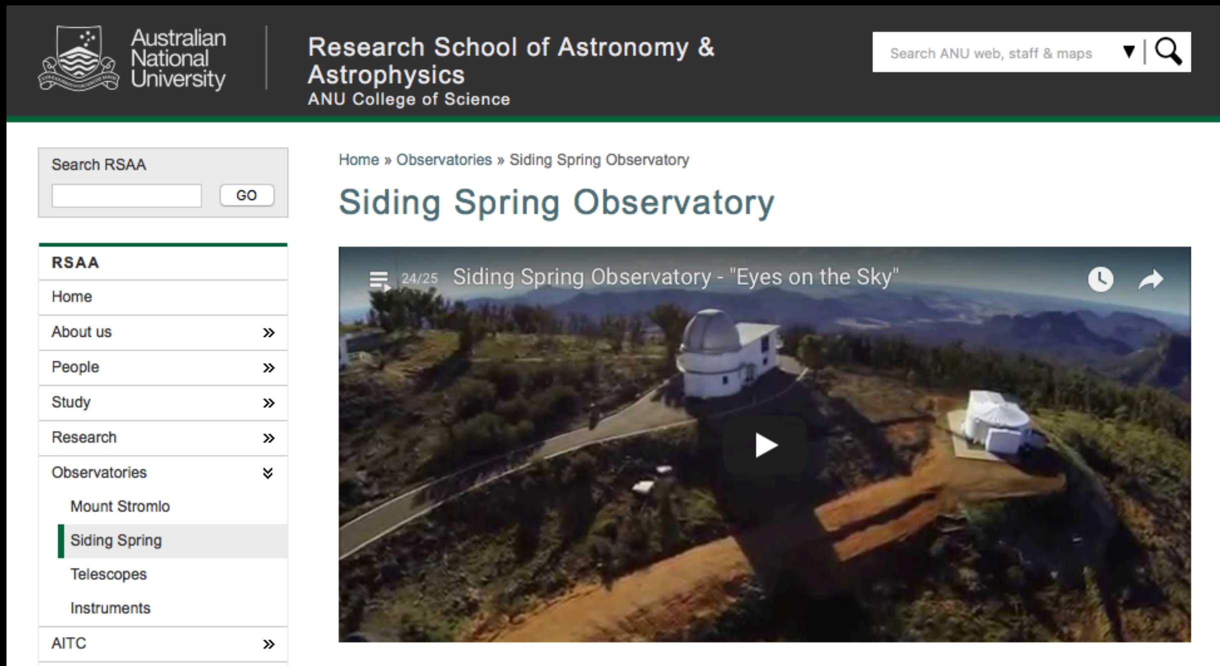
ASTRO-ESTADISTICA

DRA. CARMEN SÁNCHEZ GIL
UNIVERSIDAD DE CÁDIZ
INSTITUTO DE ASTROFÍSICA DE ANDALUCIA (CSIC)

INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA, IAA-CSIC

https://www.iaa.csic.es/



http://rsaa.anu.edu.au/observatories/siding-spring-observatory

http://angelrls.blogalia.com

El Lobo Rayado

CSIRO

Australia Telescope National Facility

https://www.atnf.csiro.au/outreach/visiting/narrabri/index.html

INFERENCIA BAYESIANA

Data Modelling & Parameter estimation

School on Statistics, Data Mining, and Machine Learning
Instituto de Astrofísica de Andalucía (CSIC)

Dr. Carmen Sánchez Gil

Tuesday, Nov 5th 2019

Universidad de Cádiz

Data modelling & parameter estimation

- We perform experiments or make observations in order to learn about a phenomenon, which only can be partially observed.
- First step: describe the resulting data (plots, summaries, statistics, etc)
- To interpret the data we usually have to model them
- Data (measurements) are always noisy
- Inference is the process of making general statements about a phenomenon, via a model, using noisy and incomplete data.
- *Generative model* is the theoretical model that generates (simulates) the observable data from the model parameters (mathematical equation)
- *Measurement or noise model* describes how the measurement process affects our data. It describes a probability distribution over possible observations given the ideal (noise-free) data, i.e. the Likelihood.

Example

$x \sim \mathcal{N}(\mu, \sigma)$, where x is the measured data, $\mu = g(\theta)$ is the ideal data, and σ is the uncertainty in the measurement: $f(x) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(x-g(\theta))^2}{2\sigma^2}\right)$

Data modelling & parameter estimation

The key to data modeling is to use the given data, together the generative and measurement models to make consistent, probabilistic inferences.

Given some data D, for a specified model M, with parameter(s) θ:

- Parameter estimation. To infer the *parameter posterior* pdf $p(\theta|D, M)$
- Model comparison. Given a set of different models $\{M_i\}$, find out which one is best supported by the data: *model posterior probability* $P(M_i|D)$, or *posterior odds ratio* of two models $P(D|M_i)/P(D|M_j)$
- Prediction. Predict some new data, $p(\tilde{x}|D, M)$

INFERENCIA BAYESIANA

Probability Theory

Bayesian Inference

Given a *model* M, with parameter(s) θ,

- Likelihood: $p(x|\theta, M) \equiv p(x|\theta)$, *sampling* or *data distribution*. Key function in data modeling, it describes both the *phenomenon* and the *measurements*.
- Prior: $p(\theta|M) \equiv p(\theta)$, pdf over the model params. θ. Information we have, independent of the data, about the possible values of θ.
- *Joint probability d.* for θ and x: $p(\theta, x) = p(\theta)p(x|\theta)$
- Posterior: $p(\theta|x, M)$ pdf over the model params., given the data and the background inform. on M, is the answer to an inference problem

Bayes' rule

$$p(\theta|x) = \frac{p(\theta, x)}{p(x)} = \frac{p(\theta)p(x|\theta)}{p(x)}$$

$$\propto \underbrace{p(\theta)p(x|\theta) = p^*(\theta|x)}_{\text{unnormalized post}}$$

HTTPS://WWW.UCA.ES/NOTICIA/CAPTAN-LOS-DISTINTOS-PULSOS-EN-LA-LLAMARADA-MAGNETICA-GIGANTE-DE-UNA-ESTRELLA-DE-NEUTRONES/

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature › articles › article

Article | Published: 22 December 2021

Very-high-frequency oscillations in the main peak of a magnetar giant flare

A. J. Castro-Tirado, N. Østgaard ✉, E. Göğüş ✉, C. Sánchez-Gil, J. Pascual-Granado, V. Reglero, A. Mezentsev ✉, M. Gabler ✉, M. Marisaldi ✉, T. Neubert, C
Kuvvetli, P. Cerdá-Durán, J. Navarro-González, J. A. Font,
Brandt, M. D. Caballero-García, I. M. Carrasco-García, A. C
C. J. Eyles, E. Fernández-García, G. Genov, S. Guziy, Y.-D.
K. Peng, C. Pérez del Pulgar, A. J. Reina Terol, E. Rodrígue
S. Yang — Show fewer authors

Nature **600**, 621–624 (2021) | Cite this article

Abstract

Magnetars are strongly magnetized, isolated neutro
around 10^{15} gauss, luminosities of approximately 10
periods of about 0.3–12.0 s. Very energetic giant flar

Abstract

Magnetars are strongly magnetized, isolated neutron stars[1,2,3] with magnetic fields up to around 10^{15} gauss, luminosities of approximately 10^{31}–10^{36} ergs per second and rotation periods of about 0.3–12.0 s. Very energetic giant flares from galactic magnetars (peak luminosities of 10^{44}–10^{47} ergs per second, lasting approximately 0.1 s) have been detected in hard X-rays and soft γ-rays[4], and only one has been detected from outside our galaxy[5]. During such giant flares, quasi-periodic oscillations (QPOs) with low (less than 150 hertz) and high (greater than 500 hertz) frequencies have been observed[6,7,8,9], but their statistical significance has been questioned[10]. High-frequency QPOs have been seen only during the tail phase of the flare[9]. Here we report the observation of two broad QPOs at approximately 2,132 hertz and 4,250 hertz in the main peak of a giant γ-ray flare[11] in the direction of the NGC 253 galaxy[12,13,14,15,16,17], disappearing after 3.5 milliseconds. The flare was detected on 15 April 2020 by the Atmosphere–Space Interactions Monitor instrument[18,19] aboard the International Space Station, which was the only instrument that recorded the main burst phase (0.8–3.2 milliseconds) in the full energy range (50×10^3 to 40×10^6 electronvolts) without suffering from saturation effects such as deadtime and pile-up. Along with sudden spectral variations, these extremely high-frequency oscillations in the burst peak are a crucial component that will aid our understanding of magnetar giant flares.

