Hierarchical Multispectral Galaxy Decomposition using a MCMC Algorithm with Multiple Temperature Simulated Annealing

Benjamin Perret[a,*], Vincent Mazet[a], Christophe Collet[a], Éric Slezak[b]

[a] *LSIIT (UMR 7005, University of Strasbourg-CNRS),address: Pôle API, Bd Sébastien Brant, BP 10413, 67412 Illkirch Cedex France*
[b] *University of Nice-Sophia Antipolis, Cassiopée Laboratory, UMR 6202, address: CNRS-Observatoire de la Côte d'Azur, Boulevard de l'Observatoire, BP 4229, 06304 Nice, Cedex 4, France*

Abstract

We present a new method for the parametric decomposition of barred spiral galaxies in multi-spectral observations. The observation is modelled with a realistic image formation model and the galaxy is composed of physically significant parametric structures. The model also includes a parametric filtering to remove non desirable aspects of the observation. Both the model and the filter parameters are estimated by a robust Monte Carlo Markov chain (MCMC) algorithm. The algorithm is based on a Gibbs sampler combined with a novel strategy of simulated annealing in which several temperatures allow to manage efficiently the simulation effort. Besides, the overall decomposition is performed following an original framework: a hierarchy of models from a coarse model to the finest one is defined. At each step of the hierarchy the estimate of a coarse model is used to initialize the estimation of the finer model. This leads to an unsupervised decomposition scheme with a reduced computation time. We have validated the method on simulated and real 5-band images: the results showed the accuracy and the robustness of the proposed approach.

Keywords: Modelling and recovery of physical attributes, Monte Carlo Markov chain algorithms, Simulated annealing, Hierarchical decomposition, Inverse problems, Astronomy

1. Introduction

Thanks to the increasing computational power available in modern computers, the traditional low-level models (or pixel models) used in image processing are gradually replaced by high-level semantic descriptions using complex models for the objects present in the image. These high-level image models, where each object describes a large set of pixels with only a few significant parameters of high-level semantic, have become very popular, especially with the emergence of marked point process methods [1]. These approaches have the advantage to fill the semantic gap which arose with the increasing spatial resolution of images, when pixels taken independently are no more significant. Applications in various domains like building reconstruction [2] in dense remote sensing images, vascular tree reconstruction [3], flamingo counting [4] or tree crown extraction [5] have been developed.

Because these approaches lead to complex ill-posed inverse problems of high dimension, the problem is usually regularized within a Bayesian framework [6, 7] or, more generally, with a regularized energy function [5]. In this difficult context, MCMC algorithms have proved their ability to provide robust solutions in a fully automated way.

---

*Corresponding author. Tel: +333 68 85 44 84. Fax: +333 68 85 44 55
Email addresses: bperret@unistra.fr (Benjamin Perret), vincent.mazet@unistra.fr (Vincent Mazet), c.collet@unistra.fr (Christophe Collet), eric.slezak@oca.eu (Éric Slezak)

In this paper, we aim at developing a high-level semantic model for the analysis of multispectral images of galaxies in astronomy as well as an efficient optimization algorithm based on MCMC methods to compute an estimation of the model. The MCMC algorithm is based on the Hastings Metropolis within Gibbs technique taking advantage of recent development in adaptive scale and direction methods [8–10]. In this way, we obtain an automatic method to characterize these objects by extracting quantitative morphological, brightness and spectral descriptors. Today, this application is a major issue since astronomers are searching for a new comprehensive and objective spectro-morphological classification scheme of local universe galaxies.

Galaxy decomposition and quantitative morphology measurement are very topical problems in astronomy (see for instance [11–14] and references therein) with many theoretical outcomes. First, it provides clues to the origin of galaxies and their formation process in relation with the local environment. In this aspect, understanding the morphology of galaxies at low and moderate redshifts (local universe) is crucial for any meaningful comparison with galaxy images at higher redshift (very distant objects at earlier epochs) and with simulation results. Second, it helps in understanding the physical differences between galaxies, since it appears that morphological types do correlate with colours, atomic hydrogen mass and other quantities. Moreover, the huge data sets related to large present and forthcoming surveys prevent a visual inspection by dedicated experts and require automatic and objective methods. Several algorithms and models have already been developed to perform galaxy decompositions [15–17]. Nevertheless, these methods suffer either from the use of too naive or simple models or from their supervised approach of the problem requiring parameter tuning. For example, [15] is an automatic algorithm but the model used is too simple for our purpose. On the other hand, other methods use more sophisticated models but they require a lot of human time for their supervision [12]. Besides nowadays, large multispectral imaging surveys produce huge amount of multiwavelength data for local and more distant (and less resolved) galaxies and a major issue in astronomy is to design methods able to process automatically and efficiently such data [18].

Following the Hubble's visual classification scheme [19] and its modifications, galaxies can be divided into a sequence of four major classes: from elliptical galaxies, through lenticular galaxies, to spiral galaxies and a parallel branch of spirals with a barred component extending out from the nucleus; each galaxy class has subclasses for which boundaries may be difficult to define (Fig. 1). Galaxies from these four classes can be decomposed into several morphologically significant features: the bulge, the stellar disc, the spiral arms and the stellar bar. So, a barred spiral galaxy in the Hubble's classification has the richest decomposition and all galaxies from the other classes can be seen as barred spiral galaxies with missing structures. Thus, we present hereafter a rich barred spiral galaxy model which can model all galaxy types. It can naturally handle multispectral information and we propose a new model for the spiral structure simulation that was still missing in the literature.
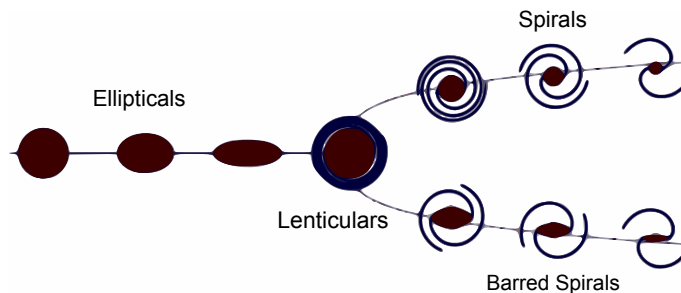

Figure 1: The Hubble tuning fork separates galaxies into four classes: elliptic, lenticular, spiral and barred spiral.

The innovative aspect of the spiral model relies in its dual behaviour. One part of the model

is a traditional constructive parametric model which provides a way to simulate this structure. On the other hand, we introduce a novel modelling technique which is based on the introduction of a parametric filtering of the input image. The parametric filter modifies the observation to remove some of its non-desirable aspects (like non-relevant features). This approach differs from a traditional pre-processing of the observation because the parameters of the filter are estimated like the other parameters of the model. Moreover, we show that, even if the filter consists in masking parts of the observation, we can still preserve useful information about the masked regions in the estimation process.

To improve the performances of the optimization algorithm we describe two new approaches that allow to reduce the number of needed iterations to only a few hundreds compared to the tens of thousands usually required. Our first idea relies on a simulated annealing strategy with multiple temperatures. Compared with traditional simulated annealing [10, 20, 21], our approach assigns a different temperature to each dimension of the posterior. This method leads to a better distribution of the computational effort.

A second idea is to introduce an iterative decomposition strategy based on a hierarchy of models. The hierarchy is composed of models of increasing complexity. The estimates of a model are used as input for the estimation of the model of higher complexity. This way, we strongly decrease the length of the burn-in period (first iterations discarded because they are not representative of the equilibrium distribution of the Markov chain in a short run) and finally, we are able to estimate the full hierarchy of models faster than the most complex model taken alone.

The good behaviour of the method is demonstrated on simulated and real images. The simulated images were produced using the parametric model of barred spiral galaxies and are used to produce quantitative measurement on the quality of the method. The real galaxies are 5-band images from near ultra-violet (UV) to near infrared (IR) selected from the EFIGI database [22]. These images were extracted from the Sloan Digital Sky Survey archive (SDSS[1]); collected with a dedicated 2.5-m wide-angle optical telescope at Apache Point Observatory in New Mexico this dataset allows one to extract a well-defined sample of galaxy images spanning all the morphological types. Results obtained with this multiband data show the accuracy and the robustness of the proposed method.

Compared to previous works on galaxy decomposition, this paper introduces a support for multispectral images and a new parametric model for spiral arms. The parametric model is described in section 2. It is based on a global approach which incorporates the expert knowledge (prior on the shape and the brightness of the galaxy structures) and the multispectral information in a natural way. The Bayesian framework (section 3) allows to rigorously take account of the noise and the intrinsic variability of the observations. The Bayesian scheme is adapted to include our new parametric filtering strategy and an estimation method is proposed. The details of the estimation algorithm are given in section 4. In section 5, we present the two new techniques designed to improve the estimation algorithm: a simulated annealing using several temperatures and a hierarchical decomposition scheme providing an accurate, robust and fast initialization procedure. The results on simulated and real images are presented in section 6 with a quantitative and a qualitative analysis. Finally, we conclude our work and discuss the perspectives in section 7.

2. Model of a Barred Spiral Galaxy

Galaxies can be decomposed into different visible structures which are both morphologically and physically significant. Each structure can be distinguished by its shape and location within the galaxy, by its mix of stellar populations, average chemical composition, internal kinematics, etc. In the case of barred spiral galaxies, four main components are evidenced (the spherical stellar halo component where globular clusters reside and the invisible and massive halo of dark matter are not

---

[1]http://www.sdss.org/

considered): a bulge ($\mathfrak{B}$), a disc ($\mathfrak{D}$), a so-called stellar bar ($\mathfrak{Ba}$) and spiral arms ($\mathfrak{s}$). The bulge is the central tight and bright spheroidal component made mostly of old red stars. Beyond the bulge is the disc: it is a planar structure with a circular shape, a small scale height and a distinct light profile; it includes most of the gas and dust of the galaxy and it owns a population of stars of various ages from old red to younger and bluer one. The angle between the line-of-sight and the plane defined by the disc gives the inclination of the galaxy on the celestial sphere: for an inclination of 90° (respectively 0°) we say that the galaxy is viewed face-on (respectively edge-on). The stellar bar is an elongated structure in the galactic disc extending out from the bulge with a population of stars similar in colour to the one of the bulge but with a different kinematic. Finally, the spiral arms are over dense and bright regions in the disc induced by a density wave. They are of great importance since the star formation regions are located there; spiral arm observations can thus differ importantly from one band to another one since the short-lived massive and energetic stars lead either to a pattern of spotty regions (called H II regions) when viewed at short wavelengths or to a continuous feature in the IR due to the heating of the dust content of the disc (Fig. 2).
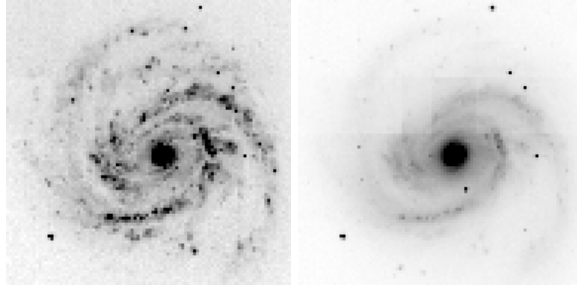

Figure 2: The barred spiral galaxy PGC (principal galaxies catalogue) 40153 from the SDSS in band u (near UV) and z (near IR). The z-band image is a mosaic of several images which explains its blocky aspect.

