Decomposition and Classification of Spectral Lines in Astronomical Radio Data Cubes

Vincent Mazet[1], Christophe Collet[1], and Bernd Vollmer[2]

[1] LSIIT (UMR 7005 University of Strasbourg–CNRS),
Bd Sébastien Brant, BP 10413, 67412 Illkirch Cedex, France
[2] Observatoire Astronomique de Strasbourg (UMR 7550 University of
Strasbourg–CNRS), 11 rue de l'Université, 67000 Strasbourg, France
{vincent.mazet,c.collet,bernd.vollmer}@unistra.fr

Abstract. The natural output of imaging spectroscopy in astronomy is a 3D data cube with two spatial and one frequency axis. The spectrum of each image pixel consists of an emission line which is Doppler-shifted by gas motions along the line of sight. These data are essential to understand the gas distribution and kinematics of the astronomical object. We propose a two-step method to extract coherent kinematic structures from the data cube. First, the spectra are decomposed into a sum of Gaussians using a Bayesian method to obtain an estimation of spectral lines. Second, we aim at tracking the estimated lines to get an estimation of the structures in the cube. The performance of the approach is evaluated on a real radio-astronomical observation.

Keywords: Bayesian inference, MCMC, spectrum decomposition, multicomponent image, spiral galaxy NGC 4254.

1 Introduction

Astronomical data cubes are 3D images with spatial coordinates as the two first axis and the frequency (velocity channels) as third axis. We consider in this paper 3D observations of galaxies made at different wavelengths, typically in the radio (> 1 cm) or near-infrared bands ($\approx 10\,\mu$m). Each pixel of these images contains an atomic or molecular line spectrum which we called in the sequel *spexel*. The spectral lines contain information about the gas distribution and kinematics of the astronomical object. Indeed, due to the Doppler effect, the lines are shifted according to the radial velocity of the observed gas. A coherent physical gas structure gives rise to a coherent structure in the cube.

The standard method for studying cubes is the visual inspection of the channel maps and the creation of moment maps (see figure 1 a and b): moment 0 is the integrated intensity or the emission distribution and moment 1 is the velocity field. As long as the intensity distribution is not too complex, these maps give a fair impression of the 3D information contained in the cube. However, when the 3D structure becomes complex, the inspection by eye becomes difficult and important information is lost in the moment maps because they are produced

by integrating the spectra, and thus do not reflect the individual line profiles. Especially, the analysis becomes extremely difficult when the spexels contain two or more components. Anyway, the need of an automatic method for the analysis of data cube is justified by the fact that eye inspection is subjective and time-consuming.

If the line components were static in position and widths, the problem would come down to be a source separation problem from which a number of works have been proposed in the context of astrophysical source maps from 3D cubes in the last years [2]. However, theses techniques cannot be used in our application where the line components (*i.e.* the sources) may vary between two spatial locations. Therefore, Flitti *et al.* [5] have proposed a Bayesian segmentation carried out on reduced data. In this method, the spexels are decomposed into Gaussian functions yielding reduced data feeding a Markovian segmentation algorithm to cluster the pixels according to similar behaviors (figure 1 c).

We propose in this paper a two-step method to isolate coherent kinematic structures in the cube by first decomposing the spexels to extract the different line profiles and then to classify the estimated lines. The first step (section 2) decomposes each spexel in a sum of Gaussian components whose number, positions, amplitudes and widths are estimated. A Bayesian model is presented: it aims at using all the available information since pertinent data are too few. The major difference with Flitti's approach is that the decomposition is not set on a unique basis: line positions and widths may differ between spexels. The second step (section 3) classifies each estimated component line assuming that two components in two neighbouring spexels are considered in the same class if their parameters are close. This is a new supervised method allowing the astronomer to set a threshold on the amplitudes. The information about the spatial dependence between spexels is introduced in this step. Performing the decomposition and classification steps separately is simpler that performing them together. It also allows the astronomer to modify the classification without doing again the decomposition step which is time consuming. The method proposed in this paper is intended to help astronomers to handle complex data cubes and to be complementary to the standard method of analysis. It provides a set of spatial zones corresponding to the presence of a coherent kinematic structure in the cube, as well as spectral characteristics (section 4).