HTTPS://WWW.UCA.ES/NOTICIA/CAPTAN-LOS-DISTINTOS-PULSOS-EN-LA-LLAMARADA-MAGNETICA-GIGANTE-DE-UNA-ESTRELLA-DE-NEUTRONES/

nature

Article | Published: 22 December 2021

Very-high-frequency oscillations in the main peak of a magnetar giant flare

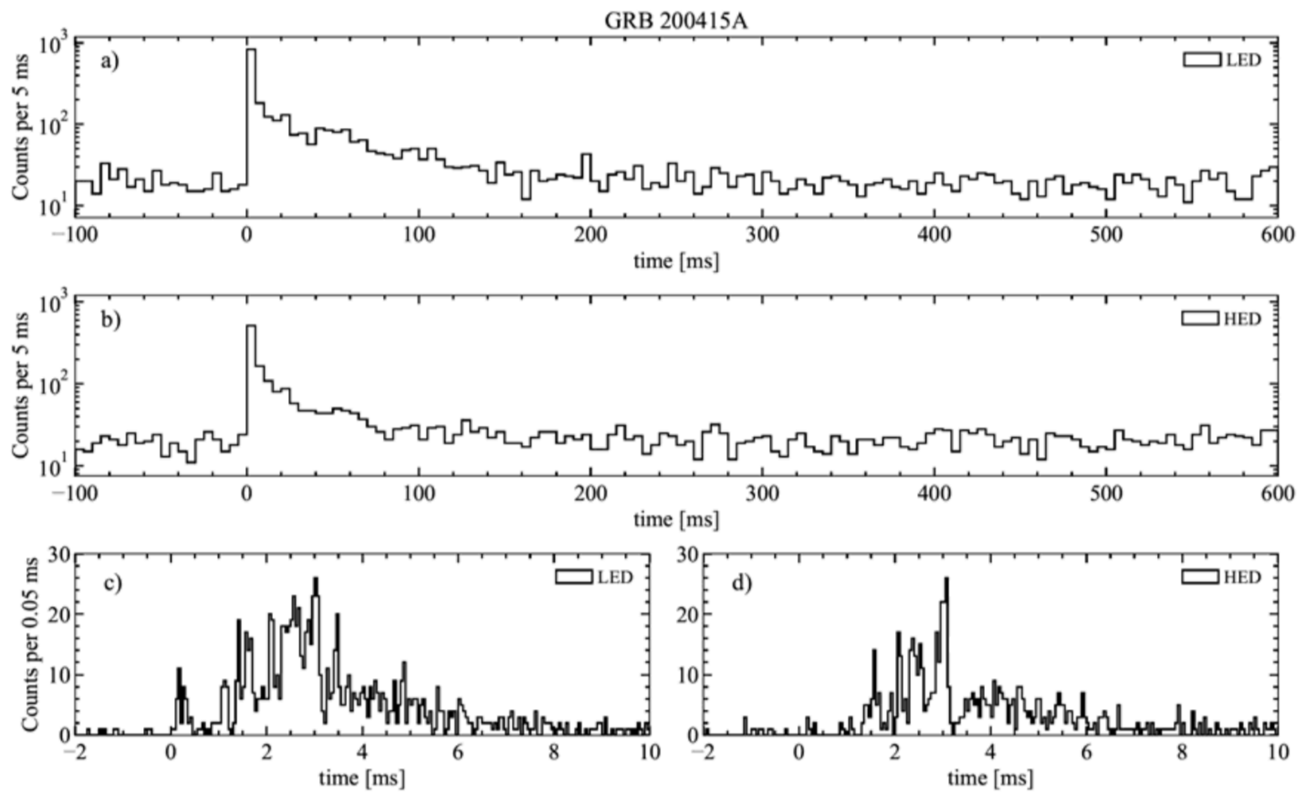
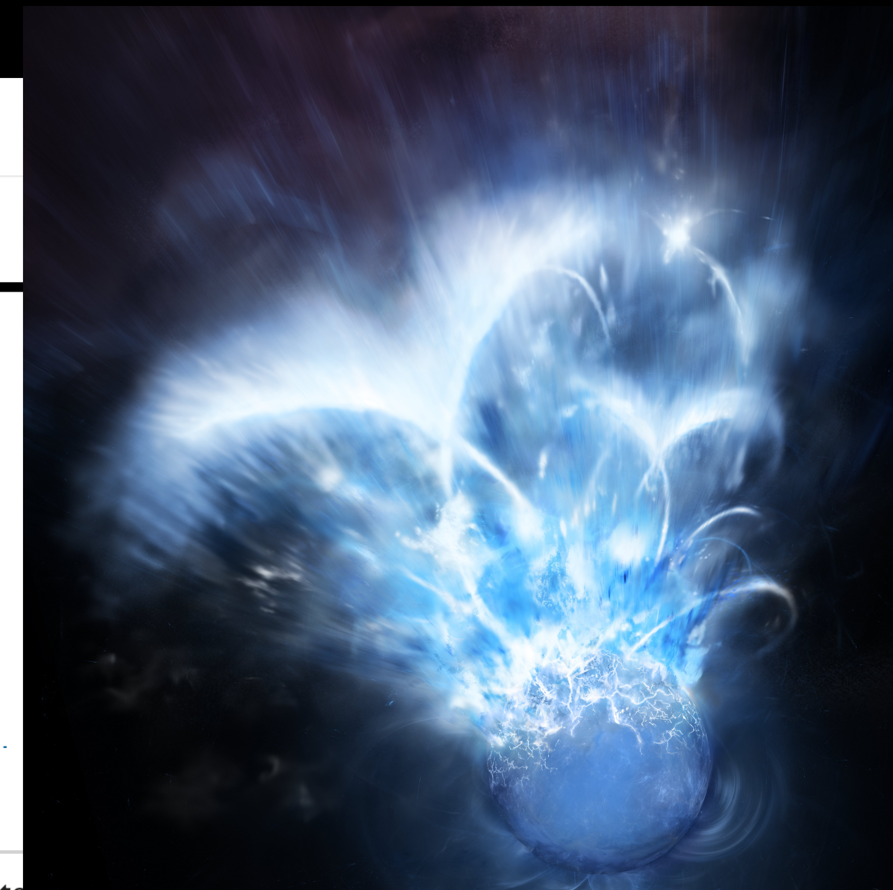
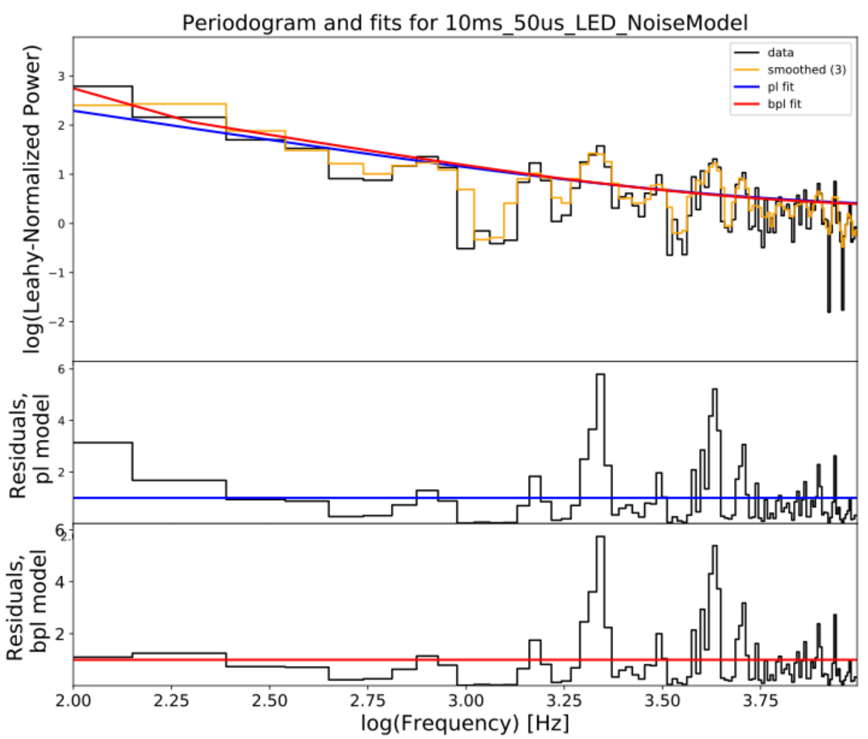


Fig. 1. Temporal variability of GRB 200415A. a) The ASIM LED (low-energy detector) hard X-ray light curve of the magnetar giant flare. b) The ASIM HED (high-energy detector) gamma-ray light curves of the magnetar giant flare. Both light curves have a complex structure, displaying six intensity peaks during the first 3.2 ms. The time interval ranges from -100 to +600 ms, with the data binned to 5 ms. c) and d) A zoom of the first 10 ms with the LED and HED data respectively binned to 50 µs.



ːimately 10^{31}–10^{36} ergs per second and rotation
ic giant flares from galactic magnetars (peak
l, lasting approximately 0.1 s) have been detected in
has been detected from outside our galaxy[5]. During
ɪns (QPOs) with low (less than 150 hertz) and high
ː been observed[6,7,8,9], but their statistical
-frequency QPOs have been seen only during the tail
servation of two broad QPOs at approximately 2,132
a giant γ-ray flare[11] in the direction of the NGC 253
ɜ milliseconds. The flare was detected on 15 April
ons Monitor instrument[18,19] aboard the
e only instrument that recorded the main burst
ergy range (50×10^3 to 40×10^6 electronvolts)
such as deadtime and pile-up. Along with sudden
frequency oscillations in the burst peak are a crucial
ıg of magnetar giant flares.

[HTTPS://WWW.UCA.ES/NOTICIA/CAPTAN-LOS-DISTINTOS-PULSOS-EN-LA-LLAMARADA-MAGNETICA-GIGANTE-DE-UNA-ESTRELLA-DE-NEUTRONES/](https://www.uca.es/noticia/captan-los-distintos-pulsos-en-la-llamarada-magnetica-gigante-de-una-estrella-de-neutrones/)



of power spectral shapes well represented in nature are power laws:

H₀: $$P(\nu) = \beta\nu^{-\alpha} + \gamma, \qquad (7)$$

where α is the power-law index, and broken power laws, which can be reduced to Equation (7) by setting $\rho = 0$:

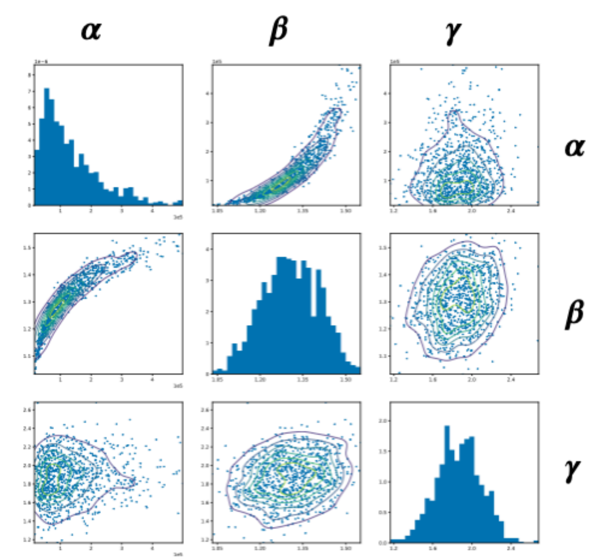
H₁: $$P(\nu) = \beta\nu^{-\alpha_1}\left(1 + \left\{\frac{\nu}{\delta}\right\}^{(\alpha_2-\alpha_1)/\rho}\right)^{-\rho} + \gamma, \qquad (8)$$

$$\theta_{\mathrm{MAP}} = \arg\max_{\theta}\, p(\theta|\mathbf{I}, H),$$

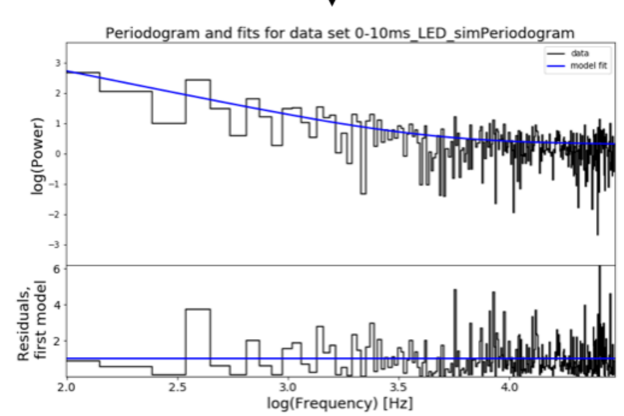
$$T_{\mathrm{LRT}} = -2\log\frac{p(\mathbf{I}|\hat{\theta}^0_{\mathrm{MAP}}, H_0)}{p(\mathbf{I}|\hat{\theta}^1_{\mathrm{MAP}}, H_1)} = D_{\min}(H_0) - D_{\min}(H_1).$$