2.1. Observation Model

According to astronomical considerations, and neglecting the obscuring effect of any irregular dust distribution, an observation of a galaxy is composed of the sum of the brightness of its different structures affected by perturbations due to the atmosphere, the optics of the telescope, the characteristics of the detector and corrupted by noise. Thus, an observation $\mathbf{Y}$ can be modelled as:

$$\mathbf{Y} = \mathbf{H} * \mathbf{F}(\boldsymbol{\phi}) + \mathbf{e} \qquad \text{with} \qquad \mathbf{F}(\boldsymbol{\phi}) = \mathbf{f}_{\mathfrak{B}}(\boldsymbol{\phi}) + \mathbf{f}_{\mathfrak{D}}(\boldsymbol{\phi}) + \mathbf{f}_{\mathfrak{Ba}}(\boldsymbol{\phi}) + \mathbf{f}_{\mathfrak{s}}(\boldsymbol{\phi}) \qquad (1)$$

where $*$ denotes the convolution, $\mathbf{f}_{\mathfrak{B}}$ (respectively $\mathbf{f}_{\mathfrak{D}}$, $\mathbf{f}_{\mathfrak{Ba}}$ and $\mathbf{f}_{\mathfrak{s}}$) is the brightness of the bulge (respectively the stellar bar, the disc and the spiral arms). $\mathbf{H}$ the convolution kernel stands for the point spread function (psf), it is generally estimated with accuracy by deterministic methods (so we assume that it is known). The noise $\mathbf{e}$ is a composition of several sources [23]: photon noise, thermal noise, electronic noise. It is approximated by a multidimensional uncorrelated zero-mean Gaussian noise $\mathcal{N}(0, \boldsymbol{\Sigma})$ whose covariance matrix $\boldsymbol{\Sigma}$ can also be easily estimated [24] and thus is considered known. Finally, the model is parametrized by unknowns $\boldsymbol{\phi}$.

All structures are modelled using parametric functions and each one can be decomposed into two descriptors:

- A shape descriptor carries the spatial information: its parameters contain the most important information about the morphology of the structure;

- A brightness descriptor models the distribution of the flux coming from the structure.

The combination of the two descriptors provides a complete discriminant model for the object.
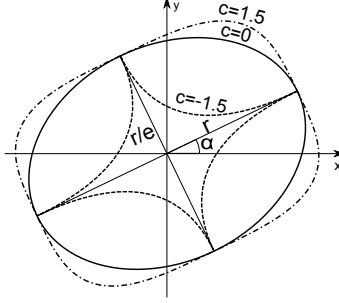
Figure 3: Three super-ellipses with the same position angle α, major axis r and minor axis r/e but different misshapenness c (displayed near each ellipse).

We will now describe the models used for the different structures. A common model for the bulge, the disc and the stellar bar already exists in the literature for monospectral images [16, 17]: we present this model generalized to multispectral images in section 2.2. Then we propose in section 2.3 a new model for the spiral arm component. At last in section 2.4, we sum up the variables gathered in $\boldsymbol{\phi}$ and give some relations and constraints between them.

2.2. Bulge, Disc, Stellar Bar

In this model, a generalized ellipse (also known as super-ellipse) is used as the shape descriptor to define the shape of the isophotes (lines of equal brightness) of the structure. The traditional Sérsic law is the brightness descriptor which describes the evolution of the brightness of the isophotes from the centre to the outer edge.

2.2.1. Shape Descriptor

Major axis r of a generalized ellipse passing trough point $(x, y) \in \mathbb{R}^2$ is given by:

$$r\left(x, y\right) = \left(\left|(x - c_x)\cos\alpha - (y - c_y)\sin\alpha\right|^{c+2} + \left|e\left((x - c_x)\sin\alpha + (y - c_y)\cos\alpha\right)\right|^{c+2}\right)^{\frac{1}{c+2}} \quad (2)$$

where (c_x, c_y) denotes the centre of the ellipse, e is the elongation (ratio of the minor to the major axis), α is the position angle (angle between the major axis and the abscissa axis) and c controls the misshapenness: if $c = 0$ the generalized ellipse reduces to a simple ellipse, if $c < 0$ the ellipse is said to be *disky* and if $c > 0$ the ellipse is said to be *boxy* (Fig. 3).

2.2.2. Brightness Descriptor

The Sérsic law [25] is used to model the brightness profile. It is a generalization of the traditional exponential and de Vaucouleurs laws usually used to model the disc and bulge brightness profile, respectively. Its high flexibility allows it to vary continuously from a nearly flat curve to a very piked one (Fig. 4). The brightness at radius r along the major axis is given by:

$$I(r) = I \exp\left(-k_n\left(\left(\frac{r}{R}\right)^{\frac{1}{n}} - 1\right)\right) \quad (3)$$

where R is the effective radius, n is the Sérsic index, and I the brightness at the effective radius. k_n is an auxiliary function such that $\Gamma(2n) = 2\gamma(2n, k_n)$ to ensure that half of total flux is contained in the effective radius (Γ and γ are respectively the complete and incomplete gamma functions).
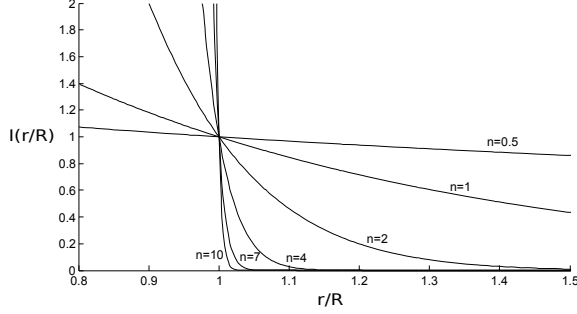
Figure 4: The Sérsic law for different Sérsic index n. $n = 0.5$ yields a Gaussian, $n = 1$ yields an exponential profile and for $n = 4$ we obtain the De Vaucouleurs profile.

2.2.3. Combination of the Descriptors

the brightness at pixel (x, y) on B spectral bands is given by:

$$\mathbf{f}(\boldsymbol{\phi}, x, y) = (f_1(x, y), \dots, f_B(x, y)) \qquad (4)$$

where the brightness in band b is defined as:

$$f_b(x, y) = \begin{cases} I_b \exp\left(-k_{n_b}\left(\left(\frac{r(x,y)}{R_b}\right)^{\frac{1}{n_b}} - 1\right)\right) & \text{if } r(x, y) \leq R_{max} \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

where R_{max} is the cut-off radius from which the brightness vanishes. $r(x, y)$ is defined in (2). In a discrete space, we assume that the pixel value is equal to the integral of the above formula on the pixel area. For computing simplicity, the integral value is approximated by the central value of the pixel in most part of the object. However, because the Sérsic function has a huge dynamic for small values of $r(x, y)$, the error increases quickly toward the centre so a special care must be taken to perform a numerical integration of the central pixels. As a compromise between approximation error and computation speed, we choose to over-sample the pixels near the centre within a radius of three pixels.

2.3. Spiral arms

To our knowledge, despite the importance of the spiral arms in the classification scheme, no parametric model for simulating the spiral arms is available, even in the monospectral case. Such a model should be able to take account of the continuous aspect of the arms which is prominent at long wavelengths as well as the spotty aspect which may dominate the structure at short wavelengths. The shape of the continuous aspect (development and tightness) is important in the Hubble's classification scheme [19, 26] and, in consequence, it has to be accurately modelled. On the contrary, the spotty aspect of the H II regions is only a marker of the spiral structure.

Thus, we propose an original model divided into two parts: a parametric model for the continuous aspect of the spiral and an adaptive masking technique for the spots which aims at masking the H II regions located on the continuous model. This second aspect is detailed in section 3.1 while the parametric model is presented below.

Following the rules used for the bulge, the disc and the stellar bar, the model for the spiral arm component is composed of a shape descriptor and a brightness descriptor.

2.3.1. Shape Descriptor

The detailed structure of the spiral arms can be very complex. However, most of the time, two main symmetric arms can be isolated. Our model concentrates on describing these two main arm

6

cases. In [27], the authors have shown that the distribution of the main components in a face-on galaxy follows a logarithmic spiral: this is our shape descriptor.

The set of points belonging to a symmetric 2-arm logarithmic spiral curve is given by:

$$E_{\mathfrak{s}} = \left\{ (r, \rho) \in \mathbb{R}^+ \times \mathbb{R} \;\middle|\; r = R_0 \exp\left(|o_{\mathfrak{s}}|\theta\right), \; \rho = o_{\mathfrak{s}}\theta + \theta_0 + k\pi, \; \theta \in \mathbb{R}^+, \; k \in \{0, 1\} \right\} \qquad (6)$$

with $o_{\mathfrak{s}}$ the pitch angle, R_0 the starting radius and θ_0 the position angle (Fig. 5).
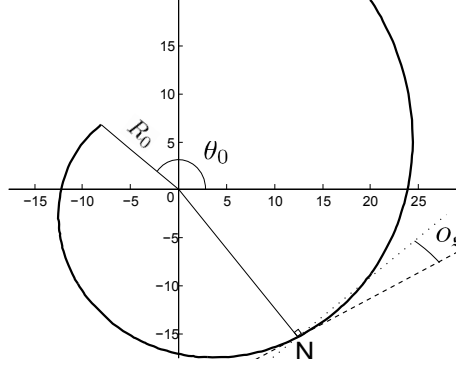

Figure 5: A logarithmic spiral and its fundamental parameters. For all points N of the spiral, the pitch angle $o_{\mathfrak{s}}$ is the constant angle between the perpendicular (dotted line) in N of the line passing through the origin and the tangent (dashed line) of the spiral in N.

As the galaxy is not necessarily seen face-on, the logarithmic spiral has to be projected on the observation plane. Because the projection is fully determined by the disc parameters, the shape of the spiral arms is linked to the parameters of the shape descriptor of the disc (which is assumed to be a thin circular structure): the centre (c_x, c_y), the position angle α and the elongation e.

A pseudo distance function that defines the isophote shape of the spiral over the whole space constitutes the second part of the model. From the fact that the spiral arm brightness decreases both with the distance from the centre and from the spiral crest, we define the distance function of a point $M(r, \theta)$ in two parts:

$$d_{\mathfrak{s}}(r, \theta) = d_{\mathfrak{s},1}(r, \theta) + d_{\mathfrak{s},2}(r, \theta) \qquad (7)$$

where $d_{\mathfrak{s},2}(r, \theta)$ is the Euclidean distance from the given point M to the closest point of the spiral $P = (r', \theta')$, and $d_{\mathfrak{s},1}(r, \theta)$ is a function of the Euclidean distance between this closest point to the centre $O = (0, 0)$ of the galaxy in the galactic plane (Fig. 6):

$$d_{\mathfrak{s},1}(r, \theta) \quad = \quad \exp\left(\frac{-r'}{R_e}\right) \qquad (8)$$

$$d_{\mathfrak{s},2}(r, \theta) \quad = \quad \left\|\overrightarrow{PM}\right\| \qquad (9)$$

with R_e a scale parameter of the spiral and

$$P = \arg\min_{P \in E_{\mathfrak{s}}} \left\|\overrightarrow{PM}\right\| \qquad (10)$$

Because there is no analytic formulation for (10), there is no analytic formula to compute $d_{\mathfrak{s}}$. But, on the crest of the spiral, $d_{\mathfrak{s},2}$ is equal to 0 and $d_{\mathfrak{s},1}$ is trivial, thus we can directly compute $d_{\mathfrak{s}}$ for the points of $E_{\mathfrak{s}}$. As $d_{\mathfrak{s},2}$ is a distance from the points of $E_{\mathfrak{s}}$ to the others, $d_{\mathfrak{s}}$ can be computed over the whole space by performing a distance transform to the set $E_{\mathfrak{s}}$ valued by the distance $d_{\mathfrak{s}}$. The distance $d_{\mathfrak{s},2}$ is measured in the galactic plane so, the equivalent distance in the projection
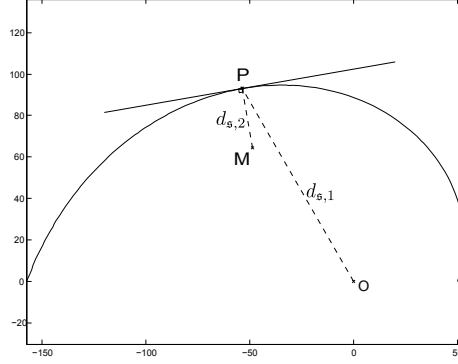
Figure 6: Definition of the distance function $d_{\mathfrak{s}} = d_{\mathfrak{s},1} + d_{\mathfrak{s},2}$ between point M and O the centre of the spiral. P is the closest point of M on the spiral crest.

plane is not the Euclidean distance any more. Nevertheless, $d_{\mathfrak{s}}$ can be computed over the whole plane using a traditional distance transform with an adapted distance mask because the unit circle in the projected distance is an ellipse which is convex [28]. The three steps to simulate the spiral arms are shown in Fig. 7.

