

Débruitage d'un cube d'images astrophysique par couplage de l'expansion de Karhunen-Loève et de la transformation en ondelette

Albert BIJAOU¹, Djamel MÉKARNIA², Jean-Pierre MAILLARD³, Céline DELLE LUCHE¹

¹Dpt CERGA, UMR CNRS 6527, OCA
B.P.4229 06304 Nice Cedex 4, France

²Dpt Fresnel, UMR CNRS 6528, OCA
B.P.4229 06304 Nice Cedex 4, France

³Institut d'Astrophysique de Paris, UMR CNRS 8617
98bis Bd Arago, 75014 Paris

bijaoui@obs-nice.fr, mekarnia@obs-nice.fr
maillard@iap.fr, celine_delle_luche@yahoo.fr

Résumé – La spectrographie intégrale de champ consiste dans l'enregistrement de l'information spectrale de tous les pixels d'une image. Des cubes de données d'une dimension typique de $128 \times 128 \times 1024$ éléments résultent actuellement de ces observations. L'ensemble des données du cube peut être exploité en tenant compte des corrélations spatiales et spectrales grâce à la décomposition de Karhunen-Loève. Néanmoins, cette représentation, nécessairement tronquée, conduit parallèlement à supprimer des détails significatifs, présents dans très peu d'images et très localisés spectralement. Pour éviter cette perte d'information nous proposons une approche en plusieurs étapes, avec une première représentation grossière, suivie d'un débruitage spatial et spectral des résidus avec une méthode basée sur la transformation en ondelettes. Avec cette procédure, tous les éléments structuraux significatifs spatiaux et spectraux sont conservés. Le rapport signal sur bruit est amélioré par un facteur de l'ordre de 4.

Abstract – Integral field spectroscopy is the ability to record spectral information for each pixel of an image. Data cubes of typically $128 \times 128 \times 104$ may result from these observations. These cubes can be processed taking into account of the spatial and the spectral correlations using the Karhunen-Loève (KL) expansion. Nevertheless, this decomposition leads to remove significant details present in a small number of spectral images. To avoid this loss, we propose a method based on the residuals of a first KL approximation of each spectral image. These residuals are spatially and spectrally denoised by applying a wavelet transform and added to the restored KL data cube. This analysis clearly shows an enhancement of both spectral and spatial information particularly in case of low signal-to-noise ratio images.

1 La spectrographie intégrale de champ.

La spectrographie intégrale de champ consiste dans l'enregistrement de l'information spectrale de tous les pixels d'une image. De nombreux instruments, parmi lesquels le spectro-imageur par transformation de Fourier (IFTS), ont été construits à cette fin. L'IFTS permet de combiner haute résolution spatiale et haute résolution spectrale dans un grand domaine de longueur d'onde [6]. Des cubes de données d'une dimension typique de $128 \times 128 \times 1024$ éléments peuvent résulter de ces observations. À temps de pose constant, on peut soit obtenir peu d'éléments spectraux, mais avec des images à bon rapport signal sur bruit, soit un plus grand nombre d'images ayant un rapport signal sur bruit dégradé.

En raison de la distribution spectrale des sources et du transfert de rayonnement, les images peuvent varier de manière très significative d'une longueur d'onde (λ) à l'autre, alors que des corrélations très grandes peuvent exister entre des images correspondant à des longueurs d'onde très différentes. À temps de pose constant, un échantillonnage fin en λ conduit à des images à faible rapport signal sur bruit. La résolution en longueur d'onde amenant des informations essentielles sur le champ de vitesse dans les sources astrophysiques étudiées, il est naturel de chercher à exploiter l'ensemble des données du

cube en tenant compte de l'ensemble des corrélations spatiales et spectrales.

La figure 1 illustre les observations multispectrales sur la nébuleuse planétaire NGC 7027 [4].

