processing, IEEE Xplore, 978-1-4244-1748-3/08.

[22] H.Li, T.Adali, W.Wang, D.Emge, A.Cichocki, Non-negative matrix   factorization with orthogonality constraints and its application to Raman spectroscopy, J. of VLSI Signal Processing, 48 (2007) 83–97.

[23] Shin-Do Kim, Chang-Hwan Kim, Jin-Su Park , Jeong-Joo Lee, A Study on the Peak Separation of Acetone and Acrolein based on High-Performance Liquid Chromatography (HPLC) Method, Bull. Korean Chem. Soc. 30 (2009) 2011-2016

[24] S.Anbumalar, P.Rameshbabu and R.Anandanatarajan, "Chromatogram separation using Matrix decomposition", *International Journal of Computer Applications*, Vol.27, No.3(2011) 24-32.

[25] S.Anbumalar, P.Rameshbabu and R.Anandanatarajan , "An Algorithm to Resolve Embedded Chromatograms in Chemical Analysis", *International Journal of Computing,*Vol.1,N4 (2011) 610-619.

[26] S.Anbumalar, P.Rameshbabu and R.Anandanatarajan, "Overlapped Chromatograms Separation Using Non-negative Matrix Factorization", *International Journal of Computer and electrical engineering*, Vol.3, N5 (2011) 654-658.

[27] Chih-Jen Lin., Tech. Rep. Information and Support Services Techincal Report, Department of Computer Science, National Taiwan University,2005a.

[28] E. F. Gonzales, Y. Zhang,Technical report, Department of Computational and Applied Mathematics, Rice University, 2005.

[29] Chih-Jen Lin., 2005b,Tech. Rep. Information and Support Services Technical Report ISSTECH-95-013, Department of Computer Science, National Taiwan University.

[30] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, Shun-Ichi Amari, Preprint, 2009 John Wiley & Sons, Ltd.

[31] C.Boutsidis, E Gallopoulos, On SVD-based initialization for nonnegative Matrix factorization, Tech. Report, HPCLAB-SCG-6/08-05, University of Patras, Patras, Greece, 2005.

[32] Z.Yuan , E.Oja, Projective nonnegative matrix factorization for image compression and feature extraction, 14th Scandinavian Conference on Image Analysis, (2005) 333–342.

[33] M.Morup, K.H.Madsen , L.K.Hansen, Shifted Non-negative Matrix Factorization, IEEE Workshop on Machine Learning for Signal Processing, 2007.

[34] S.S.Bucak, B.Gunsel , O.Gursoy, Incremental non-negative matrix factorization for dynamic background modelling, ICEIS 8th International Workshop on Pattern Recognition in Information Systems, 2007.

[35] M.Morup, L.K.Hansen, S. M.Arnfred, Algorithms for sparse higher order non- negative matrix factorization(HONMF), Technical Report, 2006

[36] B.A. Olshausen, D.J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images , Nature, 381(1996)607-609.

[37] P. O. Hoyer, In Proceedings of IEEE Workshop on Neural Networks for Signal Processing, pages 557–565,2002.

[38] J. Piper, P. Pauca, R. Plemmons, and M. Giffin., In Proceedings of AMOS Technical Conference, 2004.