Figure 7: The three steps of spiral arms simulation. Left: the crest of the spiral is drawn with the projection effect, the values on the spiral crest are the values of $d_{\mathfrak{s},1}$ (the distance increases when brightness increases), the background value is ∞. Centre: the distance transform is calculated with a projected distance mask and allows to compute $d_{\mathfrak{s}}$ over the whole simulation. Right: the brightness descriptor is applied on the distance map (image is given in inverse greyscale).

2.3.2. Brightness descriptor

Again, we choose to use the Sérsic function which is able to model various decreasing schemes. Then, the brightness of the spiral at pixel (r, θ) is given by:

$$\mathbf{f}(\boldsymbol{\phi}, r, \theta) = (f_1(r, \theta), \dots, f_B(r, \theta)) \tag{11}$$

with B the number of bands and the brightness in band b defined as:

$$f_b(r, \theta) = I_b \exp\left(-k_{n_b}\left(\left(\frac{d_{\mathfrak{s}}(r, \theta)}{R_b}\right)^{\frac{1}{n_b}} - 1\right)\right) \tag{12}$$

2.4. Constraints over the Parameters

The constraints adopted for the model were chosen from astronomical considerations or from empirical observations. The number of constraints is kept as low as possible to ensure the physical consistency of the result by avoiding discrepancies.

2.4.1. Bulge, Disc and Stellar Bar

We assume that shape parameters do not vary between bands because each structure represents a particular mix of stars and dust and gas content which do not depend from the observed band. This strong assumption, which is observed in our real images over the involved wavelength domain, is an important step in multiband processing to ensure the inter-band coherence of the model. For the same reasons we assume that the center is identical for all structures. These assumptions reduce significantly the numbers of free parameters.

On the other hand, the parameters of the Sérsic function depend on the observed band, but we set, for each structure, the maximum relative variation of R_b and $R_{b'}$ between two bands b and b' to be less or equal than 10%: $\max\left(\frac{|R_b - R_{b'}|}{\max(R_b, R_{b'})}\right) \leq 0.1$. Similarly, the maximal difference between two Sérsic indexes (n_b and $n_{b'}$) is set to 1: $\max\left(|n_b - n_{b'}|\right) \leq 1$. At last the effective brightness of all structures in all bands is positive.

The set of constraints specific for each structure at a given wavelength is given in table 1.

Table 1: Specific constraints of each structures at a given wavelength b. R_b and R_{max} are given in pixels.

	Bulge	Disc	Stellar Bar
Sérsic index	$1 \leq n_b \leq 8$	$n_b = 1$	$0.5 \leq n_b \leq 1$
elongation	$1 \leq e \leq 3.3$	$1 \leq e \leq 8$	$2 \leq e \leq 8$
misshapenness	$c = 0$		$0.5 \leq c \leq 2$
cut-off radius	$R_{max} = \infty$		$10 \leq R_{max} \leq 200$
effective radius	$1.5R_{b,\mathfrak{B}} \leq R_{b,\mathfrak{D}} \leq 200$		$0 \leq R_b \leq 200$

2.4.2. Spiral Arms

The main constraint concerns the link between the stellar bar and the arms. Because the spiral arms do start where the stellar bar ends, we impose that the difference between the position angle of the arms and the stellar bar must be less or equal to 0.3 rad and that the difference between the bar cutoff radius and the arms starting radius must be less or equal to 4 arcsec[2]. These values were determined heuristically from real observations. Because a too low pitch angle absolute value can lead to discrepancies (the spiral degenerates to a ring with a pitch angle of 0) we have fixed a low limit to the absolute value of the pitch angle at 0.1 rad.

2.4.3. Complete model

Figure 8 is a summary of the different structures and all their parameters, illustrated by a simulation of a barred spiral galaxy. The total number of parameters of the model is equal to $13 + 11B$ with B the number of bands in the observation. The shape descriptor of the bulge, the disc and the stellar bar has 5 parameters, the Sérsic function has 3 parameters. Thus we obtain $4 + 3B$ parameters for the bulge (misshapenness of the generalized ellipse is constant), $4 + 2B$ parameters for the disc (misshapenness of the generalized ellipse and Sérsic index are constant) and $5 + 3B$ for the stellar bar. As the centre is identical for each structure the total number of parameters for the 3 structures is equal to $9 + 8B$. Finally the spiral structure has $4 + 3B$ parameters, 4 parameters for the shape descriptor and $3B$ for the Sérsic function (parameters coming from the disc to determine the projection are not counted). Most nowadays astronomical surveys provide 5-band images thus leading to 68 parameters to estimate in the inverse problem.

---

[2]The arcsecond is an angular distance equals to 1/60 of arcminute $= \frac{\pi}{648000}$ rad.
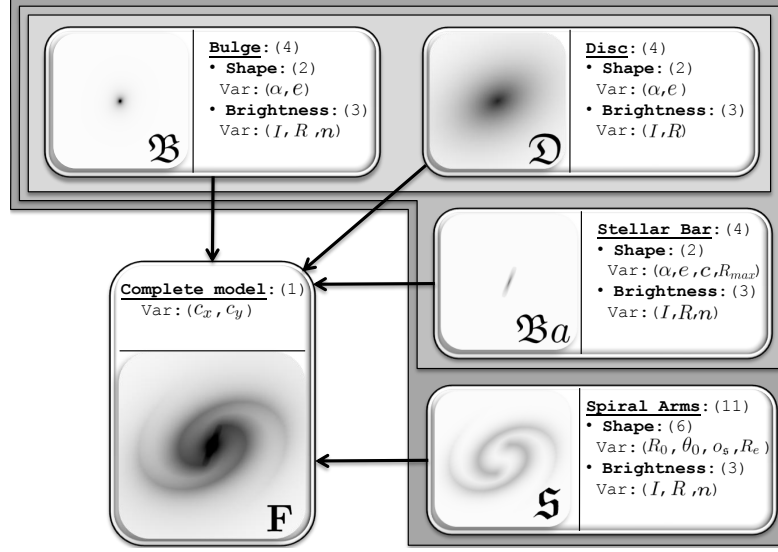
Figure 8: Example of simulated barred spiral galaxy with our model. For each structure, we indicate the equations describing the structure and its variables.

3. Proposed Approach

3.1. Adaptive Masking for H II Regions in Spiral Arms

As mentioned previously, the parametric model of the spiral arms only models their continuous aspect. But, at short wavelengths, this continuous aspect may be dominated by more or less compact H II regions containing young and very bright stars which causes a spotty aspect. Neither the shape nor the brightness profile of the H II regions carry interesting information for a classification purpose. Only their spatial distribution is an important marker of the spiral structure. Thus, rather than developing a complex parametric model for the H II regions we prefer to mask them in the observation.

The decision to mask a particular region is based on an initial map M containing all possible H II regions and on the parameters of the continuous model. The idea is to mask only the components of M which lie on the estimated spiral. Because the masking is conditional to a spatial information carried by the parameters of the continuous model, it allows to mask information in the observation while preserving the information about the location of this masking. In this way, we prevent the estimation of the brightness parameters of the spiral to be biased by the brightness of the H II regions and we preserve the spatial information of their distribution.

Formally, we mask a connected component of M if the distance $d_{\mathfrak{s},2}$ from its centre to the crest of the spiral is less or equal to three arcsec. The filter defined by the masking of selected H II regions will be denoted by $\mathbf{A}_{\mathrm{m}}\left(\cdot,\cdot\right)$. Thus, the observation $\mathbf{Y}$ is itself modified by the mask in function of parameters $\boldsymbol{\phi}$, yielding to the new data $\mathbf{A}_{\mathrm{m}}\left(\mathbf{Y},\boldsymbol{\phi}\right)$. Equation (1) reads then:

$$\mathbf{A}_{\mathrm{m}}\left(\mathbf{Y},\boldsymbol{\phi}\right) = \mathbf{H} * \mathbf{F}(\boldsymbol{\phi}) + \mathbf{e} \tag{13}$$

3.2. Bayesian Model

The ill-posed nature of the problem has been demonstrated for the model including only the bulge and the disc structure in monoband images by showing that several solutions may exist [16, 29, 30]. It seems clear that a more complex model including more structures in multispectral images

will lead to the same issues. We have adopted a Bayesian approach, which has the advantage to rigorously take account of the incertitude [7]. Equation (13) can be written as:

$$\mathbf{e} = \mathbf{A}_{\mathrm{m}}\left(\mathbf{Y}, \boldsymbol{\phi}\right) - \mathbf{H} * \mathbf{F}(\boldsymbol{\phi}) \tag{14}$$

which gives the following formulation of the posterior under the zero mean Gaussian noise hypothesis:

$$P\left(\boldsymbol{\phi}|\mathbf{Y}\right) \propto \tag{15}$$

$$P\left(\boldsymbol{\phi}\right) \frac{1}{\left(2\pi\right)^{\frac{N}{2}} \det\left(\mathbf{\Sigma}\right)^{\frac{1}{2}}} \quad \exp\left(-\tfrac{1}{2}\left(\mathbf{A}_{\mathrm{m}}\left(\mathbf{Y}, \boldsymbol{\phi}\right) - \mathbf{H} * \mathbf{F}(\boldsymbol{\phi})\right)^{T} \mathbf{\Sigma}^{-1}\left(\mathbf{A}_{\mathrm{m}}\left(\mathbf{Y}, \boldsymbol{\phi}\right) - \mathbf{H} * \mathbf{F}(\boldsymbol{\phi})\right)\right)$$

where $P\left(\boldsymbol{\phi}\right)$ denotes the priors and N is the dimension of the observation. Prior assigned to each parameter is determined by the set of constraints given in section 2.4 and is thus product of bounded uniform priors. One can note that due to the definition of $\mathbf{A}_{\mathrm{m}}\left(\cdot, \cdot\right)$ the dimension of the observation can change in function of $\boldsymbol{\phi}$ leading to an original problem of estimating parameters from an observation with varying dimension.

The conditional posterior pdf of the brightness factor parameter I_b for each structure reduces to a truncated (positive) Gaussian pdf which can be efficiently sampled using an accept-reject algorithm [31]:

$$I_b \sim \mathcal{N}^{+}\left(\mu, 1/A\right) \tag{16}$$

The proof and the values of μ and $1/A$ are given in Appendix AppendixA. This expression also leads to the integration of one brightness factor from the posterior pdf. It reads then:

$$P(\boldsymbol{\phi}_{\backslash\{I_b\}}|\mathbf{Y}) \propto \left(\frac{D}{\left(2\pi\right)^{\frac{N}{2}} |\mathbf{\Sigma}|^{\frac{1}{2}}}\right) \exp\left(\epsilon\right) P(\boldsymbol{\phi}_{\backslash\{I_b\}}) \tag{17}$$

where $\boldsymbol{\phi}_{\backslash\{I_b\}}$ denotes the set of parameters deprived of I_b. The proof and the value of D and ϵ are given in Appendix AppendixB. Unfortunately calculus become rapidly intractable when one wants to integrate several brightness factors at the same time. The integration strategy will be described in section 4.

Indeed, due to the high dimensionality of the problem, it is intractable to characterize completely the posterior pdf, thus we aim at finding the maximum a posteriori (MAP). We now give an overview of the estimation algorithm which is based on MCMC methods.