2   Spexel Decomposition

2.1   Spexel Model

Spexel decomposition is typically an object extraction problem consisting here in decomposing each spexel as a sum of spectral component lines. A spexel is a sum of spectral lines which are different in wavelength and intensity, but also in width. Besides, the usual model in radioastronomy assumes that the lines are Gaussian. Therefore, the lines are modeled by a parametric function $\boldsymbol{f}$ with unknown parameters (position $\boldsymbol{c}$, intensity $\boldsymbol{a}$ and width $\boldsymbol{w}$) which are estimated as well as the component number. We consider in the sequel that the cube

contains S spexels. Each spexel $s \in \{1, \ldots, S\}$ is a signal $\boldsymbol{y}_s$ modeled as a noisy sum of K_s components:

$$\boldsymbol{y}_s = \sum_{k=1}^{K_s} \boldsymbol{a}_{sk} \boldsymbol{f}(\boldsymbol{c}_{sk}, \boldsymbol{w}_{sk}) + \boldsymbol{e}_s = \boldsymbol{F}_s \boldsymbol{a}_s + \boldsymbol{e}_s, \tag{1}$$

where $\boldsymbol{f}$ is a vector function of length N, $\boldsymbol{e}_s$ is a $N \times 1$ vector modeling the noise, $\boldsymbol{F}_s$ is a $N \times K_s$ matrix and $\boldsymbol{a}_s$ is a $K_s \times 1$ vector:

$$\boldsymbol{F}_s = \begin{pmatrix} \boldsymbol{f}_1(\boldsymbol{c}_{s1}, \boldsymbol{w}_{s1}) \cdots \boldsymbol{f}_1(\boldsymbol{c}_{sK_s}, \boldsymbol{w}_{sK_s}) \\ \vdots \qquad\qquad \vdots \\ \boldsymbol{f}_N(\boldsymbol{c}_{s1}, \boldsymbol{w}_{s1}) \cdots \boldsymbol{f}_N(\boldsymbol{c}_{sK_s}, \boldsymbol{w}_{sK_s}) \end{pmatrix} \qquad \boldsymbol{a}_s = \begin{pmatrix} \boldsymbol{a}_{s1} \\ \vdots \\ \boldsymbol{a}_{sK_s} \end{pmatrix}.$$

The vector function $\boldsymbol{f}$ for component $k \in \{1, \ldots, K_s\}$ in pixel $s \in \{1, \ldots, S\}$ at frequency channel $n \in \{1, \ldots, N\}$ equals:

$$\boldsymbol{f}_n(\boldsymbol{c}_{sk}, \boldsymbol{w}_{sk}) = \exp\left(-\frac{(n - \boldsymbol{c}_{sk})^2}{2\boldsymbol{w}_{sk}^2}\right).$$

For simplicity, the expression of a Gaussian function was multiplied by $\sqrt{2\pi \boldsymbol{w}_{sk}^2}$ so that $\boldsymbol{a}_{sk}$ corresponds to the maximum of the line. In addition, we have $\forall s, k, \boldsymbol{a}_{sk} \geq 0$ because the lines are supposed to be non-negative. A perfect Gaussian shape is open to criticism because in reality the lines may be asymmetric, but modelling the asymmetry needs to consider one (or more) unknown and appears to be unnecessary complex.

Spexel decomposition is set in a Bayesian framework because it is clearly an ill-posed inverse problem [8]. Moreover, the posterior being a high dimensional complex density, usual optimisation techniques fail to provide a satisfactory solution. So, we propose to use Monte Carlo Markov chain (MCMC) methods [12] which are efficient techniques for drawing samples X from the posterior distribution π by generating a sequence of realizations $\{X^{(i)}\}$ through a Markov chain having π as its stationary distribution.

Besides, we are interesting in this step to decompose the whole cube, so the spexels are not decomposed independently each other. This allows to consider some global hyperparameters (such as a single noise variance allover the spexels).

