A FULLY BAYESIAN APPROACH FOR INFERRING PHYSICAL PROPERTIES WITH CREDIBILITY INTERVALS FROM NOISY ASTRONOMICAL DATA

Maxime Vono[1]*, Emeric Bron*[2]*, Pierre Chainais*[3]*, Franck Le Petit*[2]*, Sébastien Bardeau*[4]*, Sébastien Bourguignon*[5]*, Jocelyn Chanussot*[6]*, Mathilde Gaudel* [2]*, Maryvonne Gerin*[2]*, Javier R. Goicoechea*[7]*, Pierre Gratier*[8]*, Viviana V. Guzmán*[9]*, Annie Hughes*[10]*, Jouni Kainulainen*[11]*, David Languignon*[2]*, Jacques Le Bourlot*[2]*, François Levrier*[12]*, Harvey S. Listz*[13]*, Karin I. Oberg*[14]*, Jan H. Orkisz*[11]*, Nicolas Peretto*[15]*, Jérôme Pety*[4]*, Antoine Roueff*[16]*, Évelyne Roueff*[2]*, Albrecht Sievers*[4]*, Victor de Souza Magalhaes*[4] *and Pascal Tremblin*[17]

[1] IRIT, [2] LERMA, [3] Centrale Lille/CRIStAL, [4] IRAM, [5] Centrale Nantes, [6] GIPSA-Lab, [7] CSIC, [8] LAB, [9] UCatholica, [10] IRAP, [11] Chalmers, [12] ENS, [13] NRAO, [14] CfA Harvard, [15] Cardiff University, [16] Institut Fresnel, [17] Maison de la Simulation, CNRS

ABSTRACT

The atoms and molecules of interstellar clouds emit photons when passing from an excited state to a lower energy state. The resulting emission lines can be detected by telescopes in the different wavelength domains (radio, infrared, visible, UV...). Through the excitation and chemical conditions they reveal, these lines provide key constraints on the local physical conditions reigning in giant molecular clouds (GMCs), which constitute the birthplace of stars in galaxies. Inferring these physical conditions from observed maps of GMCs using complex astrophysical models of these regions remains a complicated challenge due to potentially degenerate solutions and widely varying signal-to-noise ratios over the map. We propose a Bayesian framework to infer the probability distributions associated to each of these physical parameters, taking a spatial smoothness prior into account to tackle the challenge of low signal-to-noise ratio regions of the observed maps. A numerical astrophysical model of the cloud is involved in the likelihood within an approximate Bayesian computation (ABC) method. This enables to both infer pointwise estimators (e.g., minimum mean square or maximum a posteriori) and quantify the uncertainty associated to the estimation process. The benefits of the proposed approach are illustrated based on noisy synthetic observation maps.

Index Terms— Approximate Bayesian computation, Markov chain Monte Carlo, physical conditions, radioastronomy

The first four authors are the main contributors of this paper. The remaining authors are listed in alphabetic order.

1. INTRODUCTION

The interstellar medium (ISM), filling the space between the stars of our Galaxy, is a mixture of gas and microscopic dust grains. Roughly half of the ISM mass is concentrated in localized overdense regions (occupying only 1-2% of the volume) called *interstellar clouds* (see [1, 2] for a general introduction to the ISM). The most massive of these clouds form complex filamentary structures called giant molecular clouds, in which the gravitational collapse of the densest core lead to the formation of new stars [3].

These clouds are dense enough to be opaque to UV photons, and the gas in their inner part is thus cold and mostly neutral, allowing the formation of molecules and the development of a rich chemistry. The atoms and molecules of these clouds are excited through thermal collisions, pumping by stellar UV photons or IR photons from thermal dust emission, or chemical formation pumping, and then deexcite by emitting photons at specific wavelength characteristic of the chemical species and their excitation levels. These emission lines can then be observed by telescopes in the different wavelength domain (radio domain for rotational transitions, IR domain mostly for rovibrational transitions and visible/UV domain mostly for electronic transitions).

