

DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 20 au 24 MAI 1985

UNE NOUVELLE METHODE DE SUPER-RESOLUTION : APPLICATION A LA SEPARATION DE PICS CHROMATOGRAPHIQUES

M. AHMAD*, H. RIX*, H. COLIN** et G. GUIOCHON**

*Laboratoire Signaux et Systèmes, UA814 du CNRS-Université de Nice - 41 Bd Napoléon III - 06041 NICE CEDEX
**Laboratoire de Chimie Analytique Physique, Ecole Polytechnique - 91128 PALAISEAU

RESUME

La séparation de deux pics voisins formant un doublet serré est un problème fondamental qui intervient dans toute analyse de spectre, et en particulier en chromatographie. Dans le cas où les deux pics à séparer sont de même forme et de même largeur, une méthode d'estimation de la distance et du rapport des aires à partir de profils normés par la surface a déjà été proposée et testée en simulation. On applique ici les algorithmes à des chromatogrammes réels. On parvient à séparer des couples d'isomères dont le signal résultant apparaît comme un pic unique. On discute avec l'aide de simulations numériques, le bienfondé de l'hypothèse d'égalité des largeurs pour les pics composants.

SUMMARY

The separation of two strongly overlapped peaks is a fundamental problem occurring in every spectral analysis, e.g. in chromatography. In case of two peaks with equal shapes and equal widths, an estimation method for separation and area ratio by means of normalized profiles has already been proposed and checked by simulation. Here the algorithms are applied to real chromatograms. We are able to separate isomer pairs whose resultant signal appears as a unique peak. The hypothesis of width equality is discussed using simulation on a computer.



UNE NOUVELLE METHODE DE SUPER-RESOLUTION :
APPLICATION A LA SEPARATION DE PICS CHROMATOGRAPHIQUES

1. INTRODUCTION

Le dépouillement des spectres que l'on rencontre dans les procédés d'analyse physico-chimique se heurte au problème fondamental de la séparation des pics qui se chevauchent. Même dans le cas le plus simple où l'on cherche à séparer seulement deux pics voisins, les approches sont nombreuses mais toujours limitées, en particulier lorsque le signal observé est un "doublet serré" c'est-à-dire qui se présente sous l'aspect d'un pic unique. D'un point de vue purement numérique, le problème est mal conditionné : on peut en effet ajuster un pic unique gaussien, ou lorentzien ... sans bruit, à l'aide soit d'un modèle à un pic soit d'un modèle à deux pics voisins [1] avec des erreurs résiduelles du même ordre de grandeur. Dans le cas de signaux réels le bruit ne fait qu'accroître l'incertitude.

En présence d'un signal ayant une allure de pic, le premier problème est donc de détecter s'il est "simple" (1 raie unique dans le cas de la spectroscopie...) ou "multiple" c'est à dire obtenu par sommation de deux ou plusieurs pics simples suffisamment peu résolus. Une réponse à ce problème a déjà été proposée avec la méthode des fonctions de Répartition [2,3,4] qui permet de mesurer de petites variations de forme entre deux signaux transitoires. L'efficacité de cette méthode a été prouvée pratiquement par son application à la chromatographie dans le problème de la détection des pics doubles [5] et dans d'autres problèmes tels que la mesure de l'évolution de forme d'un pic simple avec certaines conditions d'expérience [6,7].

Plaçons nous maintenant en présence d'un pic double : le problème que l'on se pose et que nous traitons ici est celui de l'estimation des paramètres du doublet. Considérons le cas où le signal observé $z(t)$ est la somme de deux pics (signaux unimodaux) de même forme :

$$z(t) = c(y(t) + (k/a)y((t-d)/a)) \quad (1)$$

où : $y(t)$ est le profil d'une des composantes ;

c est une constante ;

d représente la distance du 2nd pic au 1^{er} ;

k représente le rapport de l'aire du 2nd pic à celle du 1^{er} ;

a représente le rapport de largeur du 2nd pic à celle du 1^{er}.

Une méthode d'estimation de d et k (en supposant $a=1$, c'est à dire des pics de même largeur) a été publiée par l'un des auteurs [8]. Le but de cette communication est de montrer les premiers résultats de l'application des algorithmes à des chromatogrammes réels.