2 Couplage KL et débruitage en ondelettes.

L'expansion KL. L'expansion de Karhunen-Loève (KL) est mise en œuvre depuis plusieurs décennies pour l'exploitation des données en astrophysique [1] [8] [2] [7]. En représentant chaque image avec une faible partie de l'ensemble de la base, le bruit dans les images peut être réduit. En partant des images \mathbf{V}_n , $n \in (1, N)$, le schéma classique de calcul de l'expansion KL est le suivant :

- Calcul de la moyenne de chaque image $\bar{\mathbf{V}}_n$, ce qui conduit à l'image \mathbf{V}'_n après soustraction.
- Calcul de la matrice de variance-covariance \mathcal{C} des \mathbf{V}'_n .
- Calcul des vecteurs propres de \mathcal{C} \mathbf{u}_i de composantes u_{in} par ordre de valeur propre λ_i décroissante.

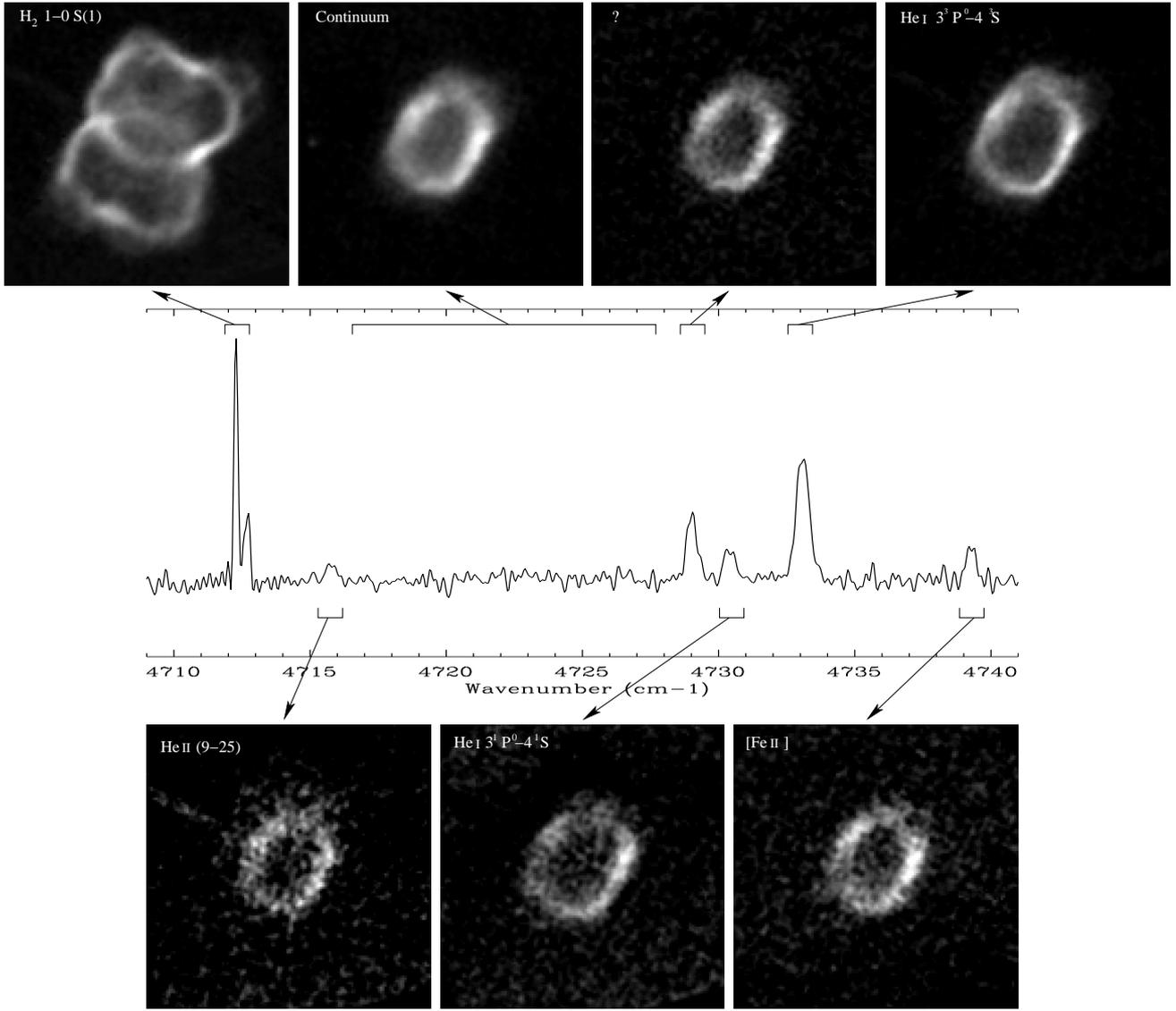


FIG. 1 – Observations multispectrales de la nébuleuse planétaire NGC 7027 obtenues avec l’IFTS [4]. Une filtre autour d’une raie de la molécule H_2 a été utilisée pour ces observations. Le spectre correspond à une zone de forte émission. Les domaines spectraux associés à chaque image sont indiqués. Des grandes variations d’une région spectrale à l’autre peuvent être notées.

– Calcul de la base KL par :

$$\mathbf{S}_i = \frac{1}{\sqrt{\lambda_i}} \sum_n u_{in} \mathbf{V}'_n \quad (1)$$

Nous pouvons écrire la relation :

$$\mathbf{S} = \mathbf{B} \cdot \mathbf{V}' \quad (2)$$

avec :

