

NON-PARAMETRIC BAYESIAN FRAMEWORK FOR DETECTION OF OBJECT CONFIGURATIONS WITH LARGE INTENSITY DYNAMICS IN HIGHLY NOISY HYPERSPECTRAL DATA.

Céline Meillier, Florent Chatelain, Olivier Michel, Hacheme Ayasso

University of Grenoble, GIPSA-Lab, 11 rue des Mathématiques,
BP 46, 38402 St Martin d'Hères, France

ABSTRACT

In this study, a method that aims at detecting small and faint objects in noisy hyperspectral astrophysical images is presented. The particularity of the hyperspectral images that we are interested in is the high dynamics between object intensities. Detection of the smallest and faintest objects is challenging, because their signal-to-noise ratio is low, and if the brightest objects are not well reconstructed, their residuals can be more energetic than faint objects. This paper proposes a marked point process within a nonparametric Bayesian framework for the detection of galaxies in hyperspectral data. The efficiency of the method is demonstrated on synthetic images, and it provides good results for very faint objects in quasi-real astrophysical hyperspectral data.

Index Terms— Detection, nonparametric Bayesian models, marked point process, hyperspectral data

1. INTRODUCTION

Over the last few decades, different research fields have used hyperspectral data to exploit spectral information, such as in biology, astrophysics [1] and remote sensing [2]. Hyperspectral imaging devices produce massive data fields that need to be treated efficiently. This requires the development of algorithms that can take into account the mass of data, in order to overcome computational issues. Hyperspectral imaging devices produce ever more spectrally and spatially resolved images, and these lead to some processing difficulties. First, there are the dimensions of the hyperspectral data cube, whereby some observations can be made on a few thousands of spectral bands, and the spatial resolution or the width of the observation field requires an image a few hundreds of pixels wide. Then there are the high dynamics between object intensities; indeed, objects that are observed can be located at different astronomical distances from the imaging device, and their intensity decreases with the distance and the observation conditions. This leads to the third point: low signal-to-noise ratio (SNR) for some faint objects, plus dilution of information due to the optical device point spread function (PSF).

In the detection framework that we are interested in, the most difficult signals to be detected have low spatial extension and a very compact spectrum, and the maximum of their spectrum is close to the noise amplitude. All of these constraints require very good reconstruction of extended and bright objects. Indeed, if the estimates of the position, shape and intensity profile are not precise enough, then the residuals of these objects can be stronger than small and faint objects. Therefore, the detection of the object configuration and the estimation of the noise residuals and error measurements must be carried out jointly to have a chance of detecting small and faint objects.

Many different detection processings have been reported in the literature; see, for instance, [3] and [4] for reviews of supervised techniques in a hyperspectral context. The main drawback of these methods is the assumption that the noise properties are known *a priori*. Recent studies have proposed the detection of sparse signals in noisy hyperspectral images [5], [6]. Sparse representation is based on a dictionary of spectral components. In [5], a constrained likelihood ratio was proposed that exploited the spread of the signal in three dimensions using the PSF. A specific dictionary that was adapted to the signal to be detected was used to decrease the false-alarm probability, and this dictionary was optimized by techniques such as kernel singular value decomposition [7] or minimax [8]. All of these methods perform pixel-wise processing and can lead to tremendous computational issues.

To overcome these issues, the method proposed here favors an object-oriented approach. Hence we propose a marked point process that has often been used to model object configuration [9], [10]. The estimation is reformulated in a nonparametric Bayesian framework. A previous study was presented in [11] for the detection of galaxies in hyperspectral data cubes provided by the Multi-Unit Spectroscopic Explorer (MUSE) panoramic integral-field spectrograph. In this paper, we propose an improvement of the Bayesian model to account for the wide dynamics between the intensities of the different galaxies to be detected. All of the regularization of the intensity intrinsic to the Bayesian observation model described in [11] has been replaced by a new approach to highlight the presence of small and faint objects. The advantages of marked point processes and the nonparametric Bayesian framework are preserved and frequentist statistical tests as adapted to small and faint objects are used to propose object positions. This allows the performance of the object configuration sampling to be enhanced.