L'hypothèse de pics de même largeur ($a=1$) couramment admise en chromatographie pour des pics voisins n'est que la limite obtenue quand la distance tend vers zéro. Afin d'étayer la discussion nous avons, au préalable, étudié en simulation l'influence de cette hypothèse : en faisant varier le paramètre a autour de la valeur 1, on calcule les estimations de d et k qui supposent " $a=1$ ", et on les compare à celles obtenues en estimant simultanément a, d, k .

2. SIMULATION NUMERIQUE

2.1. Rappel de la méthode : cas " $a=1$ "

En reprenant les notations de (1), l'application de la méthode suppose que l'on dispose : 1°/ de l'enregistrement $y(t)$ (échantillonné) de l'un des pics composants, à un facteur multiplicatif près ; 2°/ de l'enregistrement $z(t)$ somme des deux pics composants, avec la même origine des abscisses que $y(t)$. En faisant l'hypothèse " $a=1$ ", la relation (1) s'écrit :

$$z(t) = c(y(t) + ky(t-d)) \quad (2)$$

Les paramètres à estimer sont d et k . On montre [8] que le paramètre d , qui intervient de façon non linéaire, s'obtient très simplement à partir des profils normés

$$\tilde{y}(t) = y(t) / \int_{-\infty}^{+\infty} y(x) dx \quad \text{et} \quad \tilde{z}(t) = z(t) / \int_{-\infty}^{+\infty} z(x) dx$$

Comme indiqué sur la figure 1, d est la différence d'abscisse entre le point d'intersection des courbes \tilde{y} et \tilde{z} et le point de même ordonnée sur \tilde{y} . Connaissant d , la relation (2) permet de calculer k :

$$k = (\tilde{y}(t) - \tilde{z}(t)) / (\tilde{z}(t) - \tilde{y}(t-d)) \quad (3)$$

pour une valeur quelconque de t (différente de l'abscisse du point d'intersection). Pratiquement, en présence d'un signal bruité il est préférable d'utiliser une série de valeurs t_i de t ($i = J1$ à $J2$). Soit \hat{d} l'estimation de d ; si l'on pose :

$$v_i = v(t_i) = \tilde{y}(t_i) - \tilde{z}(t_i) \quad (4)$$

$$u_i = u(t_i, \hat{d}) = \tilde{z}(t_i) - \tilde{y}(t_i - \hat{d}) \quad (5)$$

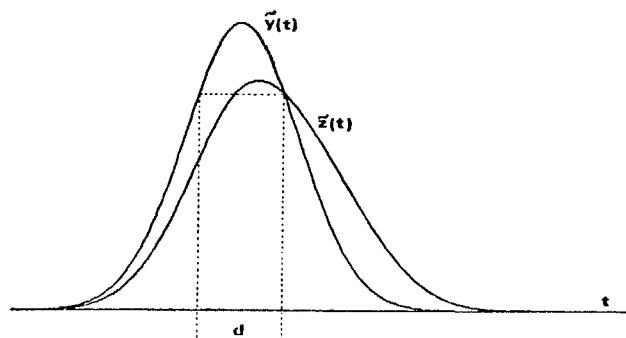


Fig. 1

UNE NOUVELLE METHODE DE SUPER-RESOLUTION :
APPLICATION A LA SEPARATION DE PICS CHROMATOGRAPHIQUES

On obtient l'estimation \hat{R} de k associée à \hat{d} en calculant la droite de régression $v = \alpha u + \beta$ des v_i par rapport aux u_i ; on a