The study of the star formation process in GMCs and its retroaction on the parental cloud thus rests in large part on the observation of molecular and atomic emission lines. The information these lines provide on the chemical composition of the gaz and on the excitation state of the chemical species can be exploited using complex astrophysical models of these regions to infer the physical conditions (gas density, temperature, pressure, ionization degree,...) reigning in star forming regions, providing us with a unique window into the conditions of star formation.

The development of instruments with increasing spectro-imaging capacities (spectral bandwidth and resolution, field of view and angular resolution) now allows studies of the spatial distribution of these physical conditions across star forming regions. However, observation maps that now cover areas not limited anymore to the few brightest positions in the cloud require dealing with lower signal-to-noise ratios over large areas of the maps. For instance, the wide field radio observations of the Orion B GMC obtained by [4], covering a dynamical range of spatial scales of more than two orders of magnitude and the full spectral band between 84.5 to 116.5 GHz at a spectral resolution of $\sim$ 200 kHz, prefigures the future of astronomical datasets.

Recently, the estimation of such physical condition maps from infrared observations (with the Hershel space telescope) of the Carina Nebula, an active star forming region where the parental cloud is submitted to strong radiative feedback from young massive star clusters such as Trumpler 14, has been undertaken in [5] using the Meudon PDR model [6]. This study used a maximum likelihood estimation (MLE) to constrain the thermal pressure of the gas and the incident UV flux, and concluded to a strong correlation between these parameters across the map, indicative of a dynamical impact of radiative feedback. However, ad hoc penalties and constraints were needed to avoid unphysical solutions in low signal-to-noise regions. Such an approach presents several shortcomings. Ad hoc constraints need to be designed by trial and error and are not easily transposable to other datasets, and the justification of the constraints used remain unclear. Moreover, even if the aforementioned authors used Monte Carlo experiments to quantify the uncertainty for the physical conditions, the use of a consistent Bayesian framework [7] has not been considered.

In this paper, we propose to extend the approach of [5] by formulating the inference problem in the Bayesian paradigm. By deriving a Bayesian hierarchical model and an associated Markov chain Monte Carlo (MCMC) algorithm, this permits to take into account both model and measurement errors, to sample possible hyperparameters instead of hand-tuning them and to derive pointwise estimators while quantifying the uncertainty for the unknown physical conditions. This quantification of uncertainties is essential to guarantee the credibility of resulting estimates. We include in this framework a spatial smoothness prior based on the idea that the structure of GMCs is such that lower density regions, less bright and thus observed at lower signal-to-noise ratio but covering a larger fraction of the cloud area, have spatial structures of larger typical scales than denser and brighter regions.

To this purpose, Section 2 presents the proposed Bayesian hierarchical model and Section 3 derives the MCMC algorithm to sample from the target posterior distribution. Section 4 illustrates the benefits of the proposed approach on a synthetic map presenting some typical regions of ISCs. Finally, Section 5 draws some final remarks and possible extensions of this work.

2. BAYESIAN APPROACH

2.1. Problem statement

We assume that we observe ℓ molecular spectral lines, where $L \leq 200$ typically, associated to a spatial map of size $N \times N$. These spectral lines are the signature of the molecules which are present within the ISC. We denote the observed intensity map associated to each line $\ell \in [L]$ by $\mathbf{y}_\ell \in \mathbb{R}_+^{N^2}$, where $[L] = \{1, \ldots, L\}$. Based on the latter, we are interested in inferring the maps of unknown physical conditions such as the thermal pressure $\mathbf{P}$, the intensity of incident radiation fields $\mathbf{G}$ and the depth of the cloud $\mathbf{A_V}$. these 3 parameters are gathered in $\boldsymbol{\theta}_{ij} = (P_{i,j}, G_{i,j}, A_{V,i,j})$ associated to each pixel at position (i, j) of the observed map. Let $\boldsymbol{\Theta} = [\mathbf{P}, \mathbf{G}, \mathbf{A_V}] \in \mathbb{R}^3 \times \mathbb{R}^{N^2}$ the matrix of parameters.