The proposed detection algorithm is applied to MUSE hyperspectral data that contain very faint light sources and bright objects, as for nearby galaxies and stars in the field of observation. The MUSE project aims to provide observations of the sky to astrophysicists, and especially for distant galaxies [12], [13]. MUSE is planned to be installed on the Very Large Telescope by the end of 2013. The dimensions of the hyperspectral data cube are finally 300×300 pixels \times 3600 wavelengths. One of the major scientific purposes of MUSE is the detection of very distant galaxies where the spectrum is composed of one faint and narrow line, called a Lyman-alpha line. The detection challenge arises as the position of the galaxies in the image and the position of the Lyman-alpha line in the spectrum are unknown, as is the redshift related to the object velocity or distance.

This report is organized as follows. The nonparametric Bayesian model is introduced in Section 2, and the detection method is presented in Section 3. Application to astrophysical data is shown in Section 4, and some conclusions are drawn in the last section.

2. THE NONPARAMETRIC BAYESIAN MODEL

To model the object configuration \mathbf{u} , a marked point process is used. The realizations of this stochastic process belong to the infinite dimension space of all of the configuration with a finite number of objects. An object is modeled as a point, i.e., a position in the image, and some marks are added. In the galaxy detection application, objects are characterized by an elliptical geometry that is governed by some marks: length of semi-axes, orientation, and some other additional marks related to the intensity profile. As we do not have physical modeling of galaxies, a generic Sérsic profile [14] is used, and the Sérsic index, which parameterizes this profile, belongs to the marks of the objects. This process leads to a sparse representation of the object configuration to be detected. The process is defined by its density with respect to the probability measure of a reference Poisson point process. This leads to a general and robust framework that avoids penalizing the detection performance by some *a-priori* information that can be misestimated.

2.1. Observation model

For the sake of simplicity, we assume the simplest case where the object detection is on a single $P \times Q$ image (i.e., at a single wavelength $\lambda \in [1, \dots, \Lambda]$). However, the same model can be extended to all of the images when the data is a stack of images.

Let \mathbf{y} be the vectorized image, \mathbf{y} is a $M \times 1$ vector where $M = P \times Q$ is the number of pixels. The objects are modeled by the configuration \mathbf{u} of marked points, which is assumed to be known in this section. Formally, the detection of the different sources from the observation \mathbf{y} is addressed by:

$$\mathbf{y} = \mathbf{X}\mathbf{w} + \mathbf{1}m + \epsilon_{Bg}, \quad (1)$$

where m is the mean of the background intensity, $\mathbf{1}$ is a unit $M \times 1$ vector, and ϵ_{Bg} is a spatially centered white Gaussian noise $M \times 1$ vector, such that:

$$\epsilon_{Bg} \sim \mathcal{N}(0, \sigma^2 I_M) \quad (2)$$

where I_M is the $M \times M$ identity matrix. The object configuration \mathbf{u} is represented by the matrix $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_{n(\mathbf{u})}]$ which is a $M \times n(\mathbf{u})$ matrix in which each column \mathbf{x}_j is the vectorized image of a source u_j that is convoluted with the impulse response of the optical device, called a PSF. The number of detected objects is $n(\mathbf{u})$ and $\mathbf{w} = [w_1, \dots, w_{n(\mathbf{u})}]^T \in \mathbb{R}^{n(\mathbf{u})}$ is the weights vector; i.e., the intensities of each source of \mathbf{X} .

Note that although a single image (at λ) is considered, the pixel value is obtained from the three-dimensional imaging system. Thus, the convolution between objects and the PSF can be written as follows:

$$x_i = u_i * H \quad (3)$$

where $*$ is a three-dimensional convolution operator, and H is the three-dimensional PSF of the instrument described in [15]. Here, r and z are the spatial variables, while λ and μ are spectral variables, and (r, λ) is the position of the pixel of interest. For a hyperspectral imaging device, this PSF can be considered as separable into a spatial component, the field spread function (FSF), and a spectral component, the line spread function (LSF). Thus the PSF centered on pixel (r, λ) reads as:

$$H_{r,\lambda}(z, \mu) = L_{r,\lambda}(z, \mu) F_{r,\lambda}(z, \mu). \quad (4)$$

Moreover, the FSF $F_{r,\lambda}$ is considered as shift invariant in the observation field: $F_{r,\lambda}(z, \mu) = F_\lambda(z - r)$, and the LSF is spatially constant and shows slow spectral variation, and thus it can be approximated by: $L_{r,\lambda}(z, \mu) = L_\lambda(\mu)$. This finally yields the following expression of the PSF:

$$H_{r,\lambda}(z, \mu) = L_\lambda(\mu) F_\lambda(z - r) \quad (5)$$

In the MUSE application, the PSF is modeled according to Equation (5), where the FSF and LSF factors are well known by instrument calibration. Finally, considering a single image (at λ) makes sense, as in the overall model for the cube, it will be assumed that the random noise fluctuations are independent *versus* λ . This allows the same model to be considered from Equation (1) for all λ , where \mathbf{u} is uniquely defined with respect to m and σ^2 , which depend on λ .