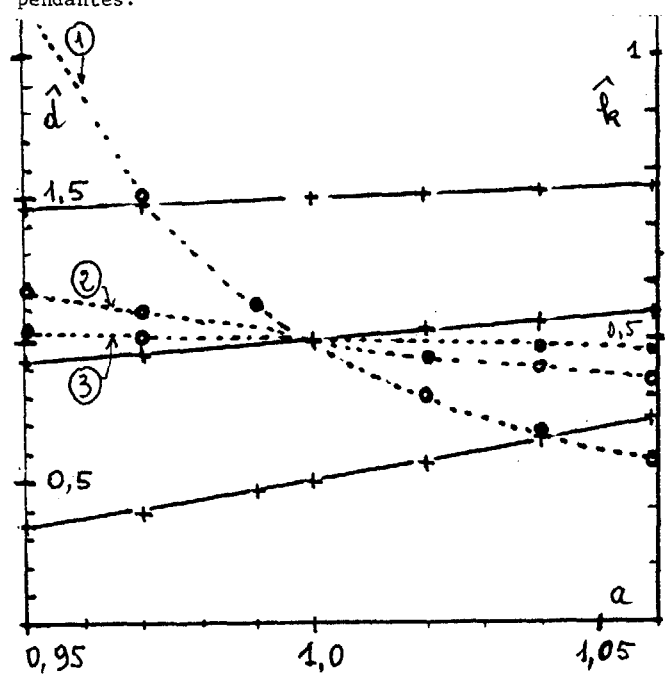
$$\hat{R} = \alpha \quad (6)$$

2.2. Etude du cas "a ≠ 1"

Pour avoir une idée de l'influence d'une fluctuation de largeur de l'un des pics par rapport à l'autre nous avons fait varier a de quelques % de part et d'autre de 1, tout en appliquant la méthode précédente qui suppose $a=1$. La figure 2 illustre les résultats obtenus sur des pics à profil gaussien pour $k=0,5$ et pour trois valeurs de d : 0,5, 1 et 1,5 (l'unité d'abscisse est l'écart-type du pic $y(t)$). En fonction de l'abscisse a on peut lire, sur les ordonnées de gauche les valeurs des estimations \hat{d} de d donnant pratiquement trois droites (trait plein) et sur les ordonnées de droite les valeurs des estimations \hat{R} de k donnant trois courbes (pointillé) correspondant aux trois valeurs données de d (0,5, 1 et 1,5). Ces résultats ont été obtenus en ajoutant aux signaux des séquences de bruit blanc gaussien de façon à obtenir un rapport $S/B=100$. La définition utilisée du rapport S/B est celle des chromatographistes :

$$S/B = h_{\max} / 4 \sigma_B \quad (7)$$

où h_{\max} désigne la hauteur maximale du signal et σ_B l'écart-type du bruit. Enfin chaque point de la figure 2 représente une moyenne obtenue à partir de dix réalisations correspondant à dix séquences de bruit indépendantes.

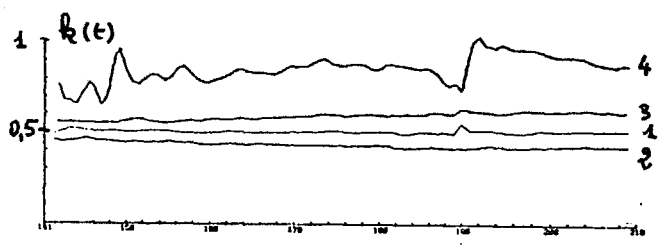


① d=0,5 ; ② d=1 ; ③ d=1,5

Fig. 2

Remarque: 1-L'estimation \hat{d} de d ne dépend pas de k ; on a vérifié de même que la variation de \hat{d} en fonction de a était indépendante de k aux fluctuations du bruit près.

2-On a également vérifié le comportement de la relation (3) permettant de calculer k pour une valeur de t donnée. La figure 3 donne des exemples de variation de $k(t)$ (qui devrait, sans bruit, être la droite $y=k$) en supposant $a=1$.



- 1. $d=1,0$; $k=0,5$; $a=1,00$
- 2. $d=1,0$; $k=0,5$; $a=1,05$
- 3. $d=1,0$; $k=0,5$; $a=0,95$
- 4. $d=0,5$; $k=0,5$; $a=0,95$

Fig. 3

Les résultats de la figure 2 seront utilisés dans la troisième partie pour justifier l'application telle quelle de la méthode (avec $a=1$) aux cas réels étudiés. Néanmoins d'un point de vue purement théorique, ou en vue d'autres applications, il nous a paru intéressant d'étendre la méthode au cas où le paramètre a , laissé variable, était aussi à estimer. Nous avons programmé la procédure d'optimisation décrite dans [8] mais qui n'avait pas fait l'objet de tests numériques. Dans le cas où le paramètre a intervient on doit tirer k de la relation (1) :

$$k = (\hat{y}(t) - \hat{z}(t)) / (\hat{z}(t) - (1/a) \hat{y}((t-d)/a)) \quad (8)$$

L'estimation de k s'obtient alors comme la pente de la droite de régression des $v_i = v(t_i)$ donnés par (4) par rapport aux quantités qu'on continuera à noter u_i , et qui maintenant dépendent de d et a :

$$u_i = u(t_i, d, a) = \hat{z}(t_i) - (1/a) \hat{y}((t_i - d)/a) \quad (9)$$

Etant donné deux estimés \hat{d} et \hat{a} de d et a , on obtient $\hat{R}(\hat{d}, \hat{a})$ en calculant la pente de la droite de régression des $v(t_i)$ par rapport aux $u(t_i, \hat{d}, \hat{a})$. La procédure consiste à rechercher les meilleures valeurs de \hat{d} et \hat{a} (donc \hat{R}) en minimisant (moindres carrés) la dispersion des v_i par rapport à la droite de régression. Les exemples de résultats du tableau ci-dessous, relatifs à un rapport $S/B=100$, montrent l'amélioration, aux prix