2.2. Challenges

Defining an accurate model to relate the intensity lines $\mathbf{Y}$ back to the physical conditions $\boldsymbol{\Theta}$ raises several challenges. Even if a coherent model taking into account the physics and the chemistry of the cloud is available, this model remains approximate and will always leave room for residual errors between the observed data and simulations. This error might be due to model misspecification, limitations from measuring devices or decisions taken by observatorial astronomers.

In order to cope with these issues, we propose to model these sources of error by relying on the approximate Bayesian computation (ABC) approach of [8]. The idea of this approach is to target a given posterior distribution with an approximate simulation-based technique. Interestingly, instead of considering this ABC approach as an approximate one, the author of [8] considers this approximation as a way to take into account potential errors coming from both the model and the measurements. A related idea can be found in [9] where the authors proposed to learn the distribution of these errors. Based on these works, we will first define the considered Bayesian model and the target posterior distribution in Section 2.3. Then, we will derive in Section 3 a ABC-MCMC algorithm to target this posterior distribution.

2.3. Bayesian model

The observations $y_{ij\ell}$ for each pixel (i, j) and line ℓ are assumed to be unknown quantities related to a physical model $M_\ell(\boldsymbol{\theta}_{ij})$ involving unknown physical conditions $\boldsymbol{\theta}_{ij}$ via the linear model

$$y_{ij\ell} = M_\ell(\boldsymbol{\theta}_{ij}) + \epsilon_{ij\ell}, \text{ where } \epsilon_{ij\ell} \sim \mathcal{N}_{\mathbb{R}_+}(0, \sigma^2), \quad (1)$$

where σ^2 is a known variance and $\mathcal{N}_{\mathbb{R}_+}$ denotes the truncated Gaussian distribution on the positive real line since the in-

tensities are assumed to be supported on $\mathbb{R}_+$. This (truncated) normal assumption is motivated by the central limit theorem which states that the limit of the sum of a large number of unknown independent sources is Gaussian. The function $M_\ell : \mathbb{R}^D \to \mathbb{R}_+$ represents a known correspondence between physical conditions $\boldsymbol{\theta}_{ij}$ related to pixel (i, j) and the corresponding intensity of line ℓ. Here, we use an interpolation of the Meudon PDR code [6]. The Meudon PDR code is a grid which maps more than 1,300 parameters configurations of (P, G, A_V) to 160 standard lines, e.g. those associated to the CO intensities. These parameters stand for the thermal pressure, the intensity of incident radiation fields and the depth of the cloud, respectively. In real settings, observations can be corrupted by the detection threshold used by the astronomers. Indeed, the latter consider that an observation is detected when the intensity line $y_{ij\ell}$ is greater than a known positive threshold ω. Thus, if the line $y_{ij\ell}$ is not detected, they do not provide $y_{ij\ell}$ but the value of the threshold ω. Let $\boldsymbol{\Delta}$ encode whether a line intensity is considered as detected or not where $\delta_{ij\ell} = 1$ if $y_{ij\ell}$ is detected and $\delta_{ij\ell} = 0$ otherwise. Under this statistical model, the likelihood function writes

$$\pi(\mathbf{Y}|\boldsymbol{\Theta}, \boldsymbol{\Delta}) \propto \prod_{l=1}^{L} \prod_{1\leq i,j\leq N} \left[\exp\left(-\frac{1}{2\sigma^2} \left(y_{ij\ell} - M_\ell(\boldsymbol{\theta}_{ij}) \right)^2 \right) \right]^{\delta_{ij\ell}} \times \left[\int_0^\omega \exp\left(-\frac{1}{2\sigma^2} (y - M_\ell(\boldsymbol{\theta}_{ij}))^2 \right) \mathrm{d}y \right]^{1-\delta_{ij\ell}}, \quad (2)$$

Note that the evaluation of this likelihood involves the numerical simulation of the model M_ℓ for physical parameters $\boldsymbol{\theta}_{ij}$).