$$\mathbf{B} = \mathbf{\Lambda}^{-\frac{1}{2}} \mathbf{U} \quad (3)$$

où $\mathbf{\Lambda}$ et \mathbf{U} désignent respectivement les matrices de valeurs et des vecteurs propres. Ceci permet d’écrire :

$$\tilde{\mathbf{V}}' = \mathbf{A} \cdot \mathbf{S} \quad (4)$$

où :

$$\mathbf{A} = \tilde{\mathbf{U}} \mathbf{\Lambda}^{\frac{1}{2}} \quad (5)$$

$\tilde{\mathbf{U}}$ étant la transposée de \mathbf{U} . Le nombre de vecteurs propres pouvant être très inférieur au nombre d’images, un écart existe alors, le résidu moyen étant égal à la somme des valeurs propres

des vecteurs propres non utilisés. En choisissant les plus grandes valeurs propres, on obtient l’erreur minimum à nombre de vecteurs donné.

Les vecteurs propres non utilisés, de faible valeur propre, correspondent à des détails significatifs, présents dans très peu d’images et très localisés spectralement. Cette représentation va donc conduire à les supprimer. Pour éviter cette perte d’information nous proposons une approche en plusieurs étapes.

Représentation KL grossière. On détermine les bases Karhunen-Loève spatiales et spectrales, en se limitant à un faible nombre de composantes. Dans l’exemple présenté nous avons limité à 8 le nombre d’éléments des bases. Les composantes suivantes ne montrent que des détails très localisés.

Le calcul d’une image de base spatiale s’effectue avec l’équation 1. Si on admet que le bruit est blanc, indépendant d’une