2.2 Bayesian Model

The chosen priors are described hereafter for all s and k. A hierarchical model is used since it allows to set priors rather than a constant on hyperparameters. Some priors are conjugate so as to get usual conditional posteriors. We also try to work with usual priors for which simulation algorithms are available [12].

- the prior model is specified by supposing that K_s is drawn from a Poisson distribution with expected component number λ [7];
- the noise $\boldsymbol{e}_s$ is supposed to be white, zero-mean Gaussian, independent and identically distributed with variance $r_{\boldsymbol{e}}$;

- because we do not have any information about the component locations $\boldsymbol{c}_{sk}$, they are supposed uniformly distributed on $[1; N]$;
- component amplitudes $\boldsymbol{a}_{sk}$ are positive, so we consider that they are distributed according to a (conjugate) Gaussian distribution with variance $r_{\boldsymbol{a}}$ and truncated in zero to get positive amplitudes. We note: $\boldsymbol{a}_{sk} \sim \mathcal{N}^+(0, r_{\boldsymbol{a}})$ where $\mathcal{N}^+(\mu, \sigma^2)$ stands for a Gaussian distribution with positive support defined as (erf is the error function):

$$p(x|\mu, \sigma) = \sqrt{\frac{2}{\pi\sigma^2}} \left[1 + \mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}} \right) \right]^{-1} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \mathbb{1}_{[0;+\infty[}(x);$$

- we choose an inverse gamma prior $\mathcal{IG}(\alpha_{\boldsymbol{w}}, \beta_{\boldsymbol{w}})$ for component width $\boldsymbol{w}_{sk}$ because this is a positive-support distribution whose parameters can be set according to the approximate component width known a priori. This is supposed to equal 5 for the considered data but, because this value is very approximative, we also suppose a large variance (equals to 100), yielding $\alpha_{\boldsymbol{w}} \approx 2$ and $\beta_{\boldsymbol{w}} \approx 5$;
- the hyperparameter $r_{\boldsymbol{a}}$ is distributed according to an (conjugate) inverse gamma prior $\mathcal{IG}(\alpha_{\boldsymbol{a}}, \beta_{\boldsymbol{a}})$. We propose to set the mean to the approximate real line amplitude (say μ) which can be roughly estimated, and to assign a large value to the variance. This yields: $\alpha_{\boldsymbol{a}} = 2 + \varepsilon$ and $\beta_{\boldsymbol{a}} = \mu + \varepsilon$ with $\varepsilon \ll 1$;
- Again, we adopt an inverse gamma prior $\mathcal{IG}(\alpha_{\boldsymbol{e}}, \beta_{\boldsymbol{e}})$ for $r_{\boldsymbol{e}}$, whose parameters are both set close to zero ($\alpha_{\boldsymbol{e}} = \beta_{\boldsymbol{e}} = \zeta$, with $\zeta \ll 1$). The limit case corresponds to the common Jeffreys prior which is unfortunately improper.

The posterior has to be integrable to ensure that the MCMC algorithm is valid. This cannot been checked mathematically because of the posterior complexity but, since the priors are integrable, a sufficient condition is fulfilled. The conditional posterior distributions of each unknown is obtained thanks to the prior defined above:

$$\boldsymbol{c}_{sk}| \cdots \propto \exp\left(-\|\boldsymbol{y}_s - \boldsymbol{F}_s\boldsymbol{a}_s\|^2 / 2r_{\boldsymbol{e}} \right) \mathbb{1}_{[1,N]}(\boldsymbol{c}_{sk}),$$

$$\boldsymbol{a}_{sk}| \cdots \sim \mathcal{N}^+(\mu_{sk}, \rho_{sk}),$$

$$\boldsymbol{w}_{sk}| \cdots \propto \exp\left(-\frac{1}{2r_{\boldsymbol{e}}}\|\boldsymbol{y}_s - \boldsymbol{F}_s\boldsymbol{a}_s\|^2 - \frac{\beta_{\boldsymbol{w}}}{\boldsymbol{w}_{sk}} \right) \frac{1}{\boldsymbol{w}_{sk}^{\alpha_{\boldsymbol{w}}+1}} \mathbb{1}_{[0;+\infty[}(\boldsymbol{w}_{sk}),$$