Since the parameters to infer stand for physical conditions of the ISC, their maps are expected to be at least piecewise continuous. In order to take into account this spatial constraint, we consider the total variation (TV) prior distribution [10] with density

$$\pi(\boldsymbol{\Theta}) \propto \prod_{d=1}^{D} \exp\left(-\tau_d \sum_{1\leq i,j\leq N} \left\| \nabla \boldsymbol{\theta}_{ij}^{(d)} \right\| \right), \quad (3)$$

where $\nabla \boldsymbol{\theta}^{(d)} \in \mathbb{R}^{N^2}$ stands for the 2-dimensional gradient of the map of parameter d, $\tau_d > 0$ is a regularization parameter and $\|\cdot\|$ is the Euclidean norm.

The application of Bayes' rule leads to a posterior distribution for parameters $\boldsymbol{\theta}$ with density

$$\pi(\boldsymbol{\Theta}|\mathbf{Y}, \boldsymbol{\Delta}) \propto \pi(\mathbf{Y}|\boldsymbol{\Theta}, \boldsymbol{\Delta})\pi(\boldsymbol{\Theta}). \quad (4)$$

Note that we consider in this paper that the variance σ^2 and the regularization parameters τ_d are fixed for simplicity reasons. However, if we want to avoid their hand-tuning, we could also consider them as random variables and estimate them through a hierarchical Bayesian model.

3. SAMPLING ALGORITHM FOR INFERENCE

3.1. ABC-MCMC

As introduced in Section 2.2, we will target the posterior in (4) with a special instance of ABC approaches called ABC-MCMC and depicted in Algorithm 1. This algorithm targets an arbitrary close approximation π_ρ of (4) defined by

$$\pi_\rho(\boldsymbol{\Theta}|\mathbf{Y}) \propto \int_{\mathbb{R}^{N^2\times L}} \pi(\boldsymbol{\Theta}|\mathbf{Z}, \boldsymbol{\Delta}) K_\rho\left(\|\mathbf{Z} - \mathbf{Y}\| \right) \mathrm{d}\mathbf{Z}, \quad (5)$$

where K_ρ stands for a kernel density. The dependence on $\boldsymbol{\Delta}$ is intentionally omitted for simplicity. Under weak assumptions on K_ρ and π, the approximate posterior π_ρ satisfies [11]

$$\lim_{\rho\to 0} \pi_\rho(\boldsymbol{\Theta}|\mathbf{Y}) = \pi(\boldsymbol{\Theta}|\mathbf{Y}). \quad (6)$$

Following [8] and assuming that K_ρ is a Gaussian kernel with variance ρ^2, targetting π_ρ instead of π boils down to consider that there is an error $\varepsilon_{ij\ell}$ associated to observation $y_{ij\ell}$ (from the model and/or the measurements) with respect to some underlying 'true value' $z_{ij\ell}$ satisfying

$$\varepsilon_{ij\ell} \sim \mathcal{N}(z_{ij\ell} - y_{ij\ell}, \rho^2). \quad (7)$$

Across iterations t, the so-called instrumental distribution with density $q(\cdot|\boldsymbol{\Theta}^{(t-1)})$ is used to propose a new random value of $\boldsymbol{\Theta}$. The likelihood and therefore a numerical simulation of the physical model with parameter $\boldsymbol{\Theta}$ is used to generate a potential observation $\mathbf{Z}$. Candidates $(\boldsymbol{\Theta}, \mathbf{Z})$ are accepted according to an ABC rejection rule, see Algo. 1.

3.2. Sampling details

In the sequel, the instrumental distribution with density $q(\cdot|\boldsymbol{\Theta}^{(t-1)})$ is chosen as follows

$$q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(t-1)}) = \pi(\boldsymbol{\Theta}) \prod_{d=1}^{D} \mathcal{N}_{\mathbb{R}_+}\left(\boldsymbol{\theta}^{(d)}; \boldsymbol{\theta}^{(d,t-1)}, \lambda_d^2 \mathbf{I}_{N^2} \right) \quad (8)$$

where for all d, $\lambda_d > 0$ stands for the standard deviation. Such an instrumental density simply amounts to consider a random walk around the current value of the parameters penalized by the prior defined in (3). In order to avoid the difficult hand-tuning of the variance λ_d, we consider a Robbins-Monro type algorithm that updates the values of λ_d with the scheme

$$\log(\lambda_d^{(t+1)}) = \log(\lambda_d^{(t)}) + \frac{a^{(t)} - a^\star}{t^{0.1}}, \quad (9)$$

where $a^\star$ is the target acceptance rate and $a^{(t)}$ is the acceptance probability at iteration t.