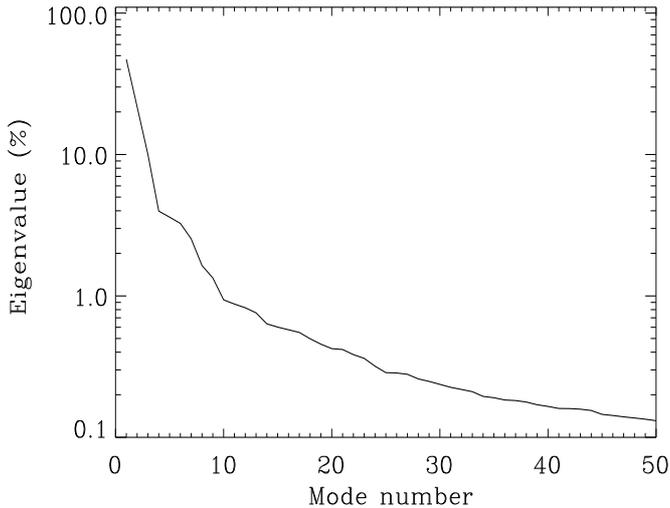


FIG. 2 – Premières valeurs propres de l’expansion KL. L’ordonnée est normalisée en pourcentage par rapport à l’énergie totale.

image à l’autre, la variance de cette image est :

$$\sigma^2(\mathbf{S}_i) = \frac{1}{\lambda_i} \sigma^2 \quad (6)$$

où σ^2 désigne la variance des images.

Les bases KL, qu’elles soient spatiales ou spectrales, sont donc bruitées, d’autant plus que la valeur propre est faible. Ce bruit est transmis à toute reconstruction par l’équation 4. Il est donc indispensable de débruiter ces images.

Pour cela nous avons appliqué une méthode de débruitage par transformation en ondelettes. Nous avons choisi d’appliquer un algorithme à trous [9], avec le seuillage doux de Donoho [5]. Le débruitage est appliqué de manière séparée spatialement et spectralement. Les images obtenues sont re-orthogonalisées, ce qui permet de reconstruire une approximation du cube débruité basée sur les composantes sélectionnées.

Traitement des résidus. Par la reconstruction obtenue avec peu de vecteurs propres et des bases débruitées une partie de l’information spatiale et spectrale a été perdue. Mais, compte tenu que cette information ne correspond qu’à une faible partie de l’énergie, elle est très localisée spatialement et spectralement. De ce fait les différences entre le cube initial et le cube débruité (résidus) peuvent être efficacement débruitées par ce même algorithme basé sur la transformation en ondelettes, successivement spatialement et spectralement. Les petits détails dans les images et dans les spectres qui auraient disparus du fait de la limitation de la base sont restaurés.

Discussion. De ce fait, grâce au débruitage en ondelettes, chaque détail très localisé spatialement et spectralement est restauré, sans accroissement significatif du bruit. On peut bien sûr procéder en effectuant un traitement image par image, ou spectre par spectre, en débruitant avec la transformation en ondelettes. Dans ces conditions, on n’exploite pas les corrélations à grande échelle dans les images et dans les spectres. Expérimentalement les résultats sont moins bons.

On peut aussi chercher à reconstruire l’image en effectuant une expansion de Karhunen-Loève complète. D’une part le nombre de vecteurs propres est excessif (1024), d’autre part très vite les images de la base deviennent très bruitées, ce qui nécessite un débruitage de plus en plus actif. À la reconstruction il n’est pas du tout sûr que des détails significatifs n’aient pas été éliminés, ces détails se décomposant sur la base en éléments non significatifs.

Dans la méthode proposée, il existe une exploitation à la fois des corrélations spatiales et spectrales, par la mise en œuvre du débruitage avec la transformation en ondelettes, mais aussi prise en compte des corrélations à grande échelle spatiale et spectrale, avec l’expansion de Karhunen-Loève.

3 Application.

La figure 2 montre les variations de la valeur propre avec l’indice. Un taux de 80% de l’énergie est contenu dans les 8 premières valeurs propres. Il est clair que les images de base correspondant à des valeurs propres plus grandes qu’une dizaine ne contiennent que très peu d’énergie. Elles ne correspondent qu’à des détails très localisés, la description KL est alors très peu pertinente. L’exploitation de la transformation en ondelettes image par image, ou spectre par spectre, est plus pertinente.

La figure 3 montre le résultat obtenu sur le cube de données associé à la nébuleuse planétaire NGC 7027. Avec cette procédure, tous les éléments structuraux significatifs spatiaux et spectraux sont conservés. Le rapport signal sur bruit semble amélioré d’un facteur de l’ordre de 4.