Bayesian Inference for θ:

- Maximum a Posteriori estimate
- posterior sample



$\{\theta_1,\ldots\theta_{1000}\}$



$$T_{\mathrm{SSE}} = \chi^2(\mathbf{I}, \hat{\theta}), \qquad (10)$$

where

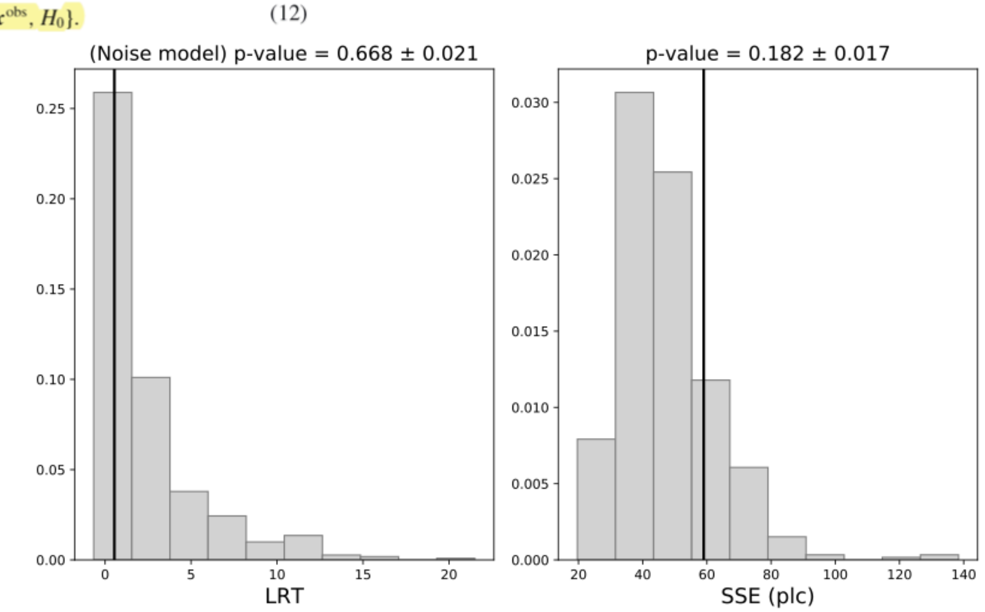
$$\chi^2(\mathbf{I}, \theta) = \sum_{j=1}^{N/2}\frac{(I_j - E[I_j|\theta])^2}{E[I_j|\theta]} = \sum_{j=1}^{N/2}\left(\frac{I_j - S_j(\theta)}{S_j(\theta)}\right)^2$$

POSTERIOR PREDICTIVE P-VALUES

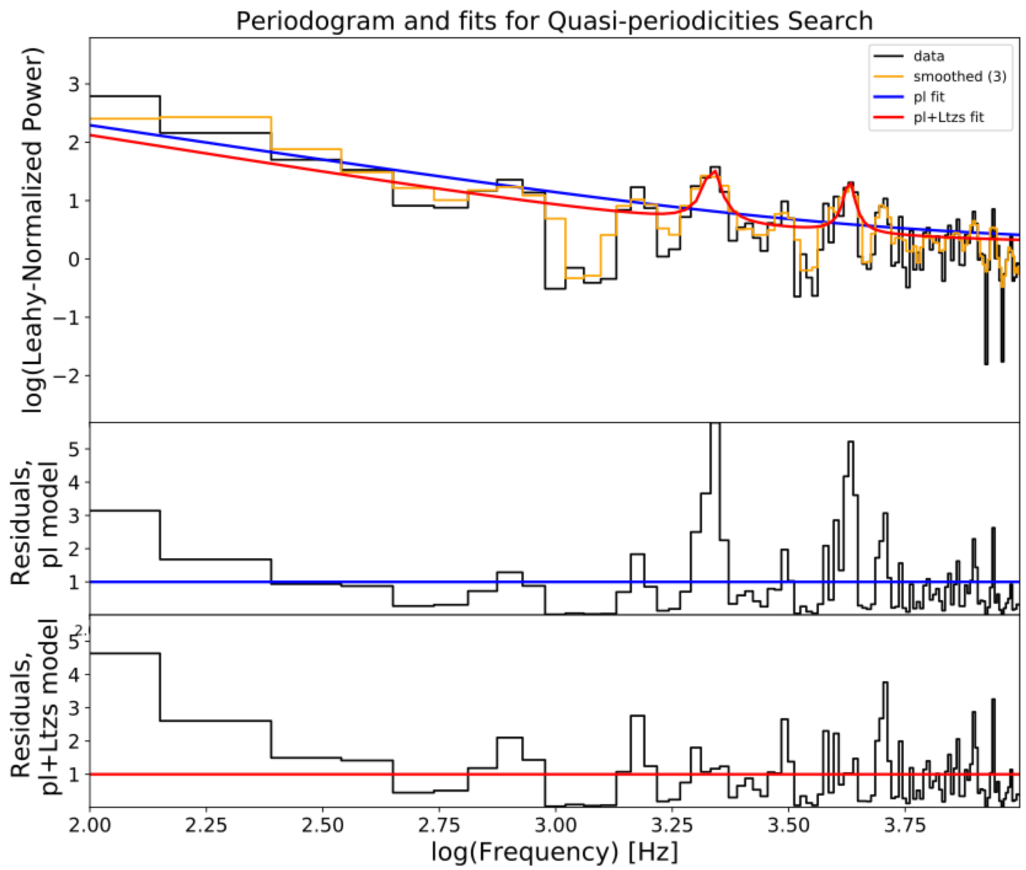
given the parameters. The Bayesian p-value is the (tail area) probability that the replicated data could give a test statistic at least as extreme as that observed:

$$p_{\mathrm{B}}(\boldsymbol{x}) = \int p_{\mathrm{C}}(\boldsymbol{x}^{\mathrm{obs}}, \boldsymbol{\theta})\, p(\boldsymbol{\theta}|\boldsymbol{x}^{\mathrm{obs}})\, d\boldsymbol{\theta}$$
$$= \Pr\{T(\boldsymbol{x}^{\mathrm{rep}}) \geq T(\boldsymbol{x}^{\mathrm{obs}})|\boldsymbol{x}^{\mathrm{obs}}, H_0\}. \qquad (12)$$

$\{LRT_1,\ldots LRT_{1000}\}$



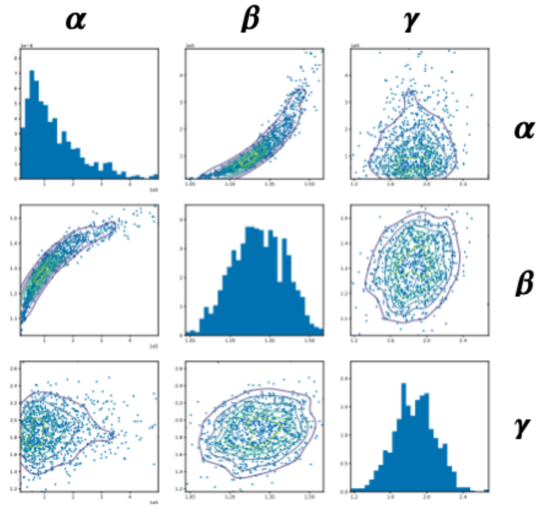
HTTPS://WWW.UCA.ES/NOTICIA/CAPTAN-LOS-DISTINTOS-PULSOS-EN-LA-LLAMARADA-MAGNETICA-GIGANTE-DE-UNA-ESTRELLA-DE-NEUTRONES/



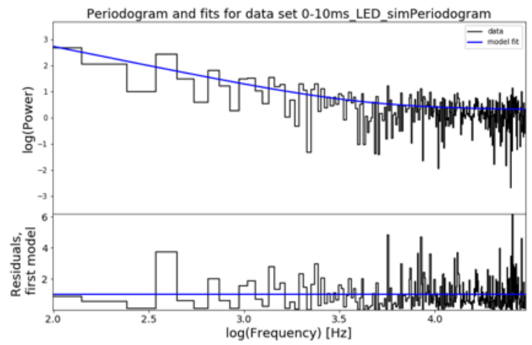
QPO Search: model selection problem

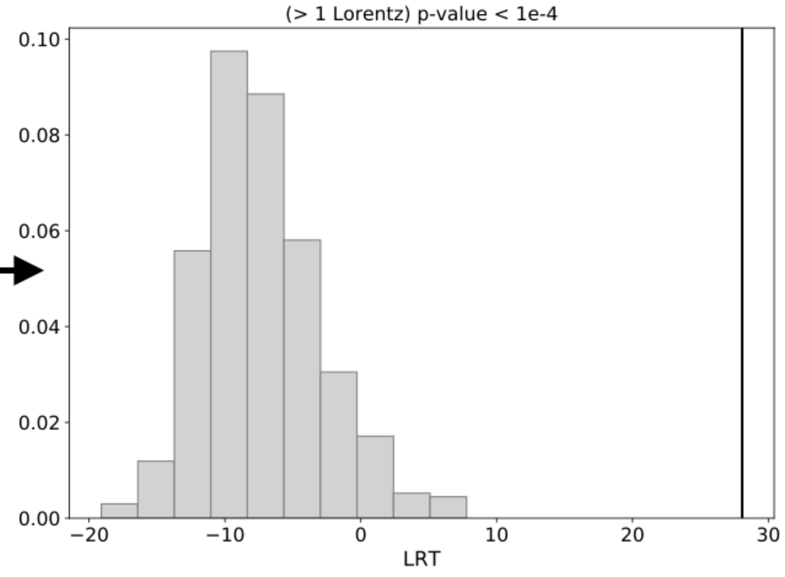
H_0: pure broadband model
H_1: mixture model broadband noise process + **Lorentzian**

$$T_{\mathrm{LRT}} = -2\log\frac{p(\mathbf{I}|\hat{\theta}^0_{\mathrm{MAP}}, H_0)}{p(\mathbf{I}|\hat{\theta}^1_{\mathrm{MAP}}, H_1)} = D_{\min}(H_0) - D_{\min}(H_1).$$