Since this proposal is not differentiable due to the use of total variations in the prior $\pi(\boldsymbol{\Theta})$, sampling from the latter

Algorithm 1: ABC MCMC

Input: Posterior (4), procedure to generate data from the likelihood, proposal density $q(\cdot|\boldsymbol{\Theta})$, kernel density K_ρ and total nb. of iterations T_{MC}

% Initialization

Generate an initial value $\boldsymbol{\Theta}^{(0)}$;

for $t \leftarrow 1$ **to** T_{MC} **do**

  % Draw a candidate $\boldsymbol{\Theta}$ (instrumental distribution)

  $\boldsymbol{\Theta} \sim q(\cdot|\boldsymbol{\Theta}^{(t-1)})$;

  % Draw the auxiliary variable $\mathbf{Z}$ (likelihood)

  $\mathbf{Z} \sim \pi(\cdot|\boldsymbol{\Theta})$;

  % Accept/reject the candidate value

  Generate $u \sim \mathcal{U}_{[0,1]}$;

  Set $(\boldsymbol{\Theta}^{(t)}, \mathbf{Z}^{(t)}) = (\boldsymbol{\Theta}, \mathbf{Z})$ if $u \leq \dfrac{K_\rho\left(\|\mathbf{Z}-\mathbf{Y}\|\right)\pi\left(\boldsymbol{\Theta}\right)q(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(t-1)})}{K_\rho\left(\left\|\mathbf{Z}^{(t-1)}-\mathbf{Y}\right\|\right)\pi\left(\boldsymbol{\Theta}^{(t-1)}\right)q(\boldsymbol{\Theta}^{(t-1)}|\boldsymbol{\Theta})}$;

  Else set $(\boldsymbol{\Theta}^{(t)}, \mathbf{Z}^{(t)}) = (\boldsymbol{\Theta}^{(t-1)}, \mathbf{Z}^{(t-1)})$;

end

Output: Collection of samples $\left\{\boldsymbol{\Theta}^{(t)}\right\}_{t=1}^{T_{\mathrm{MC}}}$ distributed according to π_ρ in (5) asymptotically.

is conducted by using the proximal MCMC algorithm of [12] called MYULA where the proximity operator of the total variation regularizer has been approximated with the iterative algorithm of [13].

Sampling $\mathbf{Z}$ from the likelihood has been done by first interpolating the Meudon PDR code with a trilinear interpolation and then by drawing $N^2 \times L$ univariate random variables.

4. EXPERIMENTS

This section presents the results of the proposed approach on a 10×10 synthetic map involving some standard regions that can be found in ISCs. This permits to quantify the performances, in particular the bias, of the proposed model since the ground truth is known.

4.1. Experimental design

From the 160 available lines, we only considered the 15 lines associated to the CO molecule. In order to test the proposed approach, we contaminate the synthetic map with some noise. To this purpose, we consider a zero-mean Gaussian noise with standard deviation σ defined by

$$\sigma = \frac{Q}{3}\mathrm{med}\left(I_{\mathrm{CO}^{13-12}}\right), \tag{10}$$

where $Q > 0$ defines the noise level and $\mathrm{med}(I_{\mathrm{CO}^{13-12}})$ stands for the median value of the CO 13-12 line intensity for all the observed pixels. As mentioned previously for real-world situations, the astronomers consider that they detect a line when the intensity line is greater than a threshold value ω. When they consider that they do not detect a line, they replace the observed intensity by an upper limit equal to ω in the data set. In the sequel, we set $Q = 10$ and the threshold values ω are computed for each observation as follows

$$\omega = \sqrt{\sigma^2 + \left(0.2 y_{ij\ell}\right)^2}, \tag{11}$$

In order to model these real cases, we replaced manually the considered non-detected observations by these upper limits.