4 Conclusion.

Le couplage de la décomposition de Karhunen-Loève avec la transformation en ondelettes a été proposé par Starck et al. [10] en appliquant une décomposition KL échelle par échelle. La procédure a été proposée pour un faible nombre d’images. Toutes les composantes sont alors utilisées alors dans le cas des grands cubes de données il n’est pas réaliste de traiter l’ensemble de la base. La procédure proposée permet ainsi d’obtenir un cube débruité exploitant au mieux les inter-corrélations, sans perdre d’information significative.

Le gain apparent en rapport signal sur bruit est très grand, de l’ordre de 4. C’est équivalent à avoir 16 fois plus de photons. Mais il faut nuancer ce résultat, les images résultantes semblent bien nettoyées de leur bruit, en tenant compte aussi bien que possible de toute l’information, mais elles ne sont pas équivalentes à des images obtenues avec 16 fois plus de flux lumineux. L’effet du débruitage est de faire disparaître des variations qui ne semblent pas significatives, compte tenu du bruit, mais cela ne peut faire apparaître des petites variations réelles, mais inférieures au seuil de signification.

L’exploitation simultanée de l’ensemble des images permet, grâce aux corrélations, de faire ressortir des petits détails, qui ne sont pas significatifs sur une image, mais qui apparaissent systématiquement sur un ensemble d’images, ou sur plusieurs longueurs d’onde.

La procédure présentée est donc un moyen d’exploiter d’une manière efficace un cube de données astrophysique. À temps