3.3. Estimation Algorithm

The estimation algorithm robustness is a crucial point. Indeed, we need to process thousands of galaxies and thus the algorithm must tolerate large approximation on the initial parameters so that it can be used in an unsupervised mode. Several authors have pointed out the difficulty of providing a fully automatic algorithm for the estimation of the parameters of the much more simpler two components (bulge and disc) model in mono band images [16, 29, 30]. Obviously, the estimation of a more complex model generates more difficulties. To overcome this problem we propose to use MCMC methods. MCMC algorithms allow to sample the parameter space according to the target distribution and theoretical results prove the convergence of the distribution of the samples to the target distribution in infinite time.

Nevertheless, the main disadvantage of MCMC algorithms is their convergence time. In our application, tens of thousands iterations (and sometimes much more) are required before the MAP can be reached. This point is of prime interest for us because one iteration of a MCMC algorithm requires several evaluations of the likelihood which have a heavy computational cost. Indeed, the simulation of the parametric model requires, for each pixel of the observation, the computation of several costly functions: exponential, power, trigonometric functions, a distance transform and a

final convolution operation. That is why the effort is put on the minimization of the number of required iterations by using robust adaptive methods and a new simulated annealing technique. We will show that, with this efficient design, we can reduce the simulation effort to only a few hundreds of iterations.

In this section, we presented the Bayesian framework coupled to a new parametric filtering strategy. We also gave the key points of the estimation algorithm whose details are provided in the next section.

4. Estimation Algorithm

4.1. Gibbs sampler

The main algorithm is a Gibbs sampler [10, 21, 32] which decomposes the sampling of a high dimensional pdf by simulating each variable separately according to its conditional pdf. The variable are not necessarily scalar and can gather a vector of parameters. In the following we call *block* a vector of parameters which are updated together in the Gibbs algorithm.

To improve the efficiency of the algorithm the blocks have been defined according to the correlation between parameters [21]. Shape parameters are weakly correlated and are thus sampled independently. On the contrary, Sérsic parameters (R_b, n_b and I_b) are highly correlated, so all variables of the block (R_b, n_b, I_b) (or (R_b, I_b) if n_b is constant) are sampled together. Because only one brightness factor can be integrated in the posterior (17), we propose to integrate a different brightness factor at different steps of the algorithm. Indeed, one can note that their conditional posterior pdf can be decomposed into:

$$P\left(R_b, n_b, I_b \mid \boldsymbol{\phi}_{\backslash\{R_b, n_b, I_b\}}, \mathbf{Y}\right) = P\left(R_b, n_b \mid \boldsymbol{\phi}_{\backslash\{R_b, n_b, I_b\}}, \mathbf{Y}\right) P\left(I_b \mid \boldsymbol{\phi}_{\backslash\{I_b\}}, \mathbf{Y}\right)$$

Thus, according to [33], the pdf $P\left(R_b, n_b, I_b \mid \boldsymbol{\phi}_{\backslash\{R_b, n_b, I_b\}}, \mathbf{Y}\right)$ can be simulated in two steps, by first sampling (R_b, n_b) according to $P\left(R_b, n_b \mid \boldsymbol{\phi}_{\backslash\{R_b, n_b, I_b\}}, \mathbf{Y}\right)$ (the conditional posterior where I_b has been integrated) and then sampling I_b according to $P\left(I_b \mid \boldsymbol{\phi}_{\backslash\{I_b\}}, \mathbf{Y}\right)$. The different steps of the Gibbs sampler are summarized in algorithm 1.

Algorithm 1 Main Gibbs sampler

Require: initial estimation $\boldsymbol{\phi}$
1: **repeat**
2:     sample $(c_x, c_y) \sim P(c_x, c_y \mid \mathbf{Y}, \boldsymbol{\phi}_{\backslash\{c_x, c_y\}})$
3:     **for all** structures $\mathbf{f} \in \{\mathbf{f}_{\mathfrak{B}}, \mathbf{f}_{\mathfrak{D}}, \mathbf{f}_{\mathfrak{B}\mathfrak{a}}, \mathbf{f}_{\mathfrak{s}}\}$ **do**
4:         **for all** shape parameters φ of component $\mathbf{f}$ **do**
5:             sample $\varphi \sim P(\varphi \mid \mathbf{Y}, \boldsymbol{\phi}_{\backslash\{\varphi\}})$
6:         **end for**
7:         **for all** bands b **do**
8:             sample $(R_b, n_b) \sim P\left(R_b, n_b \mid \mathbf{Y}, \boldsymbol{\phi}_{\backslash\{R_b, n_b, I_b\}}\right)$
9:             sample $I_b \sim P\left(I_b \mid \boldsymbol{\phi}_{\backslash\{I_b\}}, \mathbf{Y}\right)$
10:         **end for**
11:     **end for**
12: **until** Convergence criterion is satisfied

4.2. Hastings-Metropolis sampling

There are no classical expression for the posterior pdfs and thus, there is no direct simulation algorithm for them. The usual way, in this case, is to perform Hastings-Metropolis (HM) steps within the Gibbs sampler [10]. In fact, HM within Gibbs reduces to a complex HM algorithm

ensuring that all convergence properties are preserved [34]. Due to the lack of knowledge on the posterior definition, we have opted for the random walk version of the HM algorithm [10, 21]. In this case the proposal distribution q is defined as $q(. \mid X) = X + q'(.)$ where q' is a distribution which does not depend of X and is usually centred on 0. By this way, at each iteration, a random move is proposed from the actual position. The choice of a good proposal distribution is crucial to obtain an efficient algorithm, and literature suggest that, for low dimension variables, a good proposal should lead to an acceptance rate of 0.5 [35].

In case of a scalar block, the determination of a good proposal distribution can be solved by using a standard Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with an adaptive scale technique [9]. The idea is to adjust the standard deviation σ to achieve the optimal acceptance rate. If the measured acceptance rate is too high the standard deviation is decreased and vice versa. In our scheme, the scale parameter is updated every 20 iterations (adaptation time) following [9]:

$$\log(\sigma_{i+1}) = \log(\sigma_i) + \frac{\log(A_i) - \log(1 - A_i) - \log(A) + \log(1 - A)}{d} \tag{18}$$

where σ_i and σ_{i+1} are respectively the old and the new standard deviations, $A = 0.5$ is the target acceptance rate, A_i is the measured acceptance rate during the last 20 iterations and d is the dimension of the simulated law. One can note that, due to the constant adaptation time, the property of convergence of the algorithm is compromised [8]: this problem is addressed in section 4.4.

In the case of a vectorial block, we use the so called Snooker algorithm which is an adaptive direction sampling method proposed by Gilks et al. [10]. The Snooker algorithm uses a set of already generated samples to determine the preferred directions; we choose a set size of 10 samples. The first 10 samples are generated with a traditional random walk HM algorithm with a multidimensional Gaussian proposal.

4.3. Convergence criterion

Numerous criteria have been proposed to assess the convergence of MCMC algorithms (the reader can refer to [36, 37] for example). Most of these criteria require to run multiple chains, which is not an affordable solution in our application. So we draw our inspiration from [38] and [39] and we monitor the convergence of the mean value of the chain by constructing an asymptotically valid confidence interval. Because of the high dimensionality of our chain, building a confidence interval may be very difficult, thus we monitor only the evolution of the mean normalized square error (MNSE) defined as the mean log likelihood per pixel:

$$\text{MNSE}(\boldsymbol{\phi}) = \frac{1}{N} \left(\mathbf{A}_\mathrm{m}\left(\mathbf{Y}, \boldsymbol{\phi}\right) - \mathbf{H} * \mathbf{F}(\boldsymbol{\phi}) \right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{A}_\mathrm{m}\left(\mathbf{Y}, \boldsymbol{\phi}\right) - \mathbf{H} * \mathbf{F}(\boldsymbol{\phi}) \right) \tag{19}$$

Formally, let $S = \{E_1, \ldots, E_n\}$ be the set of MNSE at the end of each iteration. Let $\hat{\rho}_i^2$ be the estimated variance of the samples $\{E_{i-S+1}, \ldots, E_i\}$, where S is the size of the estimation window. The convergence is assumed to be reached if the following inequality is satisfied [38]:

$$t\frac{\hat{\rho}_n}{\sqrt{S}} < \epsilon \tag{20}$$

where t is an appropriate Student's t quantile and ϵ the desired half wide of the confidence interval. We have adopted $t = 90\%$, $\epsilon = 0.05$ and $S = 25$.

4.4. MCMC estimator

When the convergence is diagnosed, all adaptive techniques are disabled and the algorithm is run for a fixed number of iterations ($T = 50$). This number was fixed empirically to provide a fairly accurate estimate for a reasonable simulation effort. This final run is used to estimate the

MAP and its variance. This is the easiest way to ensure that the chain stationary distribution is not compromised by adaptive HM [8, 40]. Let $\boldsymbol{\Phi} = \{\boldsymbol{\phi}_0, \ldots, \boldsymbol{\phi}_T\}$ be the set of the last T generated samples. The MAP estimation is a direct research of the sample with the maximum posterior probability:

$$\hat{\boldsymbol{\Phi}}_{MAP} = \arg\max_{\boldsymbol{\phi} \in \boldsymbol{\Phi}} P(\boldsymbol{\phi}|\mathbf{Y}) \tag{21}$$

We now describe the estimator of the variance. Due to correlation in the chains, we use the overlapping batch means (OBM) estimator as suggested in [41]. The OBM estimator is given by:

$$\hat{V}^2_{OBM} = \frac{bT}{(T-b)(T-b+1)} \sum_{i=b}^{T} (\langle\boldsymbol{\Phi}\rangle_i - \langle\boldsymbol{\Phi}\rangle)^2 \tag{22}$$

where $\langle\boldsymbol{\Phi}\rangle$ is the mean of the $\boldsymbol{\phi}_i$ and $\langle\boldsymbol{\Phi}\rangle_i$ is the partial mean on b samples on a sliding window:

$$\langle\boldsymbol{\Phi}\rangle_i = \frac{1}{b} \sum_{j=i-b+1}^{i} \boldsymbol{\phi}_j \tag{23}$$

where $b = \lfloor\sqrt{T}\rfloor$ is the size of the estimation window [41].

5. Improvements and Initialization

In this section, we present two new improvements added to our estimation algorithm. The first one concerns a novel simulated annealing scheme which uses a different temperature for each block of the Gibbs sampler, as well as a general method to automatically determine the initial temperature. The second improvement concerns a hierarchical decomposition scheme which provides an efficient decrease of the burn-in period. Finally, we discuss about the initialization procedure.

5.1. Multiple Temperatures Simulated Annealing

To speed up the exploration of the space solution, we have set up a special simulated annealing scheme. The idea of using the simulated annealing in a MCMC context was introduced by [42]: the principle consists in modifying the target pdf, so the algorithm can easily jump from one mode of the target pdf to another one. Usually, the pdf is modified to become $P(\boldsymbol{\phi}|\mathbf{Y})^{\frac{1}{t}}$ whose temperature t is decreased using a logarithmic scheme to ensure the convergence [32]. This approach has some drawbacks. First, the theoretical condition of a logarithmic temperature decreasing requires too much time and another faster scheduling must be used in practice. Another issue is that the same temperature is applied to the overall distribution without considering the difference between each dimension. This is a serious drawback in applications where the conditional posterior of some parameters (e.g. the centre of the galaxy or the position angles of the structures) are quite simple with a well defined maximum and thus do not require hot temperatures whereas others (like the Sérsic parameters) are strongly coupled with several pronounced local maxima and thus require high temperatures to be efficiently sampled.

We propose to use an independent temperature for each block of the Gibbs sampler. In this approach the temperature is proportional to the complexity of the posterior pdf in the dimension of the considered block. In this way, the posterior pdf is smoothed differently in each dimension, making the posterior pdf easier to explore in all dimensions. Another interpretation of this approach is to consider that we are swapping between multiple HM algorithms sampling from different posterior distributions. But, as in the traditional simulated annealing the final MAP estimation remains the same.