2.2. Likelihood function

The observation model defined in Equation (1) leads to the following likelihood:

$$f(\mathbf{y}|\mathbf{u}, m, \sigma^2) = \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{M}{2}} \exp \left(-\frac{\mathbf{Z}^T \mathbf{Z}}{2\sigma^2} \right) \quad (6)$$

where $\mathbf{Z} = \mathbf{y} - \mathbf{X}\mathbf{w} - \mathbf{1}m$. We should recall that both m and σ^2 depend on λ .

2.3. Background parameter priors

Using a fully Bayesian model for the data aims for the building of an automatic and robust method. The Bayesian approach provides a complete framework for combining data information, using the likelihood function and external knowledge of background parameters (m, σ^2) sampled from a prior (m, σ^2) . A noninformative Jeffrey's prior is chosen for (m, σ^2) :

$$p(m, \sigma^2) = \frac{1}{\sigma^2} \mathbb{1}_{]0, +\infty[}(\sigma^2) \quad (7)$$

Note that there is no regularization on m . In this case, its maximum *a-posteriori* estimate reduces to the maximum likelihood one.

2.4. Object intensity prior

In a previous study [11], a Gaussian prior was investigated, and a *g*-prior was introduced:

$$p(\mathbf{w}|\mathbf{u}, \sigma^2, g^2) \sim \mathcal{N} \left(0, g^2 \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \right) \quad (8)$$

where the hyperparameter g^2 represents the *a-priori* SNR of the observed scene:

$$\begin{aligned} \frac{E[(\mathbf{X}\mathbf{w})^T(\mathbf{X}\mathbf{w})]}{E[\epsilon_{Bg}^T \epsilon_{Bg}]} &= \frac{\text{Tr}\{E[\mathbf{w}\mathbf{w}^T] \mathbf{X}^T \mathbf{X}\}}{\sigma^2 \text{Tr}\{I\}} \\ &= \frac{\text{Tr}\{g^2 \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X}\}}{\sigma^2 \text{Tr}\{I\}} = g^2 \end{aligned} \quad (9)$$

This prior is adapted to detect objects that have similar dynamics, because g^2 is the same SNR for all of the objects in a single image. This introduces too restrictive a constraint to detect objects with large dynamics between the brightest objects and the faintest, as occurs in the astrophysical application proposed in [11] and in Section 4 below.

In this paper, we consider the other solution that consists of using a noninformative prior on a \mathbf{w} vector. Jeffrey's prior is chosen as:

$$p(\mathbf{w}|\mathbf{u}) \propto \mathbb{1}_{\mathbb{R}}(\mathbf{w}). \quad (10)$$

2.5. Configuration prior

To avoid multiple detection, a hard core penalization is introduced to prevent objects overlapping. Let $r(u_i, u_j)$ be the overlapping ratio between the energy distribution of two objects u_i and u_j . This ratio $r(u_i, u_j) \in [0, 1]$ leads to the consideration of a hard core penalization term that is characterized by the following density with respect to the reference measure $\mu(\cdot)$:

$$h_1(\mathbf{u}) = \begin{cases} 0 & \text{if it exists } i \neq j \text{ such that } r(u_i, u_j) > t, \\ 1 & \text{otherwise,} \end{cases} \quad \text{for } t \in [0, 1]$$

Let $\mathbf{Y}(r, :)$ be the $1 \times \Lambda$ vector that corresponds to the spectrum at pixel r . To control the level of false alarms in the object detection procedure, the following hypotheses are assumed for each pixel spectrum $\mathbf{Y}(r, :) = [\mathbf{y}_1(r), \dots, \mathbf{y}_\Lambda(r)]$:

$$\begin{cases} \mathcal{H}_0 \text{ (absence of object)} & : \mathbf{Y}(r, :) = \boldsymbol{\epsilon}(:) \\ \mathcal{H}_1 \text{ (presence of object)} & : \mathbf{Y}(r, :) = \boldsymbol{\alpha}(r, :) + \boldsymbol{\epsilon}(:) \end{cases} \quad (11)$$