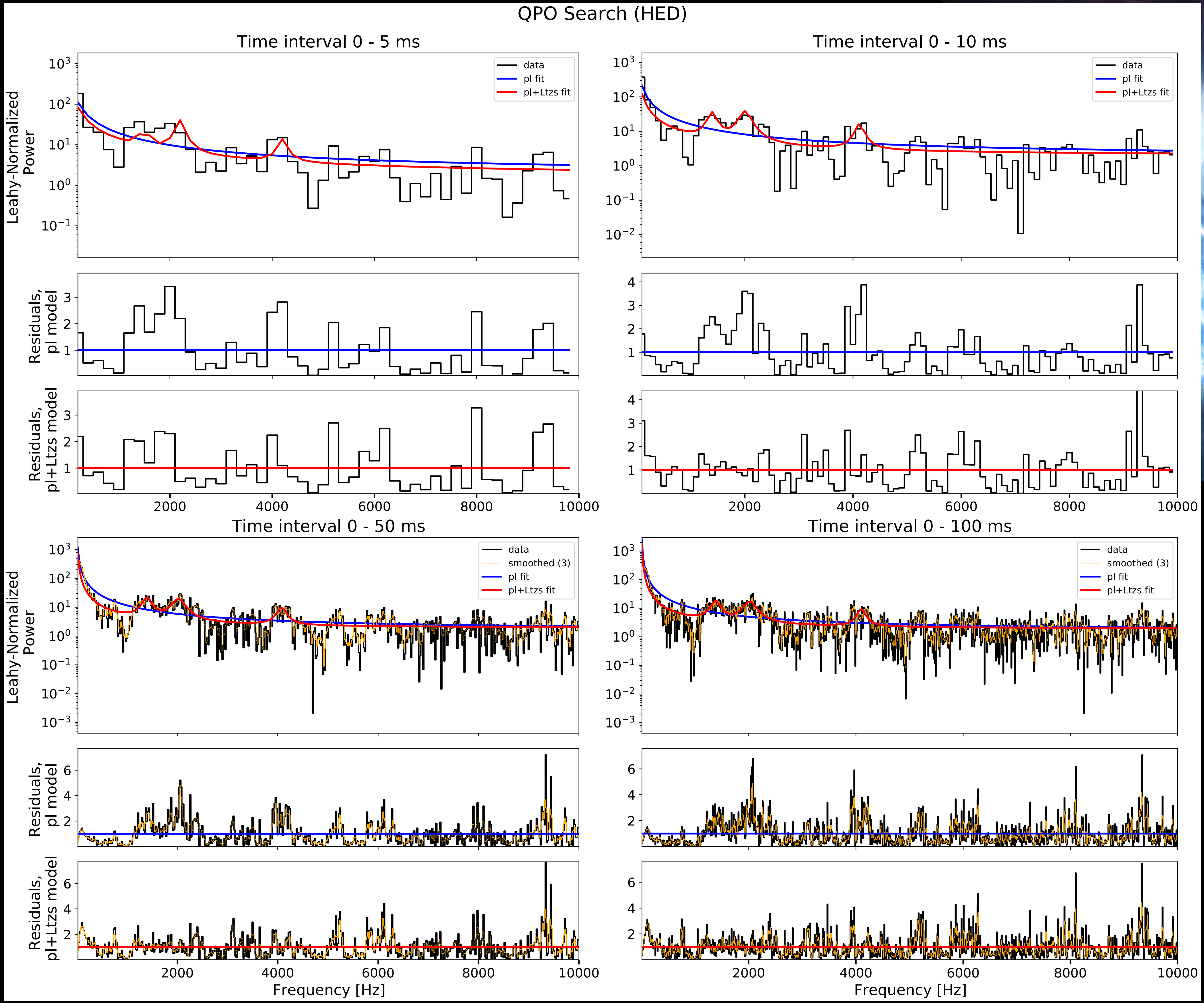


$\{\theta_1, \ldots \theta_{1000}\}$

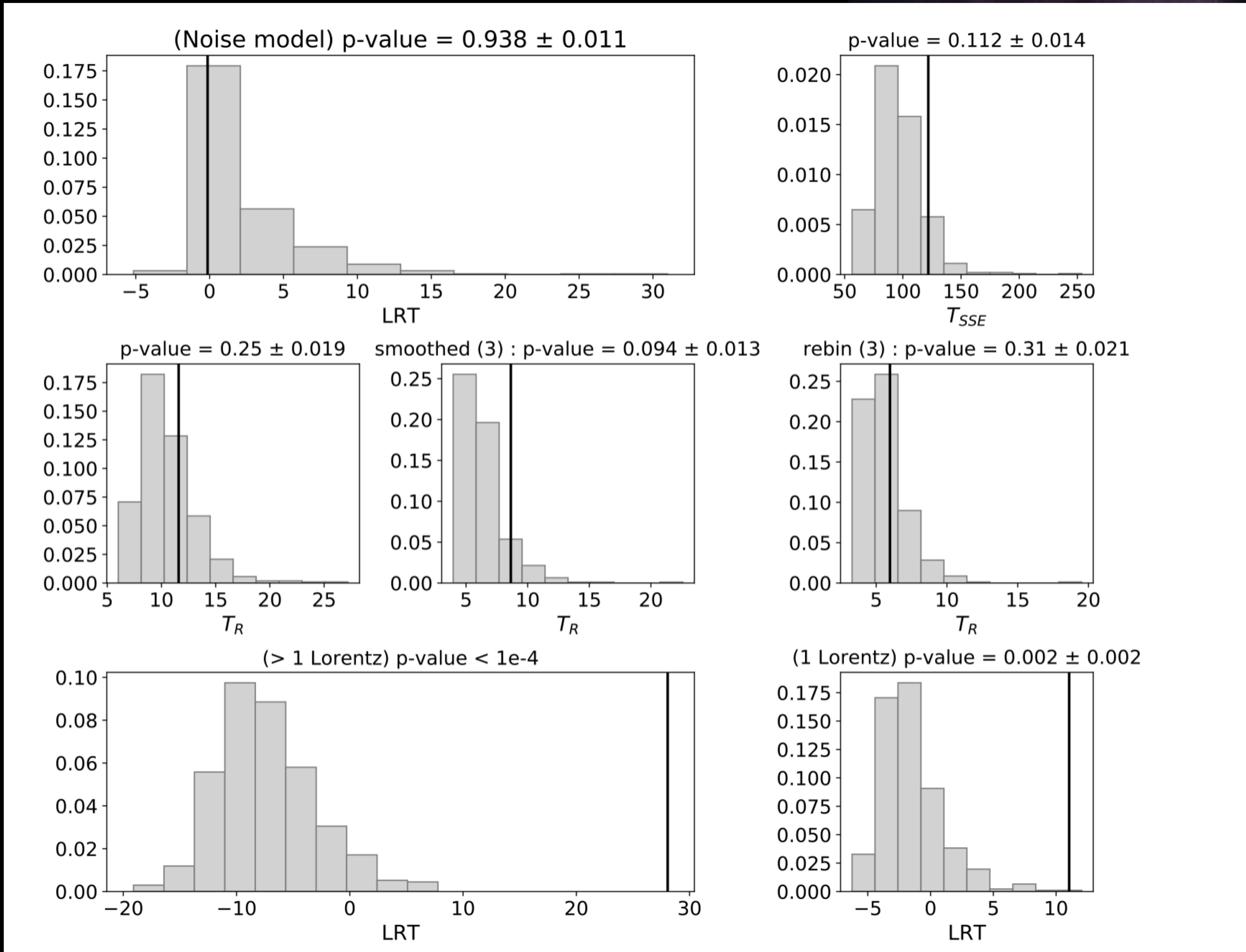




[HTTPS://WWW.UCA.ES/NOTICIA/CAPTAN-LOS-DISTINTOS-PULSOS-EN-LA-LLAMARADA-MAGNETICA-GIGANTE-DE-UNA-ESTRELLA-DE-NEUTRONES/](https://www.uca.es/noticia/captan-los-distintos-pulsos-en-la-llamarada-magnetica-gigante-de-una-estrella-de-neutrones/)

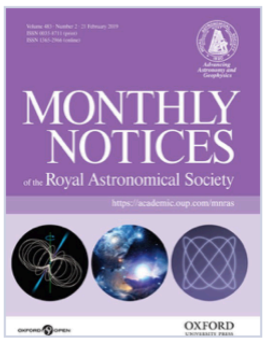


[HTTPS://WWW.UCA.ES/NOTICIA/CAPTAN-LOS-DISTINTOS-PULSOS-EN-LA-LLAMARADA-MAGNETICA-GIGANTE-DE-UNA-ESTRELLA-DE-NEUTRONES/](https://www.uca.es/noticia/captan-los-distintos-pulsos-en-la-llamarada-magnetica-gigante-de-una-estrella-de-neutrones/)



INFERENCIA BAYESIANA

Monthly Notices
of the Royal Astronomical Society



Volume 483, Issue 2
February 2019

Article Contents

Hierarchical Bayesian approach for estimating physical properties in nearby galaxies: Age Maps (Paper II) FREE

M Carmen Sánchez-Gil ✉, Emilio J Alfaro ✉, Miguel Cerviño ✉, Enrique Pérez ✉, Joss Bland-Hawthorn, D Heath Jones

Monthly Notices of the Royal Astronomical Society, Volume 483, Issue 2, February 2019, Pages 2641–2670, https://doi.org/10.1093/mnras/sty3106
Published: 16 November 2018 **Article hist**

PDF | Split View | Cite | Per

Abstract

One of the fundamental goals of modern a
physical parameters of galaxies. We prese
compute age maps from images in the H α
filter, TTF), ultraviolet band (GALEX far U
24, 70, and 160 μm). We present the burs
a sample of nearby and nearly face-on ga
good relative indicator of the very recent
nascent star-forming region evolves, the
the UV continuum, leading to a decrease i

Abstract

One of the fundamental goals of modern astrophysics is to estimate the physical parameters of galaxies. We present a hierarchical Bayesian model to compute age maps from images in the H α line (taken with Taurus tunable filter, TTF), ultraviolet band (GALEX far UV, FUV), and infrared bands (Spitzer 24, 70, and 160 μm). We present the burst ages for young stellar populations in a sample of nearby and nearly face-on galaxies. The H α to FUV flux ratio is a good relative indicator of the very recent star formation history (SFH). As a nascent star-forming region evolves, the H α line emission declines earlier than the UV continuum, leading to a decrease in the H α/FUV ratio. Using star-forming galaxy models, sampled with a probabilistic formalism, and allowing for a variable fraction of ionizing photons in the clusters, we obtain the corresponding theoretical ratio H α/FUV to compare with our observed flux ratios, and thus to estimate the ages of the observed regions. We take into account the mean uncertainties and the interrelationships between parameters when computing H α/FUV. We propose a Bayesian hierarchical model where a joint probability distribution is defined to determine the parameters (age, metallicity, IMF) from the observed data (the observed flux ratios H α/FUV). The joint distribution of the parameters is described through independent and identically distributed (i.i.d.) random variables generated through MCMC (Markov Chain Monte Carlo) techniques.