Due to the presence of the model M_ℓ in (1), the posterior distribution in (4) is highly non-convex and hence admits several maxima. Hence, the good initialization of Algorithm 1 becomes crucial. In order to analyze the behavior of the latter with respect to (w.r.t.) the initial value $\boldsymbol{\Theta}^{(0)}$, we implemented the proposed algorithm with two initializations namely the maximum likelihood (MLE) and the maximum a posteriori estimators (MAP) obtained via an optimization-based algorithm. The fixed parameters have been set to $(\tau_1, \tau_2, \tau_3) = (4.2, 3.6, 13.8)$ after a grid-search procedure while the standard deviation ρ in (7) has been adapted to each line by taking into account the range spanned by the latter. The target acceptance rate $a^\star$ has been set to 0.23 and the number of iterations T_{MC} has been set to 50,000.

Although not directly comparable, we also implemented an optimization-based approach whose initialization is the MLE and which is similar to the one used in [5]. These results will be used as a reference to study the performance of the proposed Bayesian approach.

4.2. Results

Figure 1 compares the 10×10 synthetic maps of physical conditions P, G and A_V, the MAP estimators obtained with an optimization algorithm and minimum mean square estimators (MMSE) obtained with Algorithm 1. The MMSEs obtained with the proposed approach appear to be coherent with the proposed model since the main regions of the synthetic map have been detected. Note that because we both targetted the approximate posterior π_ρ instead of π and used another pointwise estimator (MMSE instead of MAP), the maps obtained with the proposed approach are different from the ones obtained with the optimization-based algorithm. As expected, a different initialization leads to a different MMSE but the derived pointwise estimates are still coherent with the inference task.

Figure 2 depicts the 95% credibility intervals associated to each physical condition and obtained with the proposed approach with the MLE initialization. The range of these intervals is shown in $\log_{10}$-scale. The noisy structure of these credibility maps is mainly due to the very low number of pixels ($N = 10$) which have to be estimated. These credibility maps are a true benefit of the proposed Bayesian approach

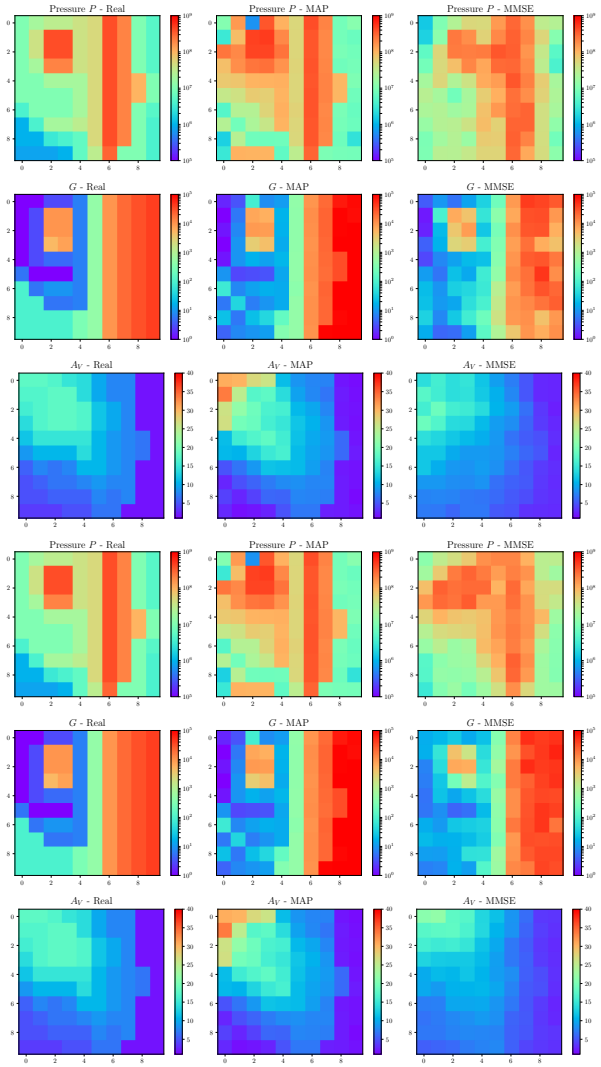

Fig. 1. 10×10 synthetic maps of physical parameters P, G and A_V, MAP estimators obtained with an optimization algorithm and minimum mean square estimators (MMSE) obtained with Algorithm 1.