where $\boldsymbol{\epsilon}$ is a $1 \times \Lambda$ noise spectrum, $\boldsymbol{\epsilon}(\lambda) \sim \mathcal{N}(m_\lambda, \sigma_\lambda)$ for each band λ with m_λ and σ_λ as the noise parameters described in Equations (1) and (2). Let $\boldsymbol{\alpha}(r, :)$ be the spectrum observed at r . As the detector aims to find faint objects of low spatial extension and very compact spectrum (a single spectral line), $\boldsymbol{\alpha}(r, :)$ will be close to the PSF. Thus, the hyperspectral data cube is processed with a filter matched to the PSF cube, and statistics based on the maximum of the filtered spectrum are used to highlight the Lyman-alpha emitter characteristics. This yields the following binary test:

$$\max_{\lambda} (\mathbf{y}_\lambda^f(r)) \stackrel{\mathcal{H}_0}{\underset{\mathcal{H}_1}{\leq}} \eta(p_{FA}), \quad (12)$$

where $\mathbf{y}_\lambda^f(r) = \sum_z \sum_\mu H_{r,\lambda}(z, \mu) \mathbf{y}_\mu(z)$ is the matched filter statistics, and $\eta(p_{FA})$ is the threshold corresponding to a given false alarm p_{FA} . Note that, because of the spectral correlation of $\mathbf{y}^f(r, :)$, it is not possible to obtain a tractable expression of $\eta(p_{FA})$. However, this can be easily approximated by Monte-Carlo simulations, with the parameters m_λ and σ_λ replaced by some pre-processing estimates that are delivered with the MUSE data cube. This approach is similar to a study presented by [5]. Another hard threshold penalization term $h_2(\mathbf{u})$ is introduced to take into account the max-test result:

$$h_2(\mathbf{u}) = \prod_{i=1}^{n(\mathbf{u})} h_2(u_i) \quad (13)$$

with

$$h_2(u_i) = \begin{cases} 1 & \text{if } \max_{\lambda} (\mathbf{y}_\lambda^f(r_{u_i}, \lambda)) > \eta(p_{FA}), \\ 0 & \text{otherwise,} \end{cases}$$

where r_{u_i} is the center of the object u_i . Finally, $h(\mathbf{u}) = h_1(\mathbf{u})h_2(\mathbf{u})$.

2.6. Prior on the reference measure

The *a-priori* information regarding the number and location of the objects being detected can be introduced into the intensity measure of the reference nonhomogeneous Poisson point process.

As the matched filter statistics indicates the most likely locations, some classes \mathcal{C}_i are created according to these statistics values. Typically, there are three classes: \mathcal{C}_3 contains each pixel under the test statistics threshold, the other pixels are in \mathcal{C}_2 , except for the local maxima, which are in \mathcal{C}_1 . A probability is associated with each class to favor the proposition in the most probable class, such that $p_{\mathcal{C}_1} = 0.8$, $p_{\mathcal{C}_2} = 0.2$, and $p_{\mathcal{C}_3} = 0$; as these last locations are not detectable, this is equivalent to the constraint introduced by $h_2(\mathbf{u})$ (Equation 13). This yields the following normalized intensity function $\lambda(x) = \frac{p_{\mathcal{C}_i}}{|\mathcal{C}_i|}$, where i is the class of the nearest pixel for the location x , and $|\mathcal{C}_i|$ is the number of pixels in \mathcal{C}_i .

Finally, the density of the process with respect to this normalized Poisson point process measure [16, p. 15] is:

$$p(\mathbf{u}|\beta) = \beta^{n(\mathbf{u})} e^{-(\beta-1)}, \quad \propto \beta^{n(\mathbf{u})} e^{-\beta}. \quad (14)$$

where $n(\mathbf{u})$ is the number of objects in the configuration \mathbf{u} , and β is the mean number of objects. As there is no available information about parameter β , a vague Gamma prior distribution is chosen:

$$p(\beta) = \frac{\beta^{a-1} e^{-\beta/b}}{\Gamma(a)b^a}, \quad \forall \beta \in \mathbb{R}^+, \quad (15)$$

where $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$ is the classical Gamma function, and the hyperparameters are fixed to $a = 1$ and $b = 10^3$ to obtain a sufficiently vague prior (with large variance). Using the Bayes formula, the joint density becomes:

$$f(\mathbf{u}, \beta) = f(\mathbf{u}|\beta)p(\beta) \propto \beta^{n(\mathbf{u})} \exp\left(-\left(\frac{1+b}{b}\right)\beta\right). \quad (16)$$

This density can be marginalized by integrating out β :

$$f(\mathbf{u}) \propto \Gamma(n(\mathbf{u}) + 1) q^{n(\mathbf{u})+1}, \quad (17)$$

where $q = b/(b+1)$.