INFERENCIA BAYESIANA

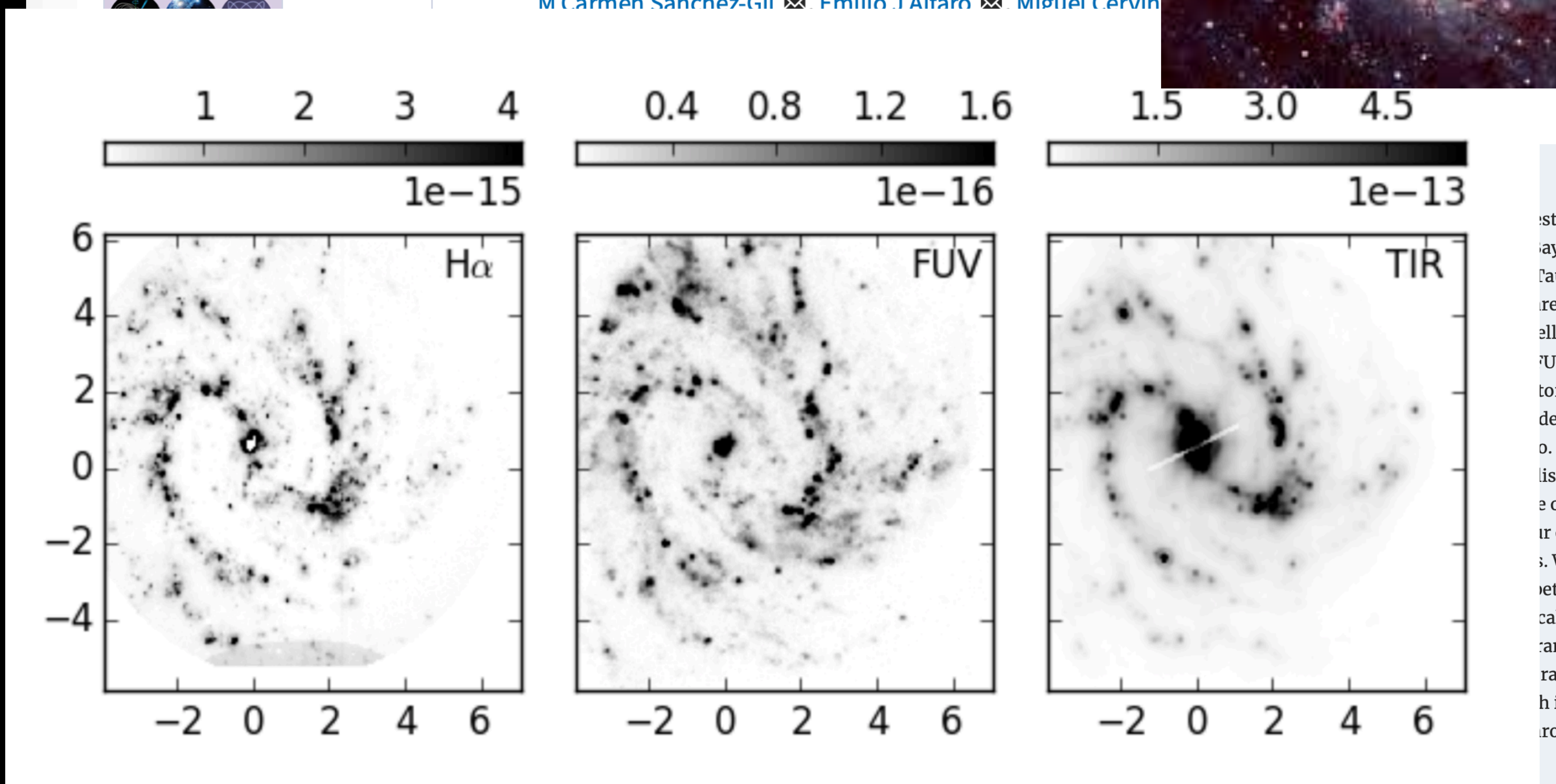

Monthly Notices
of the Royal Astronomical Society

Issues ▾ Advance articles Submit ▾ Purchase Alerts About ▾

MONTHLY NOTICES of the Royal Astronomical Society

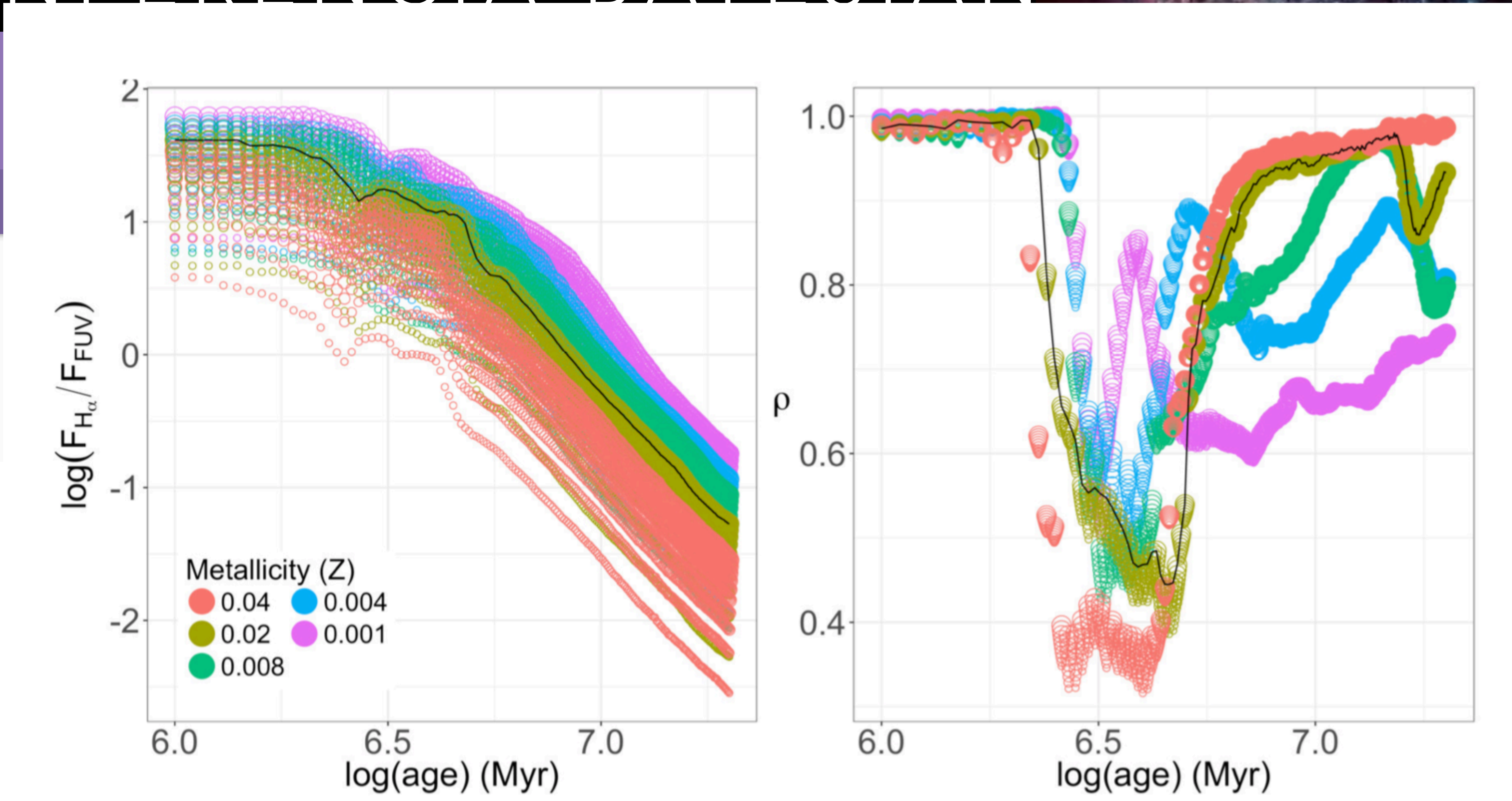
Hierarchical Bayesian approach fo
physical properties in nearby galax
(Paper II) FREE

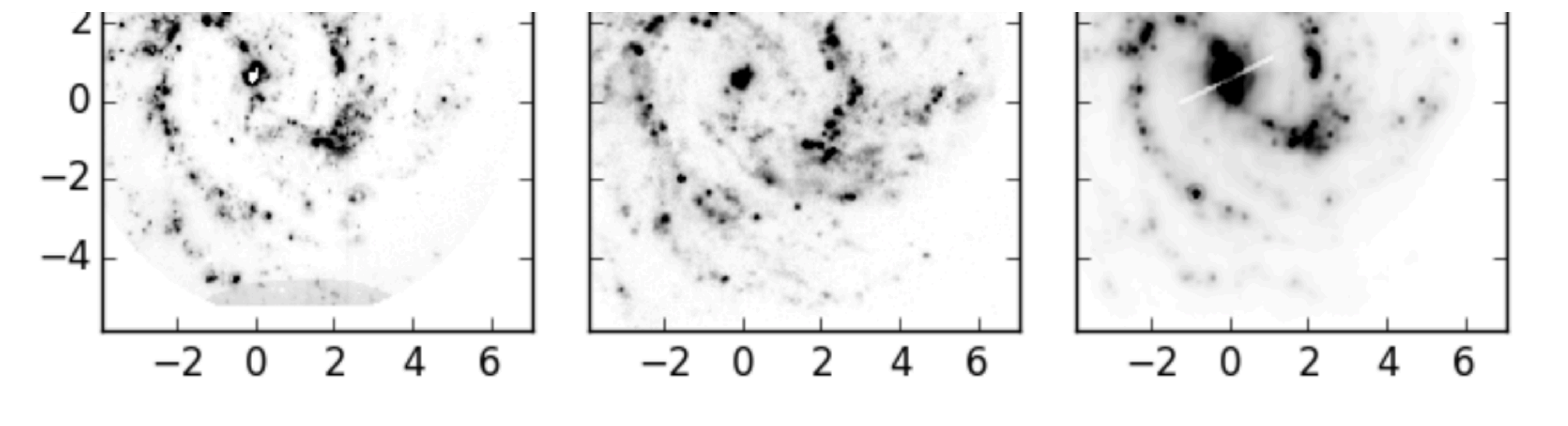
M Carmen Sánchez-Gil, Emilio J Alfaro, Miguel Cerviñ



stimate the
ayesian model to
Taurus tunable
red bands (Spitzer
ellar populations in
UV flux ratio is a
tory (SFH). As a
declines earlier than
o. Using star-
lism, and allowing
e obtain the
r observed flux
. We take into
etween parameters
cal model where a
ameters (age,
ratios H α/FUV).
h independent and
rough MCMC