UNE NOUVELLE METHODE DE SUPER-RESOLUTION :
APPLICATION A LA SEPARATION DE PICS CHROMATOGRAPHIQUES

de calculs, plus longs, grâce à l'estimation simultanée des trois paramètres a, d, k . Les valeurs estimées sont des moyennes sur 10 essais.

DONNEES	a	0,95	0,99	1,04	0,95	1,04
	d	0,5	0,5	0,5	1,5	1,5
	k	0,5	0,5	0,5	0,5	0,5
k=1	\hat{a}	0,34	0,46	0,64	1,47	1,52
	\hat{k}	1,097	0,56	0,34	0,51	0,48
	\hat{a}	0,952	0,972	1,02	0,956	1,039
	\hat{d}	0,49	0,49	0,48	1,50	1,50
	\hat{k}	0,52	0,51	0,54	0,50	0,50

TABLEAU 1

3. SEPARATION DE PICS CHROMATOGRAPHIQUES REELS

Pour tester les possibilités d'application de la méthode à la chromatographie une première série de manipulations a été effectuée sur trois isomères de position : n , iso et tertio butylbenzène, en chromatographie liquide.

Rappelons que la chromatographie est une technique d'analyse chimique qui repose sur le principe décrit ci-après. On fait passer dans une colonne le mélange à analyser sous forme gazeuse (chromatographie gazeuse) ou liquide (chromatographie liquide). Ce mélange est entraîné par un fluide vecteur qui ne réagit pas avec la colonne. En revanche les différents corps purs du mélange sont retenus par la colonne durant un temps qui est caractéristique du corps. Ce temps est en réalité une grandeur qui fluctue autour d'une moyenne (temps de rétention du corps pur) et le signal observé en sortie du chromatographe est, pour un corps pur unique, un pic qui représente la distribution des temps de séjour des molécules du corps dans la colonne. L'aire sous le pic est proportionnelle à la quantité du corps pur injecté.

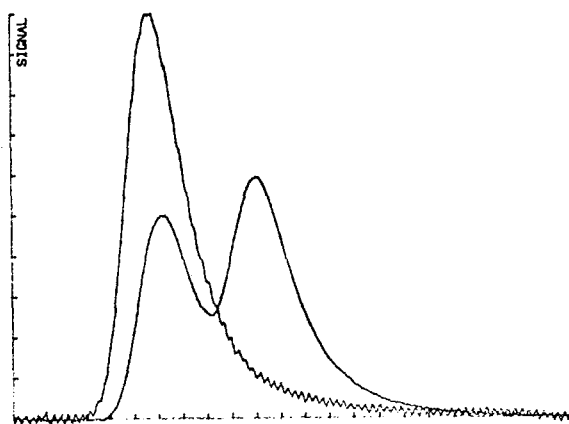
On dispose donc d'enregistrements échantillonnés des isomères seuls et d'enregistrements où l'on a, en proportion variable, mélangé deux isomères : tertio+iso, tertio+n et iso+n. L'ordre de sortie des pics est tertio-Iso-n, et l'on peut mesurer, à partir de trois injections différentes de ces corps purs, les intervalles séparant les temps de rétention ; on trouve :

- entre tertio et iso : d_{T-I} 2,70 sec
- entre iso et n : d_{I-N} 1,58 sec
- entre tertio et n : d_{T-N} 4,28 sec.

Rapportées à la demie largeur à mi-hauteur du pic tertio, ces valeurs représentent respectivement des distances de 0,9, 1,5 et 2,4. D'autre part la variation de largeur entre tertio et n est de l'ordre de 3 %. Compte tenu de la précision sur les données et compte tenu de la figure 2, on peut en première approximation négliger les variations de largeur. Le tableau ci-dessous montre des résultats d'estimation et la figure 4 donne des exemples de tracés normalisés, analogue à ceux de la figure 1, mais pour des pics réels.