This term defines the density of our reference process with respect to the normalized Poisson process measure.

2.7. Posterior distribution

The resulting posterior density can be calculated, as:

$$\begin{aligned} p(\mathbf{u}, \mathbf{w}, m, \sigma^2 | \mathbf{y}) &\propto p(\mathbf{y} | \mathbf{w}, m, \sigma^2, \mathbf{u}) p(\mathbf{w} | \mathbf{u}) p(m, \sigma^2) p(\mathbf{u}) \\ &\propto \exp\left(-\frac{(\mathbf{y} - \mathbf{X}\mathbf{w} - \mathbf{1}m)^T (\mathbf{y} - \mathbf{X}\mathbf{w} - \mathbf{1}m)}{2\sigma^2}\right) \\ &\quad \times \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{M}{2}} \times \frac{1}{\sigma^2} \mathbb{1}_{]0, +\infty[}(\sigma^2) \mathbb{1}_{\mathbb{R}}(\mathbf{w}) \mathbb{1}_{\mathbb{R}}(m) \\ &\quad \times \Gamma(n(\mathbf{u}) + 1) \times q^{n(\mathbf{u})+1} h(\mathbf{u}) \end{aligned} \quad (18)$$

This posterior distribution is marginalized by integrating out \mathbf{w} , to reduce the number of parameters to be estimated. After a straightforward calculation, the joint posterior distribution for both the configuration \mathbf{u} and parameters (m, σ^2) becomes:

$$\begin{aligned} p(\mathbf{u}, m, \sigma^2 | \mathbf{y}) &\propto p(m, \sigma^2 | \mathbf{u}, \mathbf{y}) p(\mathbf{u}) \\ &\propto e^{-\frac{(\mathbf{y} - \mathbf{1}m)^T (\mathbf{I} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T)(\mathbf{y} - \mathbf{1}m)}{2\sigma^2}} \\ &\quad \times \left(\frac{1}{\sigma^2}\right)^{\frac{M}{2}} \times \frac{1}{\sigma^2} \mathbb{1}_{]0, +\infty[}(\sigma^2) \\ &\quad \times \Gamma(n(\mathbf{u}) + 1) \times q^{n(\mathbf{u})+1} h(\mathbf{u}) \end{aligned} \quad (19)$$

Note that even if improper priors have been chosen (Equation 7), the posterior densities of the background parameters are well-defined.

3. DETECTION METHOD

To enhance the performances of the configuration sampler, the proposition map is the same as the intensity of the reference measure defined in 2.6. This allows the proposition for the most likely locations according to the max-test statistics to be favored.

The reversible jump Monte Carlo Markov chain (RJMCMC) algorithm [17] is used to sample the object configuration and background parameters. The RJMCMC algorithm is chosen first because the dimension of the problem varies together with the number of detected objects. A second reason lies in the simplicity to mix Gibbs steps [18] for background parameters, sampling and Metropolis-Hastings-Green moves to sample the configuration. In our case, both the configuration and the noise parameters must be estimated.

Given the marginalized posterior distribution (Equation 19) the conditional posterior distribution can be deduced for each background parameter given the other parameters. As Jeffrey's priors were chosen, these posterior distributions belong to classical distribution families, and they can be sampled easily (see [11] for more details). Thus, Gibbs moves are used to generate Markov chains $\{m^{(t)}\}_t$ and $\{\sigma^{2(t)}\}_t$.