INFERENCIA BAYESIAN





er
in
tory (SFH). As a
declines earlier than
o. Using star-
lism, and allowing
e obtain the
ır observed flux
s. We take into
between parameters
cal model where a
rameters (age,
ratios H α/FUV).
h independent and
rough MCMC

INFERENCIA BAYESIANA

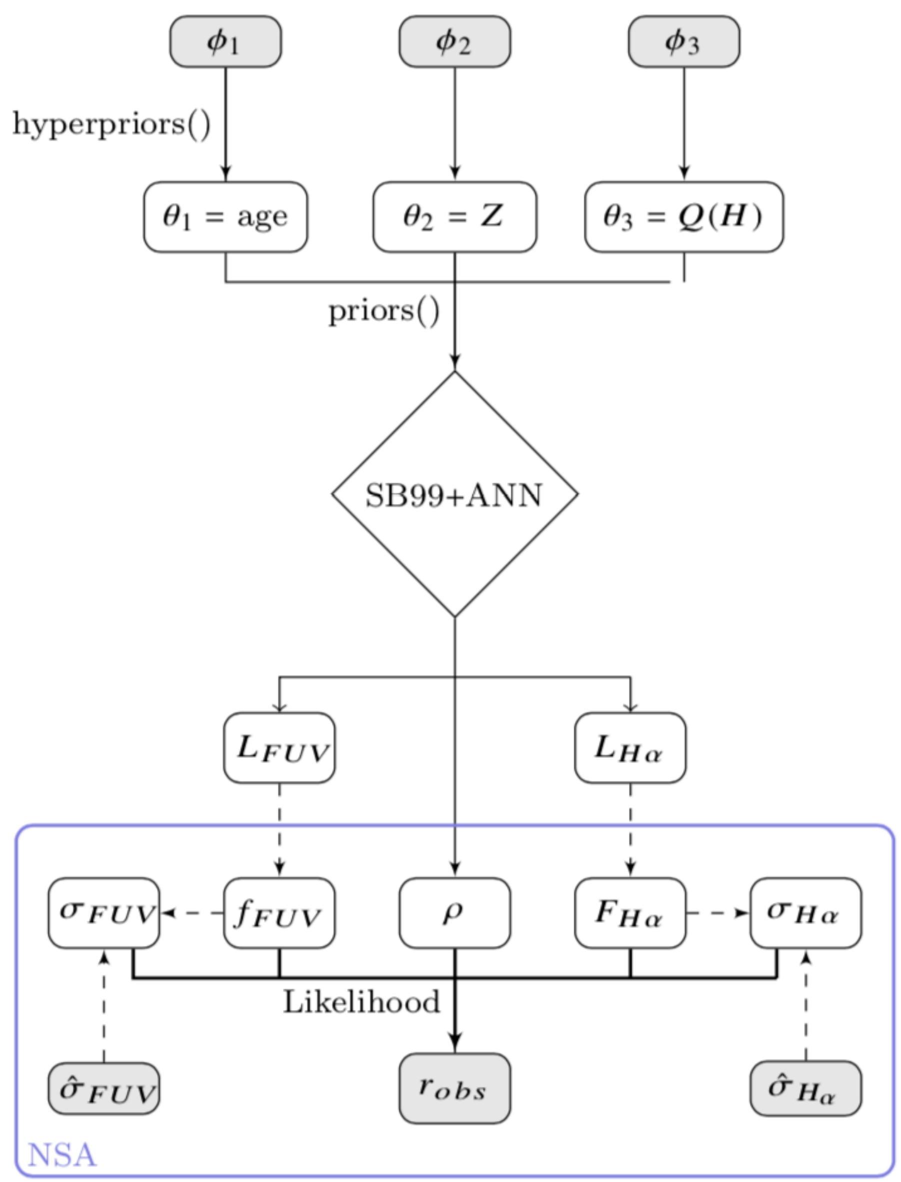


Figure 5. The hierarchical Bayesian model in plate notation. Blank shaded nodes represent random fixed (or observed) values, respectively. Arrows represent the kind of relationship between variables: probabilistic (solid arrows) and deterministic (dashed arrows). In the central part of the graphical model we show deterministic calculations done by SB99 scripts and the artificial neural network (ANN). The blue square at the bottom, encircles the variables involved in the Likelihood (7), as well as the observed ratio which are the input for the NSA.

$$\begin{aligned}\theta_1 &\sim \mathcal{U}(0.1, 20)\\ \theta_2 &\sim \mathcal{U}(0.001, 0.04)\\ \theta_3 &\sim \mathcal{U}(0.1, 1)\end{aligned}$$

$$\begin{aligned}F_{H\alpha} &= 4\pi D^2 L_{H\alpha}, (erg\, s^{-1} cm^{-2})\\ f_{FUV} &= 4\pi D^2 L_{\lambda FUV}, (erg\, s^{-1} cm^{-2} \mathring{A}^{-1})\end{aligned}$$

$$\psi(r) = \frac{b(r)d(r)}{\sqrt{2\pi}\sigma_{H\alpha}\sigma_{FUV} a^3(r)}\left(2\Phi\left(\frac{b(r)}{\sqrt{1-\rho^2}a(r)}\right)-1\right)+ + \frac{\sqrt{1-\rho^2}}{\pi\sigma_{H\alpha}\sigma_{FUV} a^2(r)}\exp\left\{-\frac{c}{2(1-\rho^2)}\right\}, \quad (6)$$

where r is the model flux ratio, and parameters $a(r)$, $b(r)$, c, and $d(r)$ are defined as

$$\begin{aligned}a(r) &= \left(\frac{r^2}{\sigma^2_{H\alpha}} - \frac{2\rho r}{\sigma_{H\alpha}\sigma_{FUV}} + \frac{1}{\sigma^2_{FUV}}\right)^{1/2}\\ b(r) &= \frac{rF^2_{H\alpha}}{\sigma^2_{H\alpha}} - \frac{\rho(F_{H\alpha} + r f_{FUV})}{\sigma_{H\alpha}\sigma_{FUV}} + \frac{f_{FUV}}{\sigma^2_{FUV}}\\ c &= \frac{F^2_{H\alpha}}{\sigma^2_{H\alpha}} - \frac{2\rho F_{H\alpha} f_{FUV}}{\sigma_{H\alpha}\sigma_{FUV}} + \frac{f^2_{FUV}}{\sigma^2_{FUV}}\\ d(r) &= \exp\left\{\frac{b^2(r) - c a^2(r)}{2(1-\rho^2)a^2(r)}\right\}.\end{aligned}$$

$$p(\hat{r}|\boldsymbol{\theta}, \mathcal{H}) = \psi(\hat{r}|F_{H\alpha}, f_{FUV}, \sigma_{H\alpha}, \sigma_{FUV}, \rho)$$

The joint posterior probability distribution $p(\boldsymbol{\theta}|\hat{r})$ can be rewritten by using Bayes' theorem according to

$$p(\boldsymbol{\theta}|\hat{r}) = \frac{p(\hat{r}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\hat{r})} \propto p(\hat{r}|\boldsymbol{\theta})p(\boldsymbol{\theta}), \quad (8)$$

$$p(\text{age}|\hat{r}) = p(\theta_1|\hat{r}) = \iint p(\boldsymbol{\theta}|\hat{r})\, d\theta_2 d\theta_3$$

INFERENCIA BAYESIANA

$$p(\boldsymbol{\theta}|\hat{r}) = \frac{p(\hat{r}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\hat{r})} \propto p(\hat{r}|\boldsymbol{\theta})p(\boldsymbol{\theta}),$$

$$p(\mathrm{age}|\hat{r}) = p(\theta_1|\hat{r}) = \iint p(\boldsymbol{\theta}|\hat{r})\, d\theta_2 d\theta_3$$

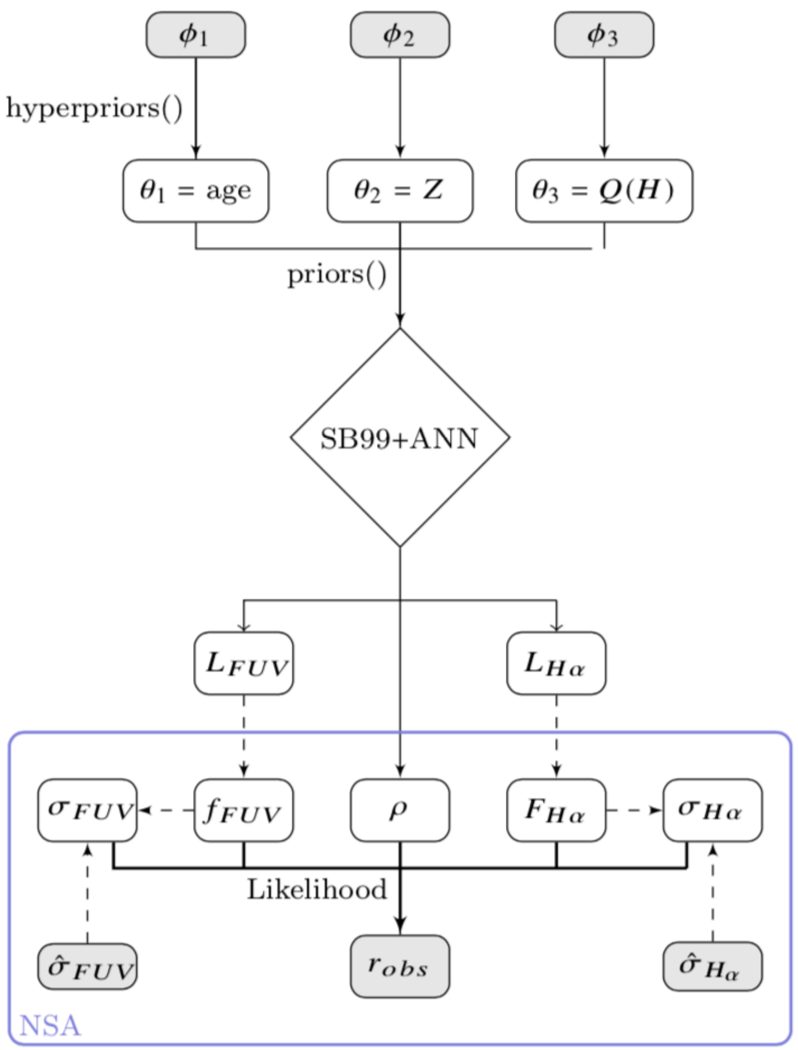


Figure 5. The hierarchical Bayesian model in plate notation. Blank shaded nodes represent random fixed (or observed) values, respectively. Arrows represent the kind of relationship between variables: probabilistic (solid arrows) and deterministic (dashed arrows). In the central part of the graphical model we show deterministic calculations done by SB99 scripts and the artificial neural network (ANN). The blue square at the bottom, encircles the variables involved in the Likelihood (7), as well as the observed ratio which are the input for the NSA.