Nevertheless, as the produced Markov chain is no longer homogeneous, the stationary distribution of the chain is not the posterior pdf and thus the samples cannot be used to estimate the

variance of the posterior pdf. This problem is easily solved by freezing the temperature to its neutral value $1\,^{\circ}$ during the final run of iterations used for the estimation.

Rather than a logarithmic decreasing, we adopt the geometric step scheduling [43] which provides an adaptive way to modify the temperature in complex optimization problems. In this scheme, the temperature is adjusted every p iterations (step length). The temperature can either decrease or stay the same depending on the evolution of the energy of the system between the current and the previous step. If the energy decreases, the temperature does not evolve. On the contrary if the energy increases the temperature is decreased by a geometric factor. Let $\boldsymbol{\phi}^j \subset \boldsymbol{\phi}$ be a block of the Gibbs sampler. Let α^j be the geometric factor. T_i^j is the temperature at step i of block j, we define an energy function $U(\boldsymbol{\phi}) = -\log P(\boldsymbol{\phi} \mid \mathbf{Y})$ and the mean energy $\langle U \rangle$ at step i:

$$\langle U \rangle_i = \frac{1}{p}\sum_{k=i}^{i+p} U(\boldsymbol{\phi}_k) \tag{24}$$

the temperature evolution is then given by the following relation:

$$T_{i+2}^j = \left\{ \begin{array}{ccc} T_{i+1}^j & \text{if} & \langle U \rangle_{i+1} \le \langle U \rangle_i \\ \alpha^j T_{i+1}^j & \text{if} & \langle U \rangle_{i+1} > \langle U \rangle_i \end{array} \right. \tag{25}$$

We set $p^j = 10$ and $\alpha^j = 0.9$ for all blocks. Experiments have shown that these parameters work well whatever the observations.

The initial temperature factor T_0^j is always a crucial point of simulated annealing algorithms. Indeed, a too low temperature make the simulated annealing procedure useless because it becomes nearly equivalent to the non-tempered algorithm. On the contrary, a too high temperature will result in a long burn-in period and may destroy all initialization effort. Thus, we propose an automatic method to determine a good initial temperature. To do so, we assume that the temperature depends on three features:

- In a Gaussian context, increasing the temperature is nearly equivalent to increasing noise variance, so we propose to first use the temperature to regularize the noise variance in the observation. This is done by dividing the initial temperature by the variance of the noise.

- Each block of the Gibbs algorithm does not affect the same number of data. Some variables like Sérsic parameters affect pixels in only one band whereas shape parameters have an influence on all pixels. One can note that because all parameters are optimized in a global way and are linked together (not necessarily directly), this is not completely true. However we can assume that the more pixels are affected by a parameter, the more difficult it is to optimize it and thus, we will set the temperature to be proportional to the number of affected pixels.

- Finally we assume that we can trust the initial parameters $\boldsymbol{\phi}_0$ in a certain limit, by saying that if the initial parameters are good (in term of MNSE) then, we should not put the initial temperature too high as we are certainly already close to the global maximum. If the initial guess is poor, we consider that it is safer to use a warmer temperature.

These assumptions suggest the following formula for the initial temperature:

$$T_0^j = \beta^j \frac{\sqrt{NSE^j(\boldsymbol{\phi}_0)}}{\text{mean}(\text{diag}(\boldsymbol{\Sigma}^j))} \tag{26}$$

where NSE^j is the normalized squared error of the pixels affected by block j and β^j is a scaling factor for block j. The term $NSE^j(\boldsymbol{\phi}_0)$ incorporates information about the quality of $\boldsymbol{\phi}_0$ and the number of affected pixels. The term $\text{mean}(\text{diag}(\boldsymbol{\Sigma}^j))$ performs the noise regularization. $\boldsymbol{\Sigma}^j$ denotes

the submatrix of $\mathbf{\Sigma}$ which concerns pixels affected by block j. One can note that taking the mean of the diagonal elements of the matrix was chosen because the noise power are nearly equivalent in all bands but may not be the best choice for all applications. The β^j factor must be determined empirically and refers to the intrinsic complexity related to this block. Adopted values are given in table 2.

5.2. Hierarchy of models

In the previous section, we suggested that a good initialization is important to start the tempering from lower temperatures. However, deriving the initialization from observations is a big issue. We can easily extract the shape properties of the dominant structure (the disc in spiral galaxies) by measuring the second order moments but other initial estimates are harder to get. Instead, we propose a new way to perform the optimization in several steps using a hierarchy of models. The idea is to estimate the parameters of a coarse model and use them to initialize the estimation algorithm of a new finer model. The final and finest model of the hierarchy is then the one presented in section 2. The hierarchy is determined by the relative importance of the different structures, the models becoming more and more complex with finer structures along the hierarchy.

In our strategy, we define three models: the first one is only composed of a disc and a bulge, the second one introduces the stellar bar, and the final one incorporates the spiral structure (Fig. 8). This hierarchy was determined by the relative importance of each structure in terms of flux and effects they can have on other structures. For example, it has been shown that the spiral arms have a relatively low influence on the disc and the bulge parameters estimation [44] whereas the stellar bar may have a big impact on the bulge parameters estimation [12].

This strategy is also based on the experience that it requires less simulation effort to estimate all the coarse models than running the burn-in period of the full model. Several factors have a good impact on the global computational effort:

- A coarse model have less blocks to estimate than a fine model. Thus, the simulation effort is lower, as well as the computation time.

- All the structures of a fine model that were also present in a coarse model can be accurately initialized using the estimates of the coarse model. This allows to start the tempering from a colder temperature and to reach the convergence time quicker.

- The estimates of a coarse model can also help to provide a good initialization of structures introduced in the finer model.

5.3. Deterministic initialization

There are three deterministic initialization procedures corresponding to the three models (see section 5.2). The first one aims at finding initial estimation for the bulge and the disc parameters from the global observation. We choose to use second order statistics to determine the shape parameters of both the bulge and the disc [24]. The second order moments give us an estimation of the centre, the elongation, the position angle and the scale length. The effective radius of the disc is set equal to the measured scale length and the bulge effective radius is initialized to half the value of the disc one. The Sérsic index of the bulge is always initialized to 4 (de Vaucouleurs law). The brightness factors of the bulge and the disc are defined such that the disc contains three quarters of the total flux and the bulge the remaining quarter.

In a second time, the parameters of the bulge and the disc are initialized using the estimates of the first model. Nevertheless, we need to determine initial parameters for the stellar bar. We have automatized and generalized a procedure imagined for face-on spiral in [45]. The overall idea of the method is to examine the evolution of the isophotes from the centre to the outer part of the galaxy and to look for the bar signature in the evolution of the elongation and the position angle of these isophotes (i.e. a strong elongation with a constant position angle). To fulfill the

Table 2: Summary of all the variables, their initial temperature factors β, how they are initialized and how they are sampled. RWHM means Random Walk Hastings Metropolis and AD means Adaptive Direction. All proposal distributions are Gaussians whose covariance matrix (or deviation for scalars) are given in the last column. Initialization methods are described in section 5.3.

Structure	Parameter	β Factor	Initialization	Estimation Algorithm
$\mathfrak{B}, \mathfrak{B}a, \mathfrak{D}$	c_x, c_y	0.01	First order moment	RWHM $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$\mathfrak{B}$	e	0.2	2nd order moment	RWHM 0.2
	α	0.01	2nd order moment	RWHM 0.2
	c	0.01	0	RWHM 0.05
	I	No annealing	1/3 total flux	direct simulation
	R	1	2nd order moment	AD $\begin{pmatrix} 0.16 & -0.02 \\ -0.02 & 0.01 \end{pmatrix}$
	n		4	
$\mathfrak{D}$	e	0.01	2nd order moment	RWHM 0.2
	α	0.01	2nd order moment	RWHM 0.2
	I	No annealing	1/3 total flux	direct simulation
	R	0.01	2nd order moment	RWHM 1
$\mathfrak{s}$	R_0	0.01	bar cut off radius	RWHM 1
	θ_0	0.01	isophotes analysis	RWHM 0.2
	$o_{\mathfrak{s}}$	0.01	bar position angle	RWHM 0.1
	R_e	0.01	disc effective radius	RWHM 0.1
	I	No annealing	1/4 disc flux	direct simulation
	R	1	3 arcsec	AD $\begin{pmatrix} 0.16 & -0.02 \\ -0.02 & 0.01 \end{pmatrix}$
	n		0.75	

face-on hypothesis, we first compute a deprojected version of the galaxy using the disc parameters estimated in the first model: this is an example on how information brought by a coarse model help to initialize new components of a fine model. This method provides estimates for the position angle and the cut off radius. The initialization of the remaining parameters is described in table 2.

Finally, the initial parameters of the spiral arms are determined using a Fourier analysis of the residual (difference between the observation and the simulation) in a log polar space. Indeed, in this particular space the spiral arms appear like two equidistant lines whose parameters can be identified in the frequency space [46, 47]. The spiral arms are identified by the maximum of frequency 2. From this maximum we can extract the pitch angle (inclination of the sinusoidal wave) and the starting radius (phase angle interpreted as a translation). The initialization of the remaining parameters is described in table 2.

6. Results

6.1. Comparison With Other Algorithms

As mentioned in section 1, available software perform supervised monoband galaxy decomposition (e.g [15–17]). In order to discuss about our method performances, we have performed a comparison with Galfit [16] which is a well-known and recognized software in the astronomical community. In this case, we used only the bulge and the disc structures in monoband images, so that both methods use exactly the same model. Nevertheless, Galfit performs a deterministic fit with the Levenberg-Marquardt method and requires a manual initialization with a sufficiently good precision. The tests were conducted on five galaxies and showed a good behaviour of the

two methods. We have also compared the MNSE of the estimations and the magnitude[3] of the structures provided by each method. The comparison is detailed in table 3 and shows a good agreement between both methods. Our approach gives slightly better results in term of MNSE than Galfit.

Table 3: Comparison between our method (O) and Galfit (G) for the bulge-disc decomposition of five monoband galaxy images. For each line, the first column gives the PGC number of the galaxy. Then we have: the MNSE, the magnitude of the disc (MD) and the magnitude of the bulge (MB) for each method.

Object	(O)-MNSE-(G)		(O)-MD-(G)		(O)-MB-(G)	
2182	2.53	2.69	12.6	12.55	15.65	15.35
2600	122.64	122.82	11.69	11.60	15.17	15.20
3563	8.48	9.65	13.13	13.13	12.24	12.24
3584	7.25	7.27	12.18	12.18	14.73	14.74
5055	1.56	1.59	14.92	14.95	13.24	13.24

6.2. Simulated images

We used simulated images to evaluate the performances of the algorithm compared to a ground truth. The results show that the proposed method is able to provide a very good solution on raw multispectral data with an unsupervised Bayesian inference scheme.

Figure 9 shows an example on a 5-band simulation of a barred spiral galaxy. The peak signal-to-noise ratio (PSNR) defined as $10\log_{10}(\max(\mathbf{Y})^2/\sigma^2)$ (σ is the standard deviation of the noise in the simulation) for each band is respectively: $37.7\,\mathrm{dB}$, $53.5\,\mathrm{dB}$, $55.6\,\mathrm{dB}$, $54.5\,\mathrm{dB}$ and $45.3\,\mathrm{dB}$. These values are typical for real images in extragalactic astronomy.

The residual image shows the trace of an ellipsoidal structure. This trace is the biggest error on the estimation and it comes from a small error in the estimates of the stellar bar shape which is confirmed by a detailed analysis of the parameters. In fact, there is a negative correlation between the stellar bar cut-off radius and the spiral starting radius, and the algorithm had some difficulties to get out of a local minimum which is very close to the optimal one. Nevertheless the error remains negligible compared to the total flux of each structures and the MNSE of the estimates is only 1.099. The maximal absolute error per pixel divided by noise deviation is respectively for each band 1.13, 2.6, 3.1, 2.5 and 1.6.