$$r_{\boldsymbol{a}}| \cdots \sim \mathcal{IG}\left(\frac{L}{2} + \alpha_{\boldsymbol{a}}, \frac{1}{2}\sum_s\sum_{k=1}^{K_s}\boldsymbol{a}_{sk}^2 + \beta_{\boldsymbol{w}} \right),$$

$$r_{\boldsymbol{e}}| \cdots \sim \mathcal{IG}\left(\frac{NS}{2} + \alpha_{\boldsymbol{e}}, \frac{1}{2}\sum_s \|\boldsymbol{y}_s - \boldsymbol{F}_s\boldsymbol{a}_s\|^2 + \beta_{\boldsymbol{e}} \right)$$

where $x|\cdots$ means x conditionally to $\boldsymbol{y}$ and the other variables, N is the spectum length, S is the spexel number, $L = \sum_s K_s$ denotes the component number and

$$\mu_{sk} = \frac{\rho_{sk}}{r_e} \boldsymbol{z}_{sk}^T \boldsymbol{F}_{sk}, \qquad \rho_{sk} = \frac{r_a r_e}{r_e + r_a \boldsymbol{F}_{sk}^T \boldsymbol{F}_{sk}}, \qquad \boldsymbol{z}_{sk} = \boldsymbol{y}_s - \boldsymbol{F}_s \boldsymbol{a}_s + \boldsymbol{F}_{sk} \boldsymbol{a}_{sk}$$

where $\boldsymbol{F}_{sk}$ corresponds to the kth column of matrix $\boldsymbol{F}_s$.

The conditional posterior expressions for $\boldsymbol{c}_{sk}$, $\boldsymbol{w}_{sk}$ and the hyperparameters are straightforward, contrary to the conditional posterior for $\boldsymbol{a}_{sk}$ whose detail of computation can be found in [10, Appendix B].

2.3   MCMC Algorithm and Estimation

MCMC methods dealing with variable dimension models are known as trans-dimensional MCMC methods. Among them, the reversible jump MCMC algo-rithm [7] appears to be popular, fast and flexible [1]. At each iteration of this algorithm, a move which can either change the model dimension or generate a random variable is randomly performed. We propose these moves:

 - B_s "birth in s": a component is created in spexel s;
 - D_s "death in s": a component is deleted in spexel s;
 - U_s "update in s": variables $\boldsymbol{c}_s$, $\boldsymbol{a}_s$ and $\boldsymbol{w}_s$ are updated;
 - H "hyperparameter update": hyperparameters r_a and r_e are updated.

The probabilities b_s, d_s, u_s and h of moves B_s, D_s, U_s and H are chosen so that:

$$b_s = \frac{\gamma}{S+1} \min\left\{1, \frac{p(K_s+1)}{p(K_s)}\right\} \qquad d_s = \frac{\gamma}{S+1} \min\left\{1, \frac{p(K_s-1)}{p(K_s)}\right\}$$

$$u_s = \frac{1}{S+1} - b_s - d_s \qquad\qquad h = \frac{1}{S+1}$$

with γ such that $b_s + d_s \leq 0.9/(S+1)$ (we choose $\gamma = 0.45$) and $d_s = 0$ if $K_s = 0$.

We now discuss the simulation of the posteriors. Many methods available in literature are used for sampling positive normal [9] and inverse gamma distribu-tions [4,12]. Besides, $\boldsymbol{c}_{sk}$ and $\boldsymbol{w}_{sk}$ are sampled using a random-walk Metropolis-Hastings algorithm [12]. To improve the speed of the algorithm, they are sampled jointly avoiding to compute the likelihood twice. The proposal distribution is a (separable) truncated Gaussian centered on the current values:

$$\tilde{\boldsymbol{c}}_{sk} \sim \mathcal{N}\left(\boldsymbol{c}_{sk}^*, r_{\boldsymbol{c}}\right) \mathbb{1}_{[1,N]}(\tilde{\boldsymbol{c}}_{sk}), \qquad \tilde{\boldsymbol{w}}_{sk} \sim \mathcal{N}^+\left(\boldsymbol{w}_{sk}^*, r_{\boldsymbol{w}}\right)$$

where $\tilde{\cdot}$ stands for the proposal and $\cdot^*$ denotes the current value. The algorithm efficiency depends on the scaling parameters $r_{\boldsymbol{c}}$ and $r_{\boldsymbol{w}}$ chosen by the user (gen-erally with heuristics methods, see for example [6]).