Composé	\hat{d} (sec.)	\hat{k}	
tertio + n : (d=4.28) k=1	4.2	0.96	
iso+tertio : (d=2.70)	. k=0.33	2.6	0.38
	. k=0.6	2.6	0.6
	. k=1	2.7	0.95
	. k=1.66	2.8	1.8
iso+n : (d= 5.8)	. k=0.14	1.6	0.15
	. k=0.33	1.6	0.38
	. k=0.6	1.7	0.7

TABLEAU 2



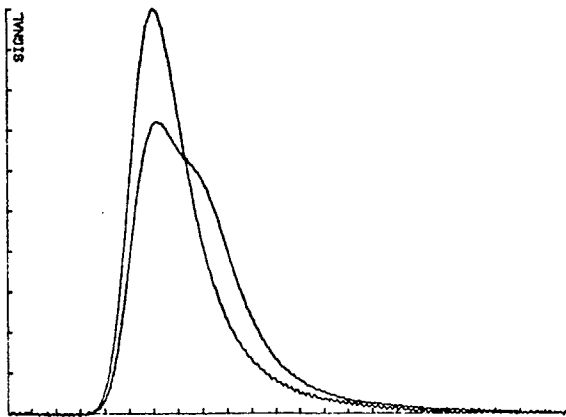
Pic simple : tertio

Pic double : 7,5 tertio + 12,5 n (unité arbitraire)

Fig. 4 (a)

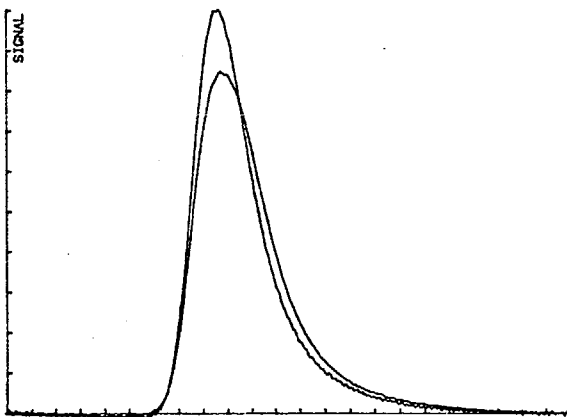


UNE NOUVELLE METHODE DE SUPER-RESOLUTION :
APPLICATION A LA SEPARATION DE PICS CHROMATOGRAPHIQUES



Pic simple : tertio
Pic double : 15 tertio + 5 iso (unité arbitraire)

Fig. 4 (b)



Pic simple : iso
Pic double : 7,5 iso + 2,5 n (unité arbitraire)

Fig. 4 (c)

Ces résultats sont très satisfaisants étant donné les conditions de bruit sur les enregistrements. Plusieurs causes peuvent, dans la pratique, expliquer la différence de précision avec les cas simulés : outre le bruit proprement dit, certains pics parasites étaient présents sur certains chromatogrammes ; la recherche de ligne de base est moins facile pour des pics avec traînée que pour des pics gaussiens ; enfin l'hypothèse de stricte égalité des formes n'est vraie que pour une réponse du détecteur parfaitement linéaire, ce qui n'est pas rigoureusement vrai [7]. Pour éliminer les cas erratiques nous avons utilisé une procédure de sélection grâce aux courbes $k(t)$ analogues à celles de la figure 2, en éliminant les profils où les variations de $k(t)$ étaient trop importantes.

En conclusion, cette première série de manipulations a permis de montrer l'efficacité de la méthode. Les raffinements et l'étude plus fine des

performances nécessitera un plus grand nombre d'expériences couplées à une analyse de forme des pics.

REFERENCES

- [1] A. BARUYA et W.F. MADDAMS, Appl. Spectrosc. 32 (1978) p 563.
- [2] H. RIX et J.P. MALENCE, 6ème Colloque sur le Traitement du Signal et ses Applications, Nice, (1977) p 42 (1).
- [3] H. RIX et J.P. MALENCE, IEEE Trans. SMC 10, n°2 (1980) p 90.
- [4] H. RIX, Thèse d'Etat es Sciences- IMAN- Université de Nice (1980).
- [5] H. RIX et J.P. MALENCE, Journal of High Resolution chromatography and chromatography Communications, 3, N°4 (1980) p.172.
- [6] H. RIX, Journal of chromatography, 204 (1981) p. 163.
- [7] J.L. EXCOFFIER, A. JAULNES, C. VIDAL-MADJAR et G. GUIOCHON, Anal. Chem, 54 (1982) p 1941.
- [8] H. RIX, Signal Processing, 5, N°1 (1983) p. 97.