Figure 10 is a plot of the estimates of all parameters with their error bars (three times the standard deviation) against their true value. The abscissa axis represents the value of the ground truth values, whereas the ordinate axis are the estimates. Ideally, every point should be on the diagonal. A point located below the diagonal means that the parameter is under-estimated and vice versa. We can see that nearly all the parameters are well estimated. Nevertheless, as usual, most of the errors are under-estimated.

A more global analysis of the results leads to several observations. The variables like the elongation, the position angle, or the centre are retrieved with a very good precision (relative error less than 0.1%) and, thanks to the spatial extension of the disc, its radius is well estimated with a relative error less than 5%. On the other hand, the Sérsic parameters are harder to estimate. For the bulge and the stellar bar, the information is hold by only a few pixels and an error in the estimation does not lead to a high variation in the likelihood. We also observed that the different errors seem to compensate each other as they nearly do not affect global features like the integrate flux.

---

[3]The magnitude is equal to $-2.5\log_{10} Fl + Fl_0$ with Fl the integrate flux of the structure and $Fl_0 = 26$ is a reference flux corresponding to the brightness of the dark sky.
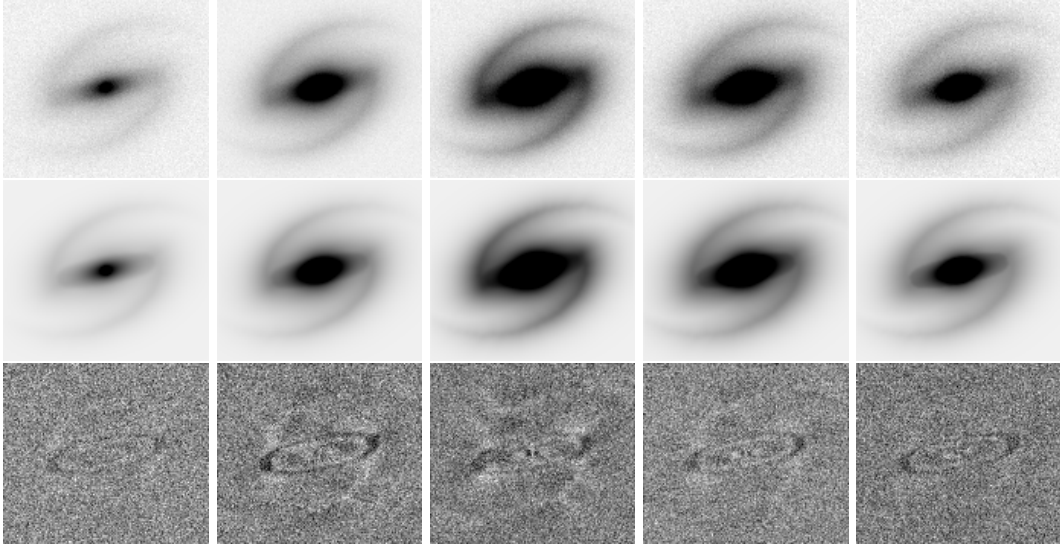
Figure 9: Application of the method on a 5-band simulated image. The first row shows the five bands of the simulation, the second row is the estimate and the last row is the residual image. All the images are given in inverse greyscale. The simulations and the estimates are all displayed with the same dynamic range, while the contrast of all residual images has been adjusted independently to emphasize the small structures; the dynamic range of the residual images is less than 2% of the dynamic range of the observations.

6.3. Real images

We present tests realized on three real 5-band images of barred spiral galaxies. The images come from the EFIGI database [22] which is based on the Sloan Digital Sky Survey. The five bands were obtained using the photometric system *ugriz* [48] which defines filters centred on the following wavelengths: u=3543 Å, g=4770 Å, r=6231 Å, i=7625 Å and z=9134 Å. Each image is about 200 by 200 pixels. During the preprocessing, other astronomical objects (mainly superimposed stars from the Milky Way) were masked using an automatic procedure. The images were chosen according to the following criteria: they contain a relatively well resolved barred spiral galaxy; the galaxy has a reasonable inclination so the stellar bar and the spiral arms are visible; the galaxy is not merging with another one; the field is not contaminated by another very bright and extended object. We remind that the H II regions are not modelled but masked using the parametric filter and thus they do not appear in the simulations.

Galaxy PGC 2182 (Fig. 11) is a very bright object viewed nearly face-on with a medium size stellar bar. It has a small bulge and its arms are quite bright with a low pitch angle. No H II region is visible. The figures show that the galaxy is well modelled with a final MNSE equals to 1.09. It is remarkable that this MNSE, corresponding to the complete model including the four structures, is 10 times lower than the MNSE of the bulge–disc model which is the first model in the hierarchy (table 4). We do not see any residual structure from the disc and the stellar bar. A small residual is observable in the bulge, it can be due to model noise or an imperfect estimation of the psf. The shape and the brightness of the arms are well retrieved. We can observe that in the real images the arms seem to spread more in the outer part of the galaxy and this effect cannot be reproduced in our model.

Galaxy PGC 23650 (Fig. 12) is a quite bright object with a low inclination. It has a massive and extended stellar bar and a small bulge. It has two faint and loosely wound spiral arms. Again, we do not observe H II region. The four structures are well estimated with a final MNSE equals to 2.06. The addition of the stellar bar and the spiral arms structures to the bulge–disc model improves the decomposition as the MNSE is reduced by 24 % (table 4). Only the cut-off radius of
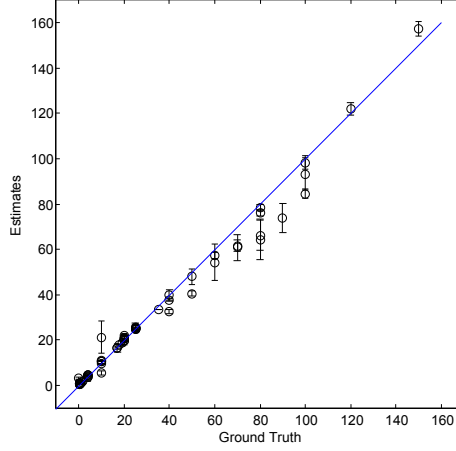
Figure 10: Estimate of each parameter against its true value for the simulated image 9

the stellar bar seems to have been over-estimated. This can be explained by the two over bright regions at the beginning of the arms that are not included in the model.

Galaxy PGC 35538 (Fig. 13) is a galaxy of medium brightness observed nearly face-on. It has a very small bulge and an extended disc. Its stellar bar is very small and confined in its inner part. Its arms are tightly wound and not very well defined. We can observe several H II regions mainly distributed along the spiral arms. We can see that the estimates are good with a final MNSE equals to 2.91. The addition of the stellar bar and the spiral arms structures to the bulge–disc model improves the decomposition as the MNSE is reduced by 20 % (table 4). The bulge, the disc and the stellar bar are very well estimated. Despite the difficulty to visually extract two distinct arms from the observation, the algorithm has found a reasonable solution: the shape of the two arms is coherent with the observation. Moreover, the map of the H II regions and the result of the adaptive masking technique showed in Fig. 14 confirms this remark. The masking technique is of major interest in such case where the spiral arms may be better defined by the spatial distribution of the H II regions than by the brightness of its continuous counterpart.

Figure 15 shows the brightness profiles of the three real galaxies and their estimates. We can see that, despite the complexity of the different curves, an accurate fit can be automatically realized.

6.4. Computation Time

Table 4 gives the computation statistics for the simulated galaxy (section 6.2) and the three real galaxies (section 6.3). The tests were run on an Intel Core 2 processor (2.66 GHz) and the code is written in Java. The total time to process the three steps of the hierarchy of models is comprised between 2 and 3 hours and requires 600 to 700 iterations. About 15% of this time is used to estimate the first model and 24% for the second model.

Table 4: Computation statistics. For each galaxy, and for each step of the hierarchy of models, we give the number of iteration to converge (It.), the total time spend (in minutes) and the MNSE of the estimates.

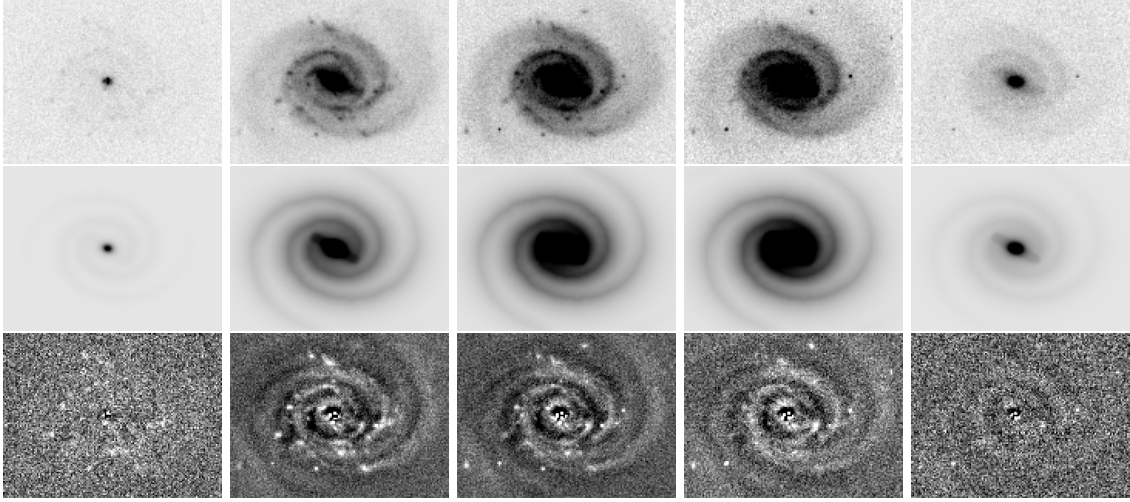| Galaxy	Simulation			PGC2182			PGC23650			PGC35538		
Model	It.	Time	MNSE	It.	Time	MNSE	It.	Time	MNSE	It.	Time	MNSE
$\mathfrak{B}+\mathfrak{D}$	165	14	10.73	156	12	3.41	142	10	3.52	140	11	3.64
$\mathfrak{B}+\mathfrak{D}+\mathfrak{Ba}$	187	23	7.72	182	21	3.25	190	18	2.86	205	20	3.48
$\mathfrak{B}+\mathfrak{D}+\mathfrak{Ba}+\mathfrak{s}$	382	145	1.09	293	101	2.6	251	98	2.68	303	137	2.91
Total	734	182		631	133		583	126		648	168	

Figure 11: Application of the method on a 5-band (*ugriz*) image of the galaxy PGC 2182. The first row shows the five bands of the observation, the second row is the estimate and the last row is the residual image. All the images are given in inverse greyscale. The simulations and the estimates are all displayed with the same dynamic range. The contrast of all residual images has been adjusted independently to emphasize the small structures.

The total number of iterations per object is low for our type of application. It is difficult to appreciate the gain due to the tempering strategy because it would have been very difficult to obtain good results without it. We have performed tests without the hierarchical decomposition strategy and we observed that a thousand of burn-in iterations were needed, leading to a much more longer computation time.

One can note that, this application could take a great advantage of GPU capabilities for a massive parallelization of the likelihood computation which is the most time consuming part of the algorithm.

7. Conclusion

This paper deals with the decomposition of galaxies in multispectral observations. An MCMC algorithm allows to estimate the model parameters in a fully automatic and efficient way. The method has been validated on simulated and real images showing its ability to recover accurate and meaningful estimates.

Our first contribution was to extend an existing parametric galaxy model to handle multispectral observations of barred spiral galaxies. Especially, we have proposed a new multispectral parametric model for the spiral arms.

Second, we proposed a new tempering strategy in which each block of the Gibbs sampler has a different temperature. This reflects the difficulty to explore the conditional posterior distribution of the parameters in the dimension of the block. In this scheme, the simulation effort is better distributed than in traditional simulated annealing. Moreover, we have proposed a way to automatically determine an initial temperature for each block of the sampler.