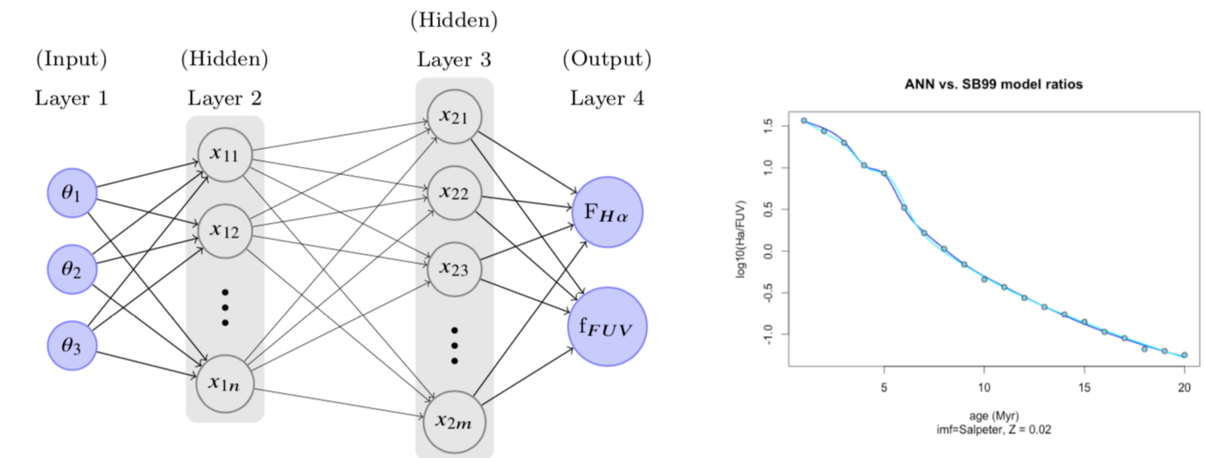


Figure B1. *On the left:* Multilayer Perceptron for the estimation of the Hα and FUV luminosities from the SED given by SB99 (output layer), depending on the input parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (age, IMF, Z)$. The ANN was developed with two hidden layers with n=20 and m=50 nodes respectively, and it was trained with 70 % and validate with 30 % of the grid points. The uncertainties due this interpolation were consider negligible. *On the right:* Comparison of the ANN (interpolation for the Hα to FUV flux ratio given different parameters than in SB99. Blue line corresponds to the ratios from the ANN given straightforward the model ratios, whereas the cyan line when the Hα and FUV fluxes are calculated separately. They are in good agreement.

Algorithm 1: Nested Sampling Algorithm

Data: the observed flux ratio, $\hat{r} = \hat{F}_{H\alpha}/\hat{F}_{FUV}$,
of a certain pixel of the ratio image.
Result: evidence $p(\hat{r})$ calculation, and the nested sequence of params. $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$
initialization;
$x_0 \leftarrow 1$
$z_0 \leftarrow 0$ ▹ z is the evidence, i.e. $p(\hat{r})$
$j \leftarrow 0$
repeat
Draw N objects $\boldsymbol{\theta}_1^{(j)}, \ldots, \boldsymbol{\theta}_N^{(j)}$ from prior (1)
for $i = 1, \ldots, N$ **do**
Use $\boldsymbol{\theta}_i^{(j)}$ to generate the model params.
$\{F_{H\alpha}, f_{FUV}, \sigma_{H\alpha}, \sigma_{FUV}, \rho\}_i$
with SB99+ANN, Eqs. (2) to (5)
Compute the likelihood $L_i^{(j)} \leftarrow$ Eq. (6)
end
$j \leftarrow j + 1$
Record $L_j = \min_{1 \le i \le N}\{L_i^{(j-1)}\}$
$x_j \leftarrow e^{-j/N}$ ▹ *crude estimation*
$w_j \leftarrow \frac{1}{2}\left(e^{-(j-1)/N} - e^{-(j+1)/N}\right)$
$z_j \leftarrow z_{j-1} + L_j w_j$ ▹ *approx. by Eq.* (B4)
until $L_{max} x_j < f z_j$ *&* $j < M$;
▹ terminate the main loop when the largest current likelihood would not increase the current evidence by more than small fraction f