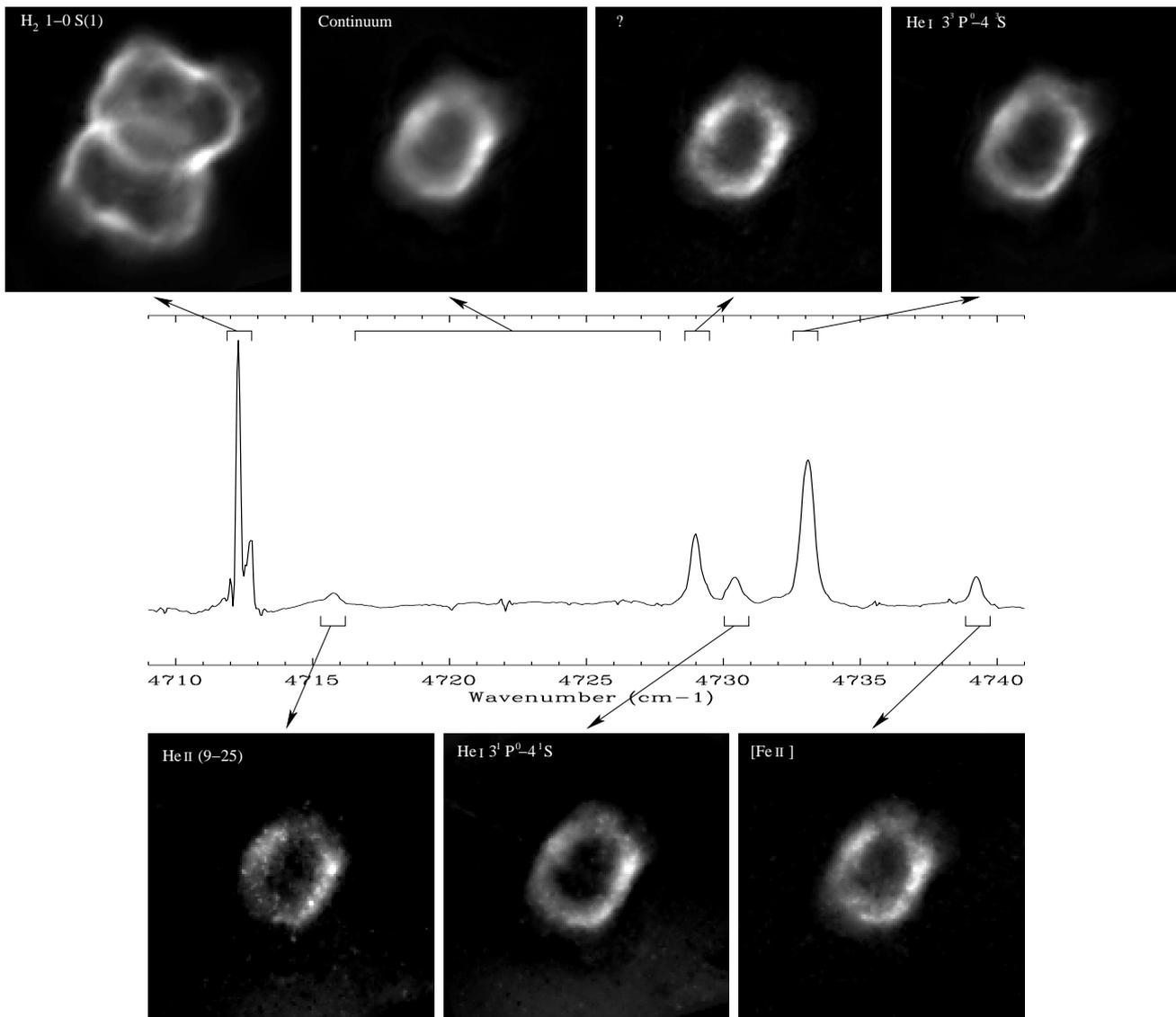


FIG. 3 – Reconstruction du cube de données à partir de la procédure indiquée.

d'exposition constant, cela permet d'augmenter la résolution spectrale, paramètre qui conditionne le plus souvent la qualité de l'information astrophysique.

La décomposition KL a permis de mettre en évidence des artefacts liés à la procédure instrumentale. Les images de bases correspondantes sont bien décorréélées, par construction, mais visuellement un mélange existe entre ce qui provient des artefacts et des sources astrophysiques. Nous avons testé l'analyse en composantes indépendantes [3]. Elle permet de mieux séparer les sources physiques, toutefois le débruitage ne nécessite pas ce degré de sophistication et la mise en œuvre de l'expansion KL, c'est-à-dire de l'analyse en composantes principales, suffit pour cette opération.

Références

- [1] Bijaoui A. 1974, *Astronomy and Astrophysics*, 30, 199.
- [2] Bijaoui A., Doazan V., 1979, *Astronomy and Astrophysics*, 70, 285.
- [3] Comon, P., 1994, Independent Component Analysis, a new concept. *Signal Processing*, 36, pp.287-314.
- [4] Cox, P.; Huggins, P. J.; Maillard, J.-P.; Habart, E.; Morriset, C.; Bachiller, R.; Forveille, T., 2002, *Astronomy and Astrophysics*, 384, 603.
- [5] Donoho D.L.; Johnstone I.M. 1994, Ideal spatial adaptation via wavelet shrinkage, *Biometrika* 81, pp.425-455.
- [6] Maillard, J.-P. 1995, in 3-D Optical spectroscopy methods in astronomy, ed. G. Comte, & M. Marcellin, IAU Colloq. 149, ASP. Conf. Ser., 71, 316
- [7] Murtagh, F., Heck, A., 1987, *Multivariate data analysis*, p. 13, Reidel, Dordrecht.
- [8] Pelat D., 1974, *Astronomy and Astrophysics*, 33, 321.
- [9] Starck J.L., Bijaoui A., 1994, Filtering and deconvolution by the wavelet transform. *Signal Processing*, 35, pp. 195-211.
- [10] Starck, J.-L., Murtagh, F., Querre, P., Bonarel, F., 2001, *Astronomy and Astrophysics*, 368, 730-746.