The estimation is computed by picking in each Markov chain the sample which minimises the mean square error: it is a very simple estimation of the maximum a posteriori which does not need to save the chains. Indeed, the number of unknowns, and as a result, the number of Markov chains to save, is prohibitive.

3   Component Classification

3.1   New Proposed Approach

The decomposition method presented in section 2 provides for each spexel K_s components with parameter $\boldsymbol{x}_{sk} = \{\boldsymbol{c}_{sk}, \boldsymbol{a}_{sk}, \boldsymbol{w}_{sk}\}$. The goal of component classification is to assign to each component (s, k) a class $\boldsymbol{q}_{sk} \in \mathbb{N}^*$. One class corresponds to only one structure, so that components with the same class belong to the same structure. We also impose that, in each pixel, a class is present once at the most.

First of all, the components whose amplitude is lower than a predefined threshold τ are vanished in the following procedure (this condition helps the astronomer to analyse the gas location with respect to the intensity). To perform the classification, we assume that the component parameters exhibit weak variation between two neighbouring spexels, *i.e.* two components in two neighbouring spexels are considered in the same class if their parameters are close. The spatial dependency is introduced by defining a Gibbs field over the decomposed image [3]:

$$p(\boldsymbol{q}|\boldsymbol{x}) = \frac{1}{Z} \exp\left(-U(\boldsymbol{q}|\boldsymbol{x})\right) = \frac{1}{Z} \exp\left(-\sum_{c \in \mathcal{C}} U_c(\boldsymbol{q}|\boldsymbol{x})\right) \tag{2}$$

where Z is the partition function, $\mathcal{C}$ gathers the cliques of order 2 in a 4-connexity system and the potential function is defined as the total cost of the classification.

Let consider one component (s, k) located in spexel $s \in \{1, \ldots, S\}$ ($k \in \{1, \ldots, K_s\}$), and a neighbouring pixel $t \in \{1, \ldots, S\}$. Then, the component (s, k) may be classified with a component (t, l) ($l \in \{1, \ldots, K_t\}$) if their parameters are similar. In this case, we define the cost of component (s, k) equals to a distance $D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl})^2$ computed with the component parameters (we see further why we choose the square of the distance). On the contrary, if no component in spexel t is close enough to component (s, k), we choose to set the cost of the component to a threshold σ^2 which codes the weaker similarity allowed. Indeed, if the two components (s, k) and (t, l) are too different (that is $D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl})^2 > \sigma^2$), it would be less costly to let them in different classes. Finally, the total cost of the classification (*i.e.* the potential function) corresponds to the sum of the component costs.

Formally, these considerations read in the following manner. The potential function is defined as:

$$U_c(\boldsymbol{q}|\boldsymbol{x}) = \sum_{k=1}^{K_s} \varphi(\boldsymbol{x}_{sk}, \boldsymbol{q}_{sk}, \boldsymbol{x}_t, \boldsymbol{q}_t) \tag{3}$$

where s and t are the two spexels involved in the clique c, and $\varphi(\boldsymbol{x}_{sk}, \boldsymbol{q}_{sk}, \boldsymbol{x}_t, \boldsymbol{q}_t)$ represents the cost associated for the component (s, k) and defined as:

$$\varphi(\boldsymbol{x}_{sk}, \boldsymbol{q}_{sk}, \boldsymbol{x}_t, \boldsymbol{q}_t) = \begin{cases} D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl})^2 & \text{if } \exists l \text{ such that } \boldsymbol{q}_{sk} = \boldsymbol{q}_{tl}, \\ \sigma^2 & \text{otherwise.} \end{cases} \tag{4}$$

In some ways, $\varphi(\boldsymbol{x}_{sk}, \boldsymbol{q}_{sk}, \boldsymbol{x}_t, \boldsymbol{q}_t)$ can be seen as a truncated quadratic function which is known to be very appealing in the context of outliers detection [13].