For sampling the object configuration, different moves are used: birth, death, modification of position, orientation, and dimension or

spatial profile. Birth moves consist of the adding of an object to the current configuration \mathbf{u} to build the new one (\mathbf{v}), while death moves randomly remove a single object of the configuration. Simple moves such as translation, rotation, enlarge or shrink are applied to one object selected with a uniform probability. All of these moves from a configuration \mathbf{u} to the configuration \mathbf{v} are randomly accepted with a probability $\min(1, r(\mathbf{u}, \mathbf{v}))$, where $r(\mathbf{u}, \mathbf{v})$ denotes the Metropolis-Hastings-Green ratio. This ratio depends on the transition kernels associated with the different moves, as expressed in [11]. In the birth case, the following Metropolis-Hastings-Green ratio is obtained:

$$r(\mathbf{u}, \mathbf{v}) = \frac{p(\mathbf{v}, m, \sigma^2 | \mathbf{y}) p_D(\mathbf{v})}{p(\mathbf{u}, m, \sigma^2 | \mathbf{y}) p_B(\mathbf{u})}, \quad (20)$$

For a given configuration \mathbf{u} , $p_B(\mathbf{u})$ denotes the probability to select the birth move, while $p_D(\mathbf{u}) = 1 - p_B(\mathbf{u})$ is the probability to choose the reversible move: the death move. In the death case, the ratio becomes:

$$r(\mathbf{u}, \mathbf{v}) = \frac{p(\mathbf{v}, m, \sigma^2 | \mathbf{y}) p_B(\mathbf{v})}{p(\mathbf{u}, m, \sigma^2 | \mathbf{y}) p_D(\mathbf{u}) p_S(u_i | \mathbf{u})}, \quad (21)$$

where $p_S(u_i | \mathbf{u})$ is the uniform probability to select the object u_i in the configuration \mathbf{u} . Simple moves on one object of the current configuration can be viewed as the composition of a death move and a birth move. Finally, the configuration that maximizes the posterior density is selected as an approximation of the maximum *a-posteriori* (MAP) estimator.

4. APPLICATION TO HYPERSPECTRAL DATA

The main challenge in hyperspectral MUSE data is the detection of the faint galaxies that are buried in both the Gaussian noise and the optical response of large and bright galaxies. Figure 1 shows typical galaxy spectra observed using the MUSE instrument.

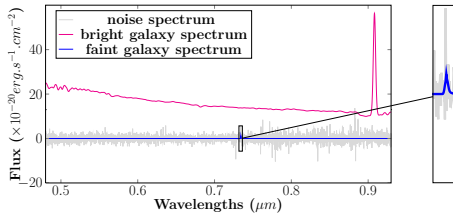


Fig. 1: Typical spectra of bright (magenta) and faint (blue) galaxies. The noise spectrum is in gray.

In this example, the faintest galaxy spectrum is much lower than the noise level and it is at least 100-fold fainter than the brightest galaxy spectrum. Lyman-alpha emitters are expected to show such sparse and compact spectra. In the faint galaxy spectrum shown in Figure 1, the position of the spectral line cannot be predicted because of the redshift effect. As the MUSE instrument will be operational in early 2014, we work here with a simulated quasi-real data cube of 100×100 pixels by 3600 wavelengths that are provided by the MUSE consortium. This contains eighteen typical objects that might be observed by the MUSE instrument, and that are generated according to some physical models or extracted from a catalog of galaxies. Some of the objects show very low SNR, and the objective is to test the performances of different algorithms on this cube. A sub-cube of 28×26 pixels by 3600 wavelengths of these data has already been used in [5], and our method here can be compared with the constrained likelihood ratio proposed in [5] on this sub-cube.

Figure 2 presents the results obtained for 10^4 RJMCMC iterations. A proposition map was calculated for a false alarm probability of 2×10^{-2} . The performances obtained for object detection from

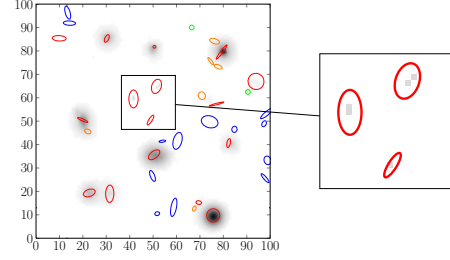


Fig. 2: The detection result over the whole cube, where the background image is the max-test result, and white pixels are under the threshold. Ellipses correspond to the support of the Sérsic profile at a decreasing rate of 63%, red ellipses represent good detection, blue ellipses are false alarms, orange ellipses represent multiple detections of extended objects, and green circles show where the detection was missed. Inset: Zoom on a sub-cube that contains three very faint Lyman-alpha objects.