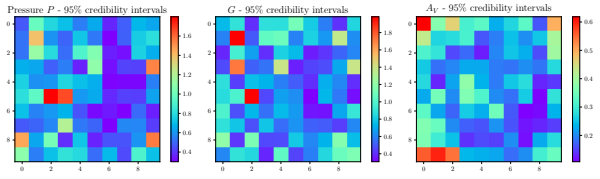

Fig. 2. 95% credibility intervals associated to physical parameters P, G and A_V obtained with Algorithm 1 by taking the MLE as starting value.

since they permit to quantify the uncertainty around resulting estimates.

These results remain preliminary since they are obtained on a synthetic map. They do not reveal to which extent the proposed approach might be efficient for high-dimensional astronomical images since the structure of the image can differ (e.g., presence of filaments). Nevertheless, this work paves the way to the full Bayesian analysis of ISCs and the quantification of uncertainties for inverse problems encountered in physics. This direction of research sounds very promising.

5. CONCLUSION

We present a fully Bayesian approach to tackle challenging inverse problems in astrophysics. We pursue the goal of quantifying the underlying uncertainty by modeling potential measurement errors and by resorting to a simulation-based ABC method. The results obtained on a synthetic map are encouraging and suggest further tests on real and high-dimensional images.

6. REFERENCES

[1] Katia M. Ferrière, "The interstellar environment of our galaxy," *Reviews of Modern Physics*, vol. 73, no. 4, pp. 1031–1066, Oct 2001.

[2] Donald P. Cox, "The Three-Phase Interstellar Medium Revisited," *Annual Review of Astronomy and Astrophysics*, vol. 43, no. 1, pp. 337–385, Sep 2005.

[3] Christopher F. McKee and Eve C. Ostriker, "Theory of Star Formation," *Annual Review of Astronomy and Astrophysics*, vol. 45, no. 1, pp. 565–687, Sep 2007.

[4] Jérôme Pety et al., "The anatomy of the Orion B giant molecular cloud: A local template for studies of nearby galaxies," *A&A*, vol. 599, pp. A98, Jan 2017.

[5] Ronin Wu et al., "Constraining physical conditions for the PDR of Trumpler 14 in the Carina Nebula," *A&A*, vol. 618, pp. A53, 2018.

[6] Franck Le Petit, Cyrine Nehme, Jacques Le Bourlot, and Evelyne Roueff, "A model for atomic and molecular interstellar gas: The Meudon PDR code," *The Astrophysical Journal Supplement Series*, vol. 164, no. 2, pp. 506–529, 2006.

[7] C. P. Robert, *The Bayesian Choice: from decision-theoretic foundations to computational implementation*, Springer, New York, 2 edition, 2001.

[8] R. Wilkinson, "Approximate Bayesian computation (ABC) gives exact results under the assumption of model error," *Statistical applications in genetics and molecular biology*, vol. 12, pp. 1–13, 2013.

[9] Oliver Ratmann, Christophe Andrieu, Carsten Wiuf, and Sylvia Richardson, "Model criticism based on likelihood-free inference, with an application to protein network evolution," *Proceedings of the National Academy of Sciences*, vol. 106, no. 26, pp. 10576–10581, 2009.

[10] A. Chambolle, M. Novaga, D. Cremers, and T. Pock, "An introduction to total variation for image analysis," in *in Theoretical Foundations and Numerical Methods for Sparse Recovery, De Gruyter*, 2010.

[11] S.A. Sisson, Y. Fan, and M. Beaumont, Eds., *Handbook of Approximate Bayesian Computation*, Chapman and Hall/CRC Press, 1 edition, 2018.

[12] A. Durmus, E. Moulines, and M. Pereyra, "Efficient Bayesian computation by proximal Markov chain Monte Carlo: When Langevin meets Moreau," *SIAM Journal on Imaging Sciences*, vol. 11, no. 1, pp. 473–506, 2018.

[13] Antonin Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, pp. 89–97, Jan 2004.