We choose for the distance $D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl})$ a normalized Euclidean distance:

$$D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl}) = \sqrt{\left(\frac{\boldsymbol{c}_{sk} - \boldsymbol{c}_{tl}}{\delta_c}\right)^2 + \left(\frac{\boldsymbol{a}_{sk} - \boldsymbol{a}_{tl}}{\delta_a}\right)^2 + \left(\frac{\boldsymbol{w}_{sk} - \boldsymbol{w}_{tl}}{\delta_w}\right)^2}. \quad (5)$$

The distance is normalized because the three parameters have not the same unity. δ_c and δ_w are the normalizing factors in the frequency domain whereas δ_a is the one in the intensity domain. We consider that two components are similar if their positions or widths do not differ for more than 1.2 wavelength channel, or if the difference between the amplitudes do not exceed 40% of the maximal amplitude. So, we set $\delta_c = \delta_w = 1.2$, $\delta_a = \max(\boldsymbol{a}_{sk}, \boldsymbol{a}_{s'k'}) \times 40\%$ and $\sigma = 1$.

To resume, we look for:

$$\hat{\boldsymbol{q}} = \arg\max_{\boldsymbol{q}} p(\boldsymbol{q}|\boldsymbol{x}) \quad \Leftrightarrow \quad \hat{\boldsymbol{q}} = \arg\min_{\boldsymbol{q}} \sum_{c \in \mathcal{C}} \sum_{k=1}^{K_s} \varphi(\boldsymbol{x}_{sk}, \boldsymbol{q}_{sk}, \boldsymbol{x}_t, \boldsymbol{q}_t) \quad (6)$$

subject to the uniqueness of each class in each pixel.

3.2 Algorithm

We propose a greedy algorithm to perform the classification because it yields good results in an acceptable computation time (≈ 36 s on the cube considered in section 4 containing 9463 processed spexels). The algorithm is presented below. The main idea consists in tracking the components through the image by starting from an initial component and looking for the components with similar parameters spexel by spexel. These components are then classified in the same class, and the algorithm starts again until every estimated component is classified. We note z^* the increasing index coding the class, and the set $\mathcal{L}$ gathers the estimated components to classify.

1. set $z^* = 0$
2. while it exists some components that are not yet classified:
    3. $z^* = z^* + 1$
    4. choose randomly a component (s, k)
    5. set $\mathcal{L} = \{(s, k)\}$
    6. while $\mathcal{L}$ is not empty:
        7. set (s, k) as the first element of $\mathcal{L}$
        8. set $\boldsymbol{q}_{sk} = z^*$
        9. delete component (s, k) from $\mathcal{L}$
        10. among the 4 neighbouring pixels t of s, choose the components l that satisfy the following conditions:
            (C1) they are not yet classified
            (C2) they are similar to component (s, k) that is $D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl})^2 < \sigma^2$
            (C3) $D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tl}) = \arg\min_{m \in \{1, \dots, K_t\}} D(\boldsymbol{x}_{sk}, \boldsymbol{x}_{tm})$
            (C4) their amplitude is greater than τ
        11. Add (t, l) to $\mathcal{L}$