INFERENCIA BAYESIANA

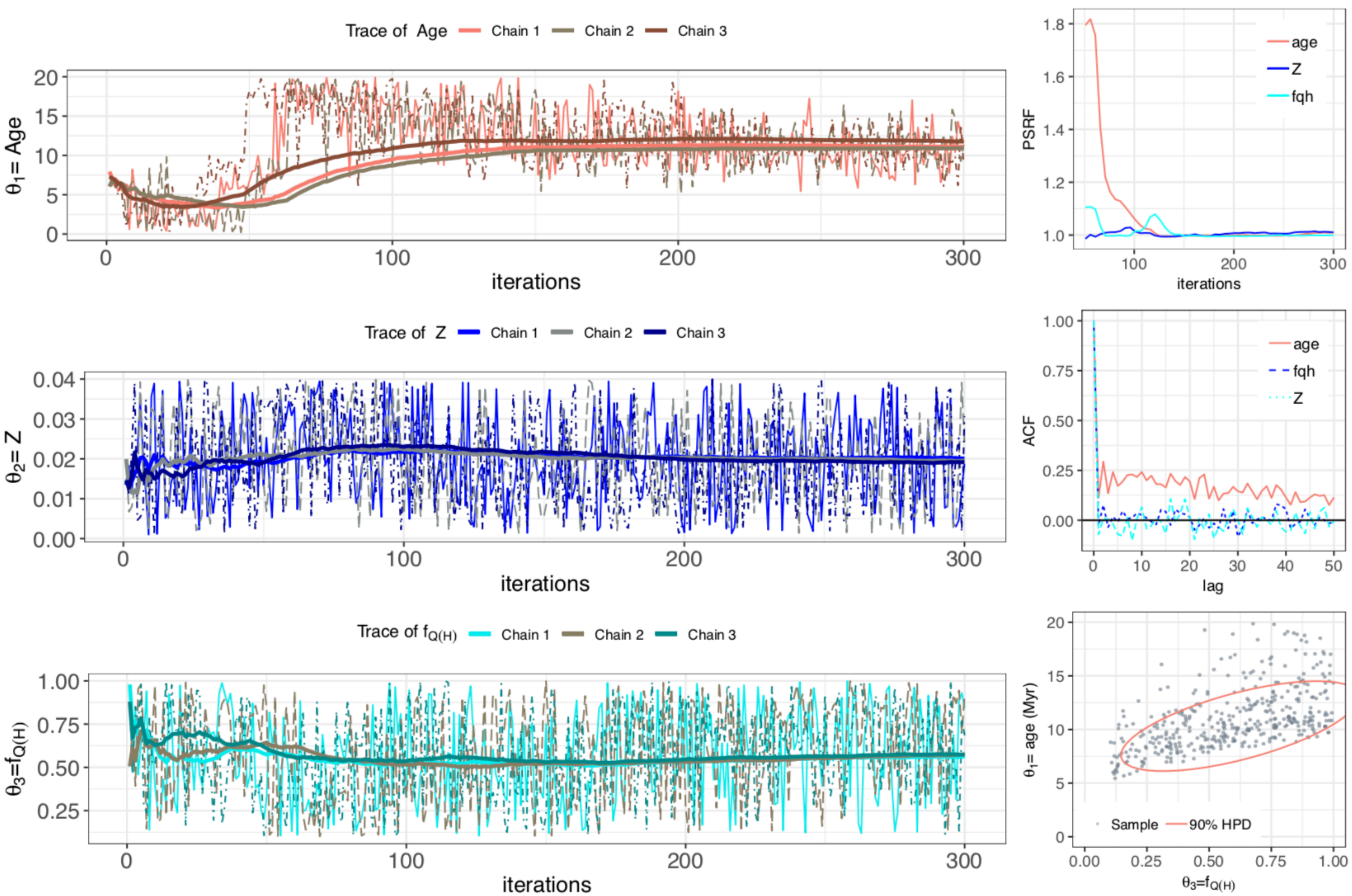
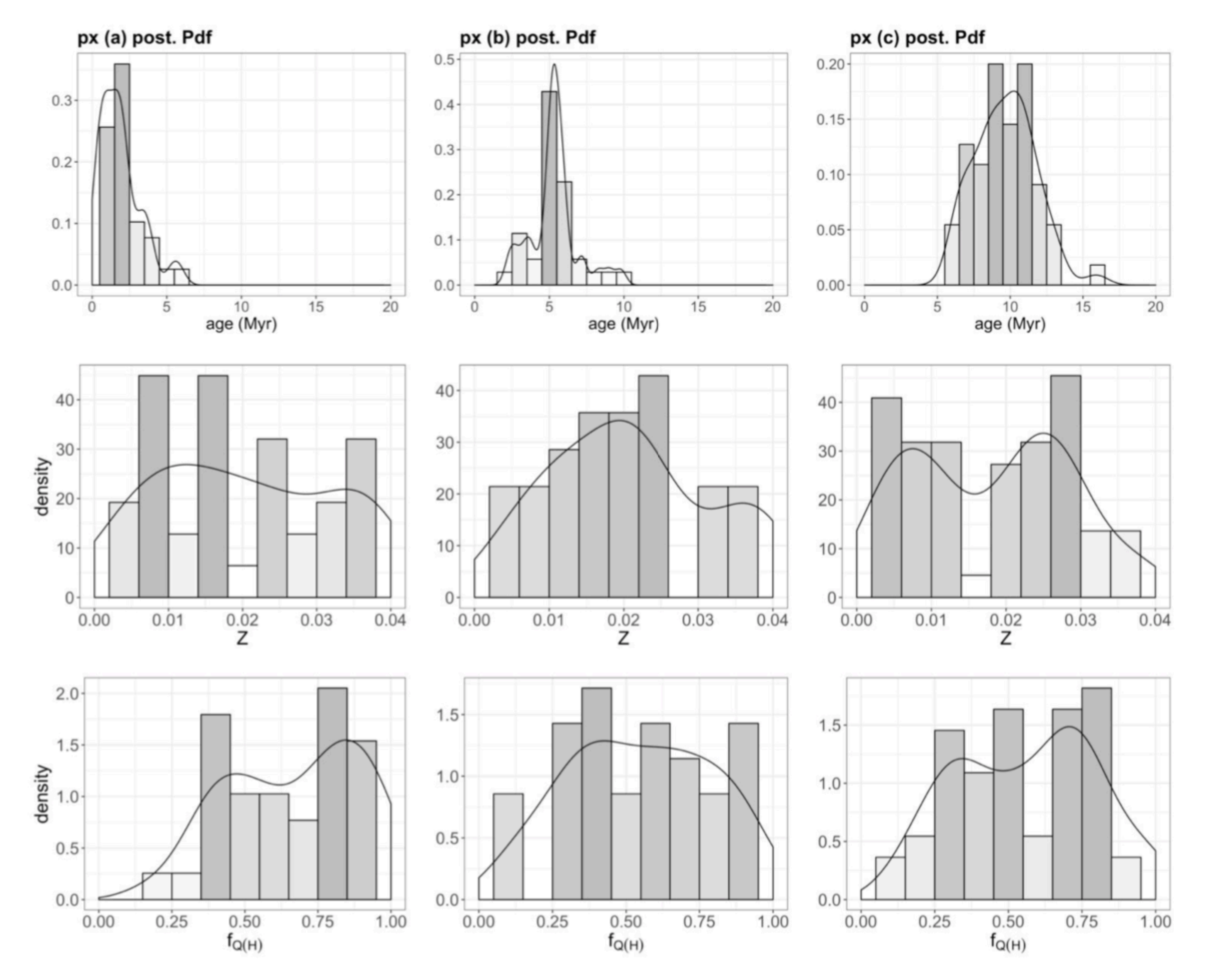
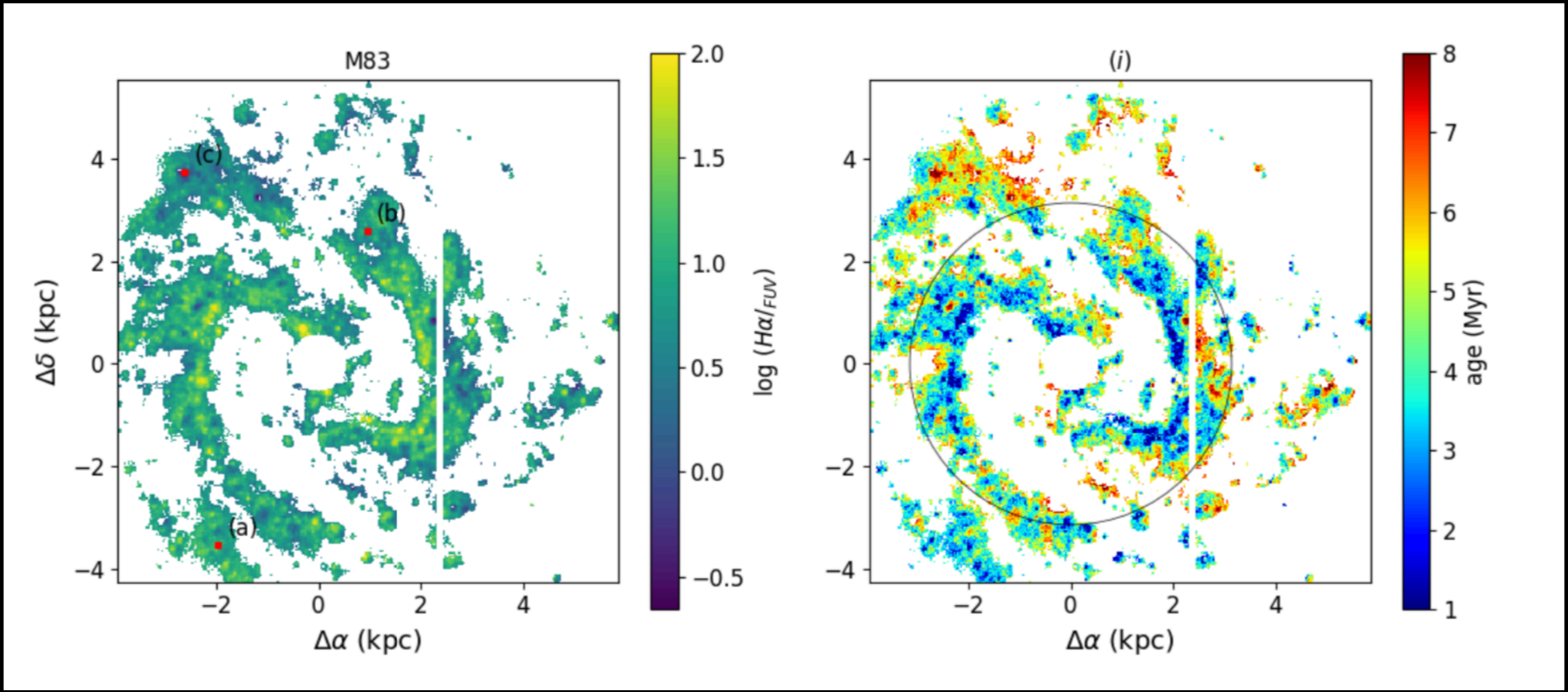


Figure B2. Assessing mixing and convergence. *On the left:* Trends of the estimated parameters, for three different chains obtained from the NSA algorithm application. After the first 100 iterations, all the chains are mixed and stable. *On the right:* On the top, the Potential Scale Reduction Factor (PSRF) for the three parameters, always lower than 1.1, indicates a fast convergence of the algorithm. At the middle a plot with the level of autocorrelation in each property, up to a maximum number of 50 lags. At the bottom, an example of the convergence towards the target distribution (red line indicates the 90% Highest Posterior Density interval). After a burn-in period of 100 iterations, the chain rapidly converges towards a specific region of the space of parameters in regard of the age, according to the given Hα to FUV flux ratio. Whereas it remains spread on the $f_{Q(H)}$, as it is expected and already shown in the bi-dimensional posteriors distributions at Figure 7.

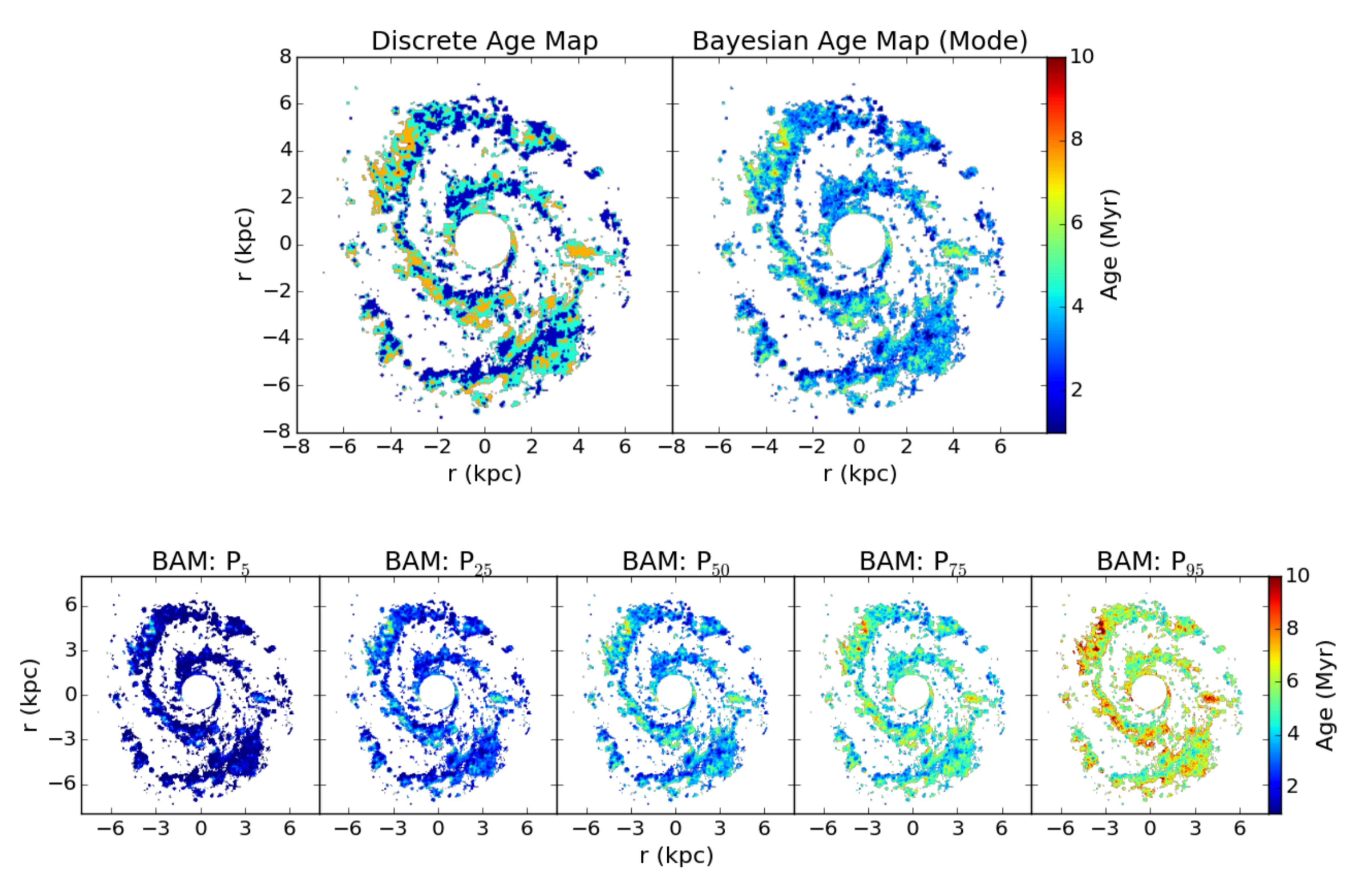
INFERENCIA BAYESIANA



INFERENCIA BAYESIANA



INFERENCIA BAYESIANA



INFERENCIA BAYESIANA