Third, we have also completed the traditional parametric model with a parametric filtering of the observation. In this approach, the parameters of the filter are fully part of the model and they are estimated as the other parameters. In our case, the parametric filter is used to remove non desirable aspects (the H II regions) of the observation that we do not want to model. Nevertheless, these H II regions carry interesting spatial information and our approach preserves this spatial information contrary to a preprocessing.
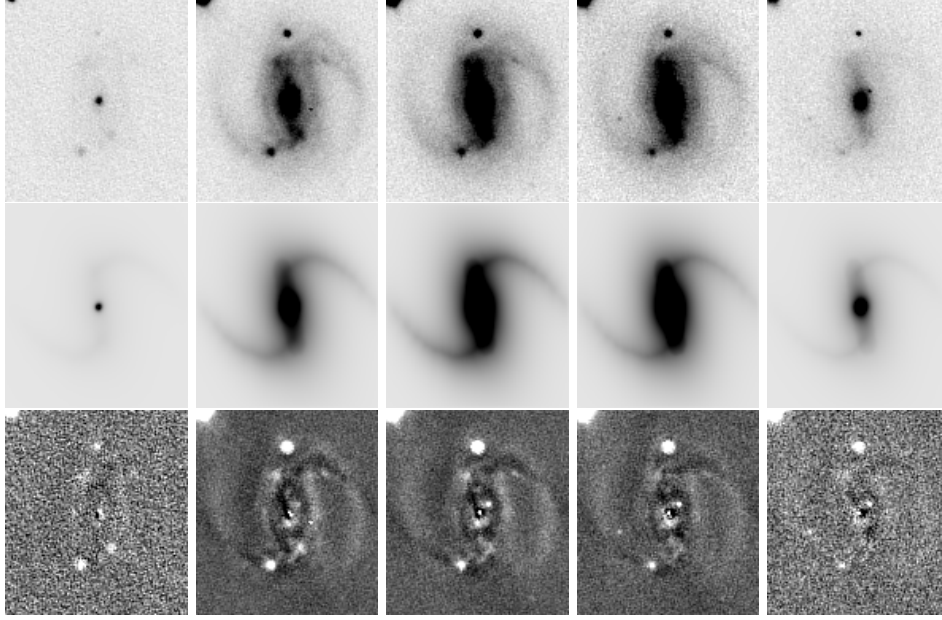
Figure 12: Application of the method on a 5-band (*ugriz*) image of the galaxy PGC 23650. See the caption of Fig. 11 for explanations. The three bright stars in the corner, bellow and above the galaxy were masked.

Our last contribution concerns the initialization procedure. We introduced a hierarchy of models, starting from a coarse model and finishing by the finest and complete model of the barred spiral galaxy. In this framework, the estimates of a coarse model are used to initialize a finer model, reducing the burn-in period of the fine model.

There are several perspectives to our work. From a methodological point of view, we think that the parametric filtering technique deserves an in-depth study as it allows to replace a preprocessing by a robust adaptive method. From an astronomical point of view, we plan to enrich the hierarchy of models with a model per galaxy type. Our idea is to define decision rules to navigate through the possible models of galaxies and to end up with the model that better fits with the multispectral observations. This last point is of great importance in astronomy, where millions of galaxies are observed by even more accurate telescopes.

AppendixA. Distribution of a Brightness Factor

Each structure of the model owns a brightness factor I_b in each band b. We can show that the conditional posterior of this factor is distributed according to a truncated positive Gaussian. Let $\mathbf{f}$ be the brightness of a structure and I_b the brightness factor of $\mathbf{f}$ in band b. We define $\mathbf{f}' = \mathbf{f}/I_b$. Let $\mathbf{z} = \mathbf{A}_\mathrm{m}\left(\mathbf{Y}, \boldsymbol{\phi}\right) - \mathbf{H} * \mathbf{F}_{\setminus\{\mathbf{f}\}}(\boldsymbol{\phi})$ and $\mathbf{g} = \mathbf{H} * \mathbf{f}'(\boldsymbol{\phi})$ where $\mathbf{F}_{\setminus\{\mathbf{f}\}}$ denotes the sum of the brightness of all structures excepted $\mathbf{f}$. Equation (15) can be rewritten as:

$$
\begin{aligned}
P(I_b|\mathbf{Y},\boldsymbol{\phi}_{\setminus I_b}) &\propto \frac{1}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left(-\tfrac{1}{2}(\mathbf{z} - I_b\mathbf{g})^T \boldsymbol{\Sigma}^{-1}(\mathbf{z} - I_b\mathbf{g})\right)\, \mathbb{I}_{\mathbb{R}+}(I_b) \\
&\propto \frac{1}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left(-\tfrac{1}{2}(\mathbf{z}^T\boldsymbol{\Sigma}^{-1}\mathbf{z} + I_b(-\mathbf{z}^T\boldsymbol{\Sigma}^{-1}\mathbf{g} - \mathbf{g}^T\boldsymbol{\Sigma}^{-1}\mathbf{z}) + I_b^2\mathbf{g}^T\boldsymbol{\Sigma}^{-1}\mathbf{g})\right)\, \mathbb{I}_{\mathbb{R}+}(I_b) \\
&\propto \frac{1}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left(-\tfrac{1}{2}(I_b^2 A + I_b B + C)\right)\, \mathbb{I}_{\mathbb{R}+}(I_b)
\end{aligned}
\tag{A.1}
$$

with

$$
A = \mathbf{g}^T\boldsymbol{\Sigma}^{-1}\mathbf{g}\ , \qquad B = -\mathbf{z}^T\boldsymbol{\Sigma}^{-1}\mathbf{g} - \mathbf{g}^T\boldsymbol{\Sigma}^{-1}\mathbf{z}\ , \ and \qquad C = \mathbf{z}^T\boldsymbol{\Sigma}^{-1}\mathbf{z}
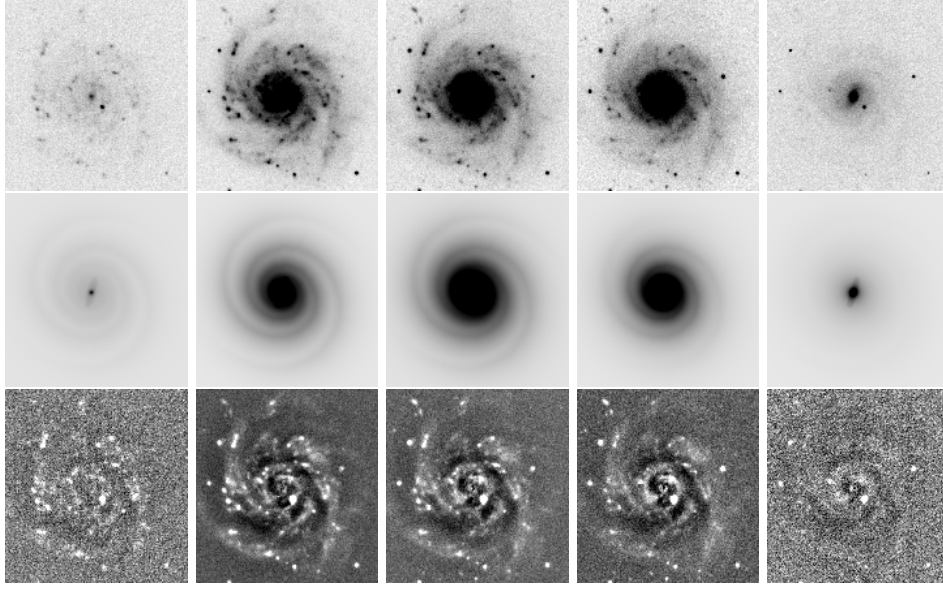\tag{A.2}
$$

Figure 13: Application of the method on a 5-band (*ugriz*) image of the galaxy PGC 35538. See the caption of Fig. 11 for explanations.

so:

$$
\begin{aligned}
I_b^2 A + I_b B + C &= I_b^2 A - 2\left(-\tfrac{1}{2}\tfrac{B}{A}\right) I_b A + \left(-\tfrac{1}{2}\tfrac{B}{A}\right)^2 A + C - \left(-\tfrac{1}{2}\tfrac{B}{A}\right)^2 A \\
&= A\left[I_b - \left(-\tfrac{1}{2}\tfrac{B}{A}\right) \right]^2 + C - \left(-\tfrac{1}{2}\tfrac{B}{A}\right)^2 A
\end{aligned}
\tag{A.3}
$$

let $\mu = -\tfrac{1}{2}\tfrac{B}{A}$ and $\epsilon = C - \mu^2 A$, (A.1) becomes:

$$
\begin{aligned}
P(I_b | \mathbf{Y}, \boldsymbol{\phi}_{\setminus I_b}) &\propto \frac{1}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left(-\tfrac{1}{2}(A(I_b - \mu)^2 + \epsilon)\right) \mathbb{I}_{\mathbb{R}^+}(I_b) \\
&\propto \left(\frac{D}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}} \right) \frac{1}{D} \exp\left(-\tfrac{1}{2}\left(\frac{I_b - \mu}{1/\sqrt{A}}\right)^2\right) \exp(\epsilon) \; \mathbb{I}_{\mathbb{R}^+}(I_b)
\end{aligned}
\tag{A.4}
$$

with

$$
D = \sqrt{\frac{\pi}{2A}} \left(1 + \mathrm{erf}\left(\frac{\mu}{\sqrt{2}\frac{1}{\sqrt{A}}} \right) \right)
\tag{A.5}
$$

where erf is the Gauss error function. We recognize the expression of a truncated Gaussian law, thus:

$$
\begin{aligned}
P(I_b | \mathbf{Y}, \boldsymbol{\phi}_{\setminus I_b}) &\propto \frac{1}{D} \exp\left(-\tfrac{1}{2}\left(\frac{I_b - \mu}{1/\sqrt{A}}\right)^2\right) \mathbb{I}_{\mathbb{R}^+}(I_b) \\
&\sim \mathcal{N}^+(\mu, 1/A)
\end{aligned}
\tag{A.6}
$$

where $\mathcal{N}^+(\mu, 1/A)$ is a truncated positive Gaussian pdf of mean μ and variance $1/A$ which can be efficiently simulated [31].