4   Application to a Modified Data Cube of NGC 4254

The data cube is a modified radio line observations made with the VLA of NGC 4254, a spiral galaxy located in the Virgo cluster [11]. It is a well-suited test case because it contains mainly only one single line (the HI 21 cm line). For simplicity, we keep in this paper pixel numbers for the spatial coordinates axis and channel numbers for the frequency axis (the data cube is a $512 \times 512 \times 42$ image, figures show only the relevant region). In order to investigate the ability of the proposed method to detect regions of double line profiles, we added an artificial line in a circular region north of the galaxy center. The intensity of the artificial line follows a Gaussian profile. Figure 1 (a and b) shows the maps of the first two moments integrated over the whole data cube and figure 1 c shows the estimation obtained with Flitti's method [5]. The map of the HI emission distribution (figure 1 a) shows an inclined gas disk with a prominent one-armed spiral to the west, and the additional line produces a local maximum. Moreover, the velocity field (figure 1 b) is that of a rotating disk with perturbations to the north-east and to the north. In addition, the artifical line produces a pronounced asymmetry. The double-line nature of this region cannot be recognized in the moment maps.
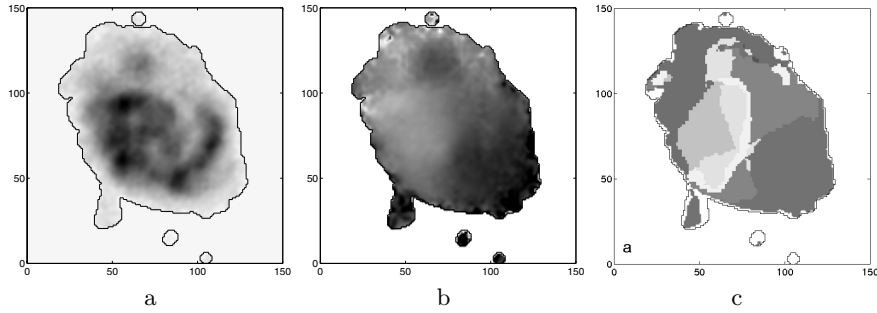

Fig. 1. Spiral galaxy NGC 4254 with a double line profile added: emission distribution (left) and velocity field (center); the figures are shown in inverse video (black corresponds to high values). Right: Flitti's estimation [5] (gray levels denote the different classes). The mask is displayed as a thin black line. The x-axis corresponds to right ascension, the y-axis to declination, the celestial north is at the top of the images and the celestial east at the left.

To reduce the computation time, a mask is determined to process only the spexels whose maximum intensity is greater than three times the standard deviation of the channel maps. A morphological dilation is then applied to connect close regions in the mask (a disk of diameter 7 pixels is chosen for structuring element).

The algorithm ran for 5000 iterations with an expected component number $\lambda = 1$ and a threshold $\tau = 0$. The variables are initialized by simulating them from the priors. The processing was carried out using Matlab on a double core (each 3.8 GHz) PC and takes 5h43. The estimation is very satisfactory because

the difference between the original and the estimated cubes is very small; this is confirmed by inspecting by eye some spexel decomposition. The estimated components are then classified into 9056 classes, but the majority are very small and, consequently, not significant. In fact, only three classes, gathering more than 650 components each, are relevant (see figure 2): the large central structure (a & d), the "comma" shape in the south-east (b & e) and the artificially added component (c & f) which appears clearly as a third relevant class. Thus, our approach operates successfully since it is able to distinguish clearly the three main structures in the galaxy.
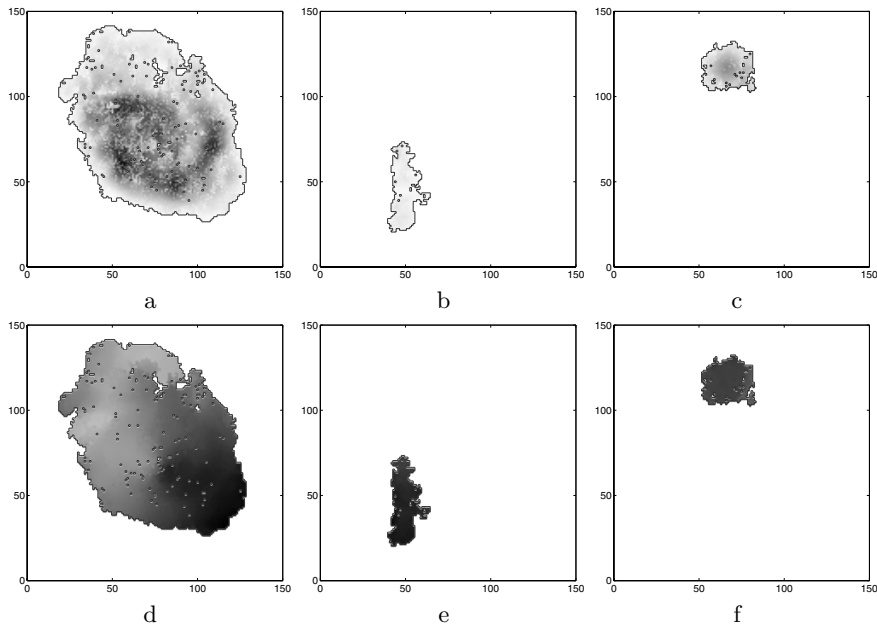