these data are similar to those obtained from synthetic images (for gravity reasons). The mean absolute error for the position is 0.5 pixels. Setting $p_{FA} = 2 \times 10^{-2}$, 16 out of 18 objects are well detected. The undetected objects have very low SNR, and they are likely to be below the detection limit for this kind of data. Setting the p_{FA} to the same value as in [5], this provides results that are comparable to those obtained by the pixel-wise constrained likelihood ratios based method presented in [5] on the sub-cube displayed in the inset of Figure 2. Note, however, that with our model, the position is considered as continuous in the image. This is expected to improve the detection performances, as it allows a better registration of the PSF with the observed pixelized image. Moreover, the object marks corresponding to the Sérsic profile parameters, which are related to the spatial profile for each galaxy, are also sampled and estimated to account more accurately for their spatial response. Finally, the nonparametric Bayesian framework provides the advantage that the multiple detection problem can be addressed directly, and the background parameters m and σ^2 can be re-estimated taking into account the objects already detected. These new values can also be used to compute the configuration prior h_2 (Equation 13), by replacing the first estimate of m and σ^2 provided with the cube. This would allow the map of proposed objects to be updated over the detection process, and the object detection performance to be enhanced. This refinement is the purpose of our ongoing study.

5. CONCLUSION

In this communication, we have explored the ability of a point process model formulated in a nonparametric Bayesian framework to detect astrophysical objects in a large hyperspectral data cube. Although there have been similar studies, we have developed an original approach where strong regularization is replaced by a statistical test that is better suited to detect sparse faint objects. The proposed approach shows improved performance where objects might have extremely different SNR. Furthermore, this approach provides a means to detect super-resolved positions, and it can easily be extended to large datasets at reasonable computational cost.

6. ACKNOWLEDGMENT

The authors thank Roland Bacon and Johan Richard for providing the hyperspectral MUSE data. They also thank David Mary, Andre Ferrari and Eric Slezak for fruitful discussions.

7. REFERENCES

- [1] B. R. Oppenheimer et al., "Project 1640: the world's first exao coronagraphic hyperspectral imager for comparative planetary science," in *Adaptive Optics systems III, Proceeding of SPIE*, 2012, vol. 8447.
- [2] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectroscopy for earth remote sensing," *Science*, vol. 228, pp. 1147–1153, 1985.
- [3] D. Manolakis and G. A. Shaw, "Detection algorithms for hyperspectral imaging application," *IEEE Signal Processing Magazine*, January 2002.
- [4] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14, 2003.
- [5] Silvia Paris, R.F.R. Suleiman, David Mary, and Andre Ferrari, "Constrained likelihood ratios for detecting sparse signals in highly noisy 3d data," in *International Conference on Acoustics, Speech and Signal Processing*, 2013.
- [6] Sebastien Bourguignon and Slezak Eric Mary David, "Processing muse hyperspectral data: Denoising, deconvolution and detection of astrophysical sources," *Statistical Methodology*, 2011.
- [7] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [8] R.F.R. Suleiman, D. Mary, and A. Ferrari, "Minimax sparse detection based on one-class classifiers," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 5553–5557.
- [9] C. Lacoste, X. Descombes, and J. Zerubia, "Point processes for unsupervised line network extraction in remote sensing," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, 2005.
- [10] Mathias Ortner, Xavier Descombes, and Josiane Zerubia, "Building outline extraction from digital elevation models using marked point processes," *International Journal of Computer Vision*, vol. 72, no. 2, pp. 107–132, 2007.
- [11] F. Chatelain, A. Costard, and O. Michel, "A bayesian marked point process for object detection. Application to MUSE hyperspectral data," in *International Conference on Acoustics, Speech and Signal Processing*, 2011.
- [12] "Muse project website," <http://muse.univ-lyon1.fr>.
- [13] R. Bacon et al., "Probing unexplored territories with MUSE : a second generation instrument for the VLT," in *SPIE 6265*, 2006.
- [14] J. L. Sersic, "Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy," *Bulletin of the Astronomical Association of Argentina*, pp. 41–43, 1963.
- [15] Denis Serre, Emma Villeneuve, Hervé Carfantan, Laurent Jolissaint, Vincent Mazet, Sébastien Bourguignon, and Aurélien Jarno, "Modeling the spatial psf at the vlt focal plane for muse wfm data analysis purpose," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2010, vol. 7736.
- [16] Xavier Descombes, Ed., *Stochastic geometry for image analysis*, Wiley-ISTE, 2011.
- [17] P.J. Green, "Reversible jump markov chain monte carlo computation and bayesian model determination," *Biometrika*, vol. 52, pp. 711–732, 1995.
- [18] S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution and bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 721–741, 1984.