AppendixB. Integration of a Brightness Factor

This Appendix uses the same notations as Appendix AppendixA. Because the brightness factor I_b is distributed according to a truncated positive Gaussian distribution (Appendix AppendixA),
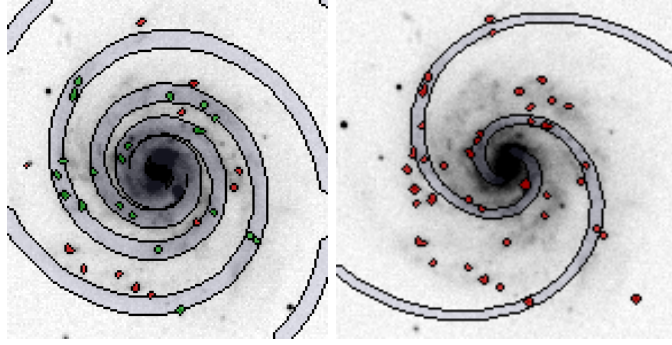
Figure 14: The results with (left) and without (right) adaptive masking technique on the galaxy PGC 35538 (Fig. 13). The red and green spots are all possible H II regions, the green ones are masked, the red ones were omitted because they were too far from the spiral crest (grey-blue area). Without adaptive masking (right), the estimation is bad because the spiral arms do not pass through a majority of H II regions as it should do.
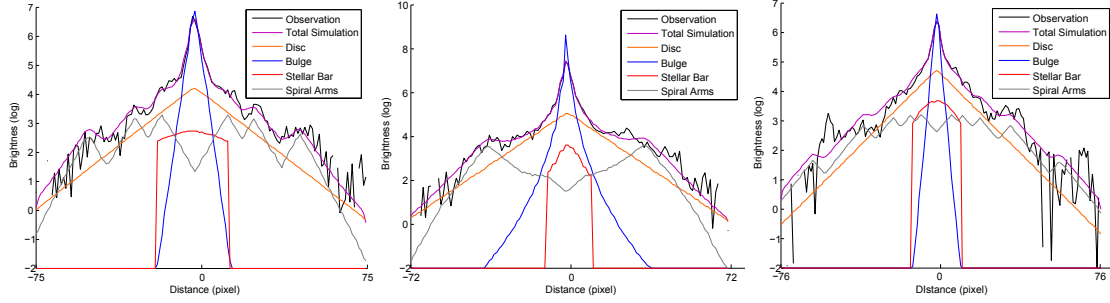

Figure 15: Profile cuts of the observation and the estimate of galaxies (from left to right) PGC 2182 (Fig. 11), PGC 23650 (Fig. 12), and PGC 35538 (Fig. 13) in band g. For the galaxies PGC 2182 and PGC 35538, the profiles passes through the stellar bar major axis. For the galaxy PGC 23650, the cut passes through the diagonal line from bottom left corner to upper right one. Each plot shows a profile of the observation (black), the complete simulation of the estimate with the psf (purple), and the different structures of the simulation without the psf: the disc (orange), the bulge (blue), the stellar bar (red), the spiral arms (grey). The intensities are given in a log scale.

we can integrate it from the posterior pdf. We have:

$$
\begin{aligned}
P(\boldsymbol{\phi}_{\setminus I_b}|\mathbf{Y}) &\propto \int_0^\infty P(I_b, \boldsymbol{\phi}_{\setminus I_b}|\mathbf{Y})\, dI_b \\
&\propto \int_0^\infty P(I_b|\mathbf{Y}, \boldsymbol{\phi}_{\setminus I_b}) P(\boldsymbol{\phi}_{\setminus I_b})\, dI_b \\
&\propto \int_0^\infty \left(\frac{D}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}}\right) \exp(\epsilon) \frac{1}{D} \exp\left(-\frac{1}{2}\left(\frac{I_b - \mu}{1/\sqrt{A}}\right)^2\right) \mathbb{I}_{\mathbb{R}^+}(I_b) P(\boldsymbol{\phi}_{\setminus I_b})\, dI_b \\
&\propto \left(\frac{D}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}}\right) \exp(\epsilon)\, P(\boldsymbol{\phi}_{\setminus I_b}) \int_0^\infty \frac{1}{D} \exp\left(-\frac{1}{2}\left(\frac{I_b - \mu}{1/\sqrt{A}}\right)^2\right) \mathbb{I}_{\mathbb{R}^+}(I_b)\, dI_b \\
&\propto \left(\frac{D}{(2\pi)^{\frac{N}{2}} \det(\boldsymbol{\Sigma})^{\frac{1}{2}}}\right) \exp(\epsilon)\, P(\boldsymbol{\phi}_{\setminus I_b})
\end{aligned}
\tag{B.1}
$$

Acknowledgment

[1] J. Møller, R. P. Waagepetersen, Statistical Inference and Simulation for Spatial Point Process, Chapman & Hall/CRC, 2004.

[2] F. Lafarge, X. Descombes, J. Zerubia, M. Pierrot-Deseilligny, Structural approach for building reconstruction from a single DSM, IEEE Trans. Pattern Analysis and Machine Intelligence 32 (1) (2010) 135–147.

[3] K. Sun, N. Sang, T. Zhang, Marked point process for vascular tree extraction on angiogram, in: LNCS: Energy Minimization Methods in Computer Vision and Pattern Recognition, Berlin, 2007, pp. 467–478.

[4] S. Descamps, X. Descombes, A. Béchet, J. Zerubia, Automatic flamingo detection using a multiple birth and death process, in: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, USA, 2008, pp. 1113–1116.

[5] G. Perrin, X. Descombes, J. Zerubia, A non-bayesian model for tree crown extraction using marked point processes, Research Report 5846, INRIA, France (February 2006).

[6] C. Collet, J. Chanussot, K. Chehdi (Eds.), Multivariate Image Processing, John Wiley and Sons, 2009.

[7] A. Djafari (Ed.), Inverse Problems in Imaging and Vision, John Wiley and Sons, 2010.

[8] C. Andrieu, J. Thoms, A tutorial on adaptive MCMC, Statistics and Computing 18 (4) (2008) 343–373.

[9] W. Gilks, G. Roberts, S. Sahu, Adaptive Markov chain Monte Carlo through regeneration, Journal of the American Statistical Association 93 (443) (1998) 1045–1054.

[10] W. Gilks, S. Richardson, D. Spiegelhalter (Eds.), Markov Chain Monte Carlo In Practice, Chapman & Hall/CRC, Washington, D.C., 1996.

[11] E. Cameron, S. P. Driver, A. W. Graham, J. Liske, The Millennium Galaxy Catalogue: Exploring the Color-Concentration Bimodality via Bulge-Disk Decomposition, Astrophysical Journal 699 (2009) 105–117.

[12] D. A. Gadotti, Image decomposition of barred galaxies and AGN hosts, Monthly Notices of the Royal Astronomical Society 384 (2008) 420–439.

[13] C. M. Gutiérrez, I. Trujillo, J. A. L. Aguerri, A. W. Graham, N. Caon, Quantitative Morphology of Galaxies in the Core of the Coma Cluster, Astrophysical Journal 602 (2004) 664–677.

[14] I. Trujillo, J. A. L. Aguerri, Quantitative morphological analysis of the Hubble Deep Field North and Hubble Deep Field South - I. Early- and late-type luminosity-size relations of galaxies out to z~ 1, Monthly Notices of the Royal Astronomical Society 355 (2004) 82–96.

[15] L. Simard, C. N. A. Willmer, N. P. Vogt, V. L. Sarajedini, A. C. Phillips, B. J. Weiner, D. C. Koo, M. Im, G. D. Illingworth, S. M. Faber, The DEEP Groth Strip Survey. II. Hubble Space Telescope Structural Parameters of Galaxies in the Groth Strip, Astrophysical Journal, Supplement 142 (2002) 1–33.

[16] C. Y. Peng, L. C. Ho, C. D. Impey, H. Rix, Detailed Structural Decomposition of Galaxy Images, Astronomical Journal 124 (2002) 266–293.

[17] R. E. de Souza, D. A. Gadotti, S. dos Anjos, BUDDA: A New Two-dimensional Bulge/Disk Decomposition Code for Detailed Structural Analysis of Galaxies, Astrophysical Journal, Supplement 153 (2004) 411–427.

[18] C. Collet, F. Murtagh, Multiband segmentation based on a hierarchical markov model, Pattern Recognition 37 (12) (2004) 2337 – 2347.

[19] E. P. Hubble, Extragalactic nebulae., Astrophysical Journal 64 (1926) 321–369.

[20] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.

[21] C. Robert, G. Casella, Monte Carlo Statistical Methods, 2nd Edition, Springer-Verlag, 2004.

[22] A. Baillard, E. Bertin, Y. Mellier, H. J. McCracken, T. Géraud, R. Pelló, F. Leborgne, P. Fouqué, Project EFIGI: Automatic Classification of Galaxies, in: C. Gabriel, C. Arviset, D. Ponz, S. Enrique (Eds.), Astronomical Data Analysis Software and Systems XV, Vol. 351 of Astronomical Society of the Pacific Conference Series, 2006, pp. 236–240.

[23] I. S. McLean, Electronic Imaging in Astronomy : Detectors and Instrumentation, 2nd Edition, Springer, Praxis Publishing, Chichester, UK, 2008.

[24] E. Bertin, S. Arnouts, SExtractor: Software for source extraction., Astronomy and Astrophysics, Supplement 117 (1996) 393–404.

[25] J. L. Sérsic, Atlas de galaxias australes, Cordoba, Argentina : Observatorio Astronomico, 1968.

[26] G. de Vaucouleurs, Classification and Morphology of External Galaxies., Handbuch der Physik 53 (1959) 275–310.

[27] E. Athanassoula, The spiral structure of galaxies., Physics Reports 114 (1984) 319–403.

[28] R. Strand, Distance functions and image processing on point-lattices : with focus on the 3d face- and body-centered cubic grids, Ph.D. thesis, Uppsala University, Centre for Image Analysis (2008).

[29] A. W. Graham, An Investigation into the Prominence of Spiral Galaxy Bulges, Astronomical Journal 121 (2001) 820–840.

[30] P. D. Allen, S. P. Driver, A. W. Graham, E. Cameron, J. Liske, R. de Propris, The Millennium Galaxy Catalogue: bulge-disc decomposition of 10095 nearby galaxies, Monthly Notices of the Royal Astronomical Society 371 (2006) 2–18.

[31] V. Mazet, D. Brie, J. Idier, Simulation of positive normal variables using several proposal distributions, Statistical Signal Processing, 2005 IEEE/SSP 13th Workshop (2005) 37–42.

[32] S. Geman, D. Geman, Stochastic relaxation, gibbs distributions and the bayesian restoration of images, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6) (1984) 721–741.

[33] L. Devroye, Non-Uniform Random Variate Generation, Springer-Verlag, New York, 1986.

[34] S. Chib, E. Greenberg, Understanding the Metropolis-Hastings algorithm, The American Statistician 49 (4) (1995) 327–335.

[35] A. Gelman, G. Roberts, W. Gilks, Efficient Metropolis jumping rules, in: Bayesian Statistics, Vol. 5, Oxford University Press, 1996, pp. 599–607.

[36] S. Brooks, G. Roberts, Assessing convergence of markov chain monte carlo algorithms, Statistics and Computing 8 (1997) 319–335.

[37] J. Flegal, M. Haran, G. Jones, Markov chain Monte Carlo: Can we trust the third significant figure?, Statistical Science 23 (2) (2009) 250–260.

[38] G. Jones, M. Haran, B. Caffo, R. Neath, Fixed-width output analysis for markov chain monte carlo, Journal of the American Statistical Association 101 (2006) 1537–1547.

[39] J. Flegal, Monte Carlo standard errors in Markov chain Monte Carlo, Ph.D. thesis, University of Minnesota (2008).

[40] G. Roberts, J. Rosenthal, Coupling and ergodicity of adaptive MCMC, Tech. rep., Department of Mathematics and Statistics, Lancaster University (March 2007).

[41] J. Flegal, G. Jones, Batch means and spectral variance estimators in Markov chain Monte Carlo, Annals of Statistics to appear.

[42] C. Geyer, E. Thompson, Annealing Markov chain Monte Carlo with applications to ancestral inference, Journal of the American Statistical Association 90 (431) (1995) 909–920.

[43] C. Lacoste, X. Descombes, J. Zerubia, N. Baghdadi, Hydrographic network extraction from radar satellite imagesusing a hierarchical model within a stochastic geometry framework, Research Report 5697, INRIA, France (September 2005).

[44] A. Misiriotis, N. D. Kylafis, J. Papamastorakis, E. M. Xilouris, Is the exponential distribution a good approximation of dusty galactic disks?, Astronomy and Astrophysics 353 (2000) 117–123.

[45] M. W. Regan, D. M. Elmegreen, K-Band observations of barred spiral galaxies., Astronomical Journal 114 (1997) 965–975.

[46] I. Puerari, H. A. Dottori, Fourier analysis of structure in spiral galaxies, Astronomy and Astrophysics, Supplement 93 (1992) 469–493.

[47] C. Garcia Gomez, E. Athanassoula, Analysis of the distribution of HII regions in external galaxies. II - Analysis of the spiral structure, Astronomy and Astrophysics, Supplement 100 (1993) 431–464.

[48] M. Fukugita, T. Ichikawa, J. E. Gunn, M. Doi, K. Shimasaku, D. P. Schneider, The Sloan Digital Sky Survey Photometric System, Astronomical Journal 111 (1996) 1748–1756.