Fig. 2. Moment 0 (top) and 1 (bottom) of the three main estimated classes

The analysis of the two first moments of the three classes is also instructive. Indeed, the velocity field of the large central structure shows a rotating disk (figure 2 d). As well, the emission distribution of the artificial component shows that the intensity of the artificial line is maximum at the center and falls off radially, while the velocity field is quite constant (around 28.69, see figure 2, c and f). This is in agreement with the data since the artificial component is a Gaussian profile in intensity and has a center velocity at channel number 28.

Flitti *et al.* propose a method that clusters the pixels according to the six most representative components. Then, it is able to distinguish two structures that crosses while our method cannot because it exists at least one spexel where the components of each structure are too close. However, Flitti's method is unable to distinguish superimposed structures (since each pixel belongs to a single class) and a structure may be split into different kinematic zones if the spexels inside

are evoluting too much: these drawbacks are clearly shown in figure 1 c. Finally, our method is more flexible and can better fit complex line profiles.

5   Conclusion and Perspectives

We proposed in this paper a new method for the analysis of astronomical data cubes and their decomposition into structures. In a first step, each spexel is decomposed into a sum of Gaussians whose number and parameters are estimated via a Bayesian framework. Then, the estimated components are classified with respect to their shape similarity: two components located in two neighbouring spexels are set in the same class if their parameters are similar enough. The resulting classes correspond to the estimated structures.

However, no distinction between classes can be done if the structure is continuous because it exists at less one spexel where the components of each structure are too close. This is the major drawback of this approach, and future works will be dedicated to handle the case of crossing structures.

References

1. Cappé, O., Robert, C.P., Rydèn, T.: Reversible jump, birth-and-death and more general continuous time Markov chain Monte Carlo samplers. J. Roy. Stat. Soc. B 65, 679–700 (2003)
2. Cardoso, J.-F., Snoussi, H., Delabrouille, J., Patanchon, G.: Blind separation of noisy Gaussian stationary sources. Application to cosmic microwave background imaging. In: 11th EUSIPCO (2002)
3. Chellappa, R., Jain, A.: Markov random fields. Theory and application. Academic Press, London (1993)
4. Devroye, L.: Non-uniform random variate generation. Springer, Heidelberg (1986)
5. Flitti, F., Collet, C., Vollmer, B., Bonnarel, F.: Multiband segmentation of a spectroscopic line data cube: application to the HI data cube of the spiral galaxy NGC 4254. EURASIP J. Appl. Si. Pr. 15, 2546–2558 (2005)
6. Gelman, A., Roberts, G., Gilks, W.: Efficient Metropolis jumping rules. In: Bernardo, J., Berger, J., Dawid, A., Smith, A. (eds.) Bayesian Statistics 5, pp. 599–608. Oxford University Press, Oxford (1996)
7. Green, P.J.: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82, 711–732 (1995)
8. Idier, J. (ed.): Bayesian approach to inverse problems. ISTE Ltd. and John Wiley & Sons Inc., Chichester (2008)
9. Mazet, V., Brie, D., Idier, J.: Simulation of positive normal variables using several proposal distributions. In: 13th IEEE Workshop Statistical Signal Processing (2005)
10. Mazet, V.: Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies. PhD. thesis, Nancy University, France (2005)
11. Phookun, B., Vogel, S.N., Mundy, L.G.: NGC 4254: a spiral galaxy with an $m = 1$ mode and infalling gas. Astrophys. J. 418, 113–122 (1993)
12. Robert, C., Casella, G.: Monte Carlo statistical methods. Springer, Heidelberg (2002)
13. Rousseeuw, P., Leroy, A.: Robust Regression and Outlier Detection. Series in Applied Probability and Statistics. Wiley-Interscience, Hoboken (1987)