Context

The objective of this file is to provide a comprehensive summary of the doubts and current knowledge of Yllari González Koda (YGK, hereafter) in the context of full spectral fitting and, particularly, pyPipe3D and pyFIT3D. As such, python code will be avoided and, instead, a reference to file and line will be provided whenever necessary (e.g., `pyFIT3D/modelling/stellar.py 351`).

The main objectives will be:

- Obtaining precise star formation histories (SFHs) in the analysed galaxies, and do so in a rigurous manner
- Correctly calculate and interpret the associated uncertainties
- Full reproductibility of results in a user-friendly manner

Analysis on **SDSS** and **CAVITY** data.

Stellar population models

Stellar population models

Three posibilities are considered for this work:

Model	Pros	Cons
maStar	fully-consistent, empirical	variable lin. res. , $\sigma_{SDSS} \sim \sigma_{maStar}$
`gsd156`	not consistent	fixed lin. res., high resolution
MILES full	consistent	no young pop., high number of models

None of them is perfect for my purposes. The default star populations templates provided are the GRANADA + MILES (`gsd156`) files.

Questions:

- How should the maStar SPs be used for SDSS and CAVITY data? $\sigma_{\text{inst}} = 0$ for SDSS? Variable resolution can create many problems

Doubts pyFIT3D

Emission lines

- How come emission line fits can report negative integrated flux values when both the LM and RND method are bounded by the introduced limits?
- Even if large errors are given for the flux for a range of values, the program fits it either way (show example). The spectral analysis is performed with a weighted least squares, so one would expect for those values to be ignored.
- The coefficients for the stellar continuum fit are not the mean of the Monte Carlo iterations, as reported in the paper, a latter search is done with random noise (`modelling/stellar.py 3020`). Why? What are the implications?

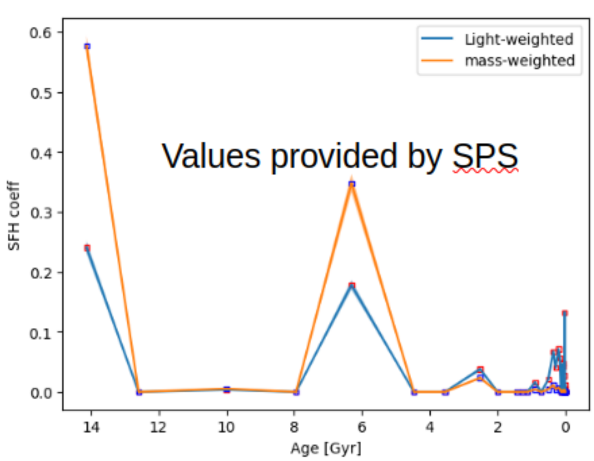
- In mass and light-to-mass calculation, no cosmology, or dust correction is taken into account, right? Also, the mass is calculated with $3500 \times M/L \times F$ (`modelling/stellar.py 3532`), where F is the median flux, M/L the mass to light ratio. Where does the 3500 factor come from? In pyPipe3D this factor does not appear (`common/tools.py 2667`).

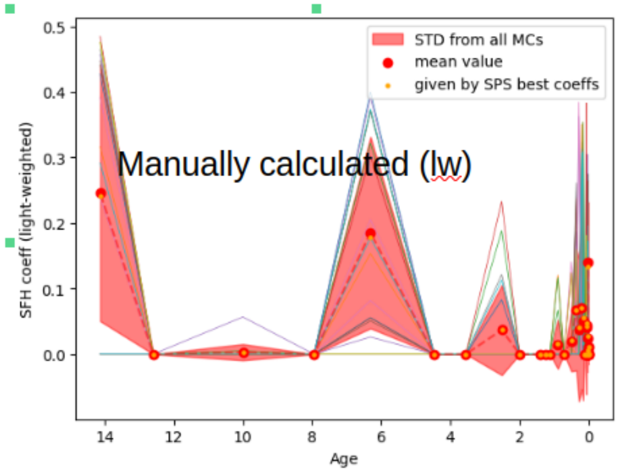
Uncertainties

Why only half of the flux error is used for the MC iteration? i.e, the added noise is gaussian-distributed with $\sigma(f_\lambda)/2$ (`modelling/gas.py 1189`). The full profile is used for the SPs fitting.

How come the error in the flux for the emission lines does not depend on the number of pixels in the fitting? Particularly in the continuum. This is from the formula $\sigma = \sqrt{\sigma_{MC}^2 + \sigma_c^2}$ (`modelling/stellar.py`) where $\sigma_c = FWHM \cdot \sigma_{\rm cont}$ and $\sigma_{\rm cont}$ is the standard deviation of the residuals for gas-only spectrum

What is the best consistent way to calculate SFH uncertainties? Right now, even if in the article it is reported that the given value for the SP coefficient χ_j is $\sigma(\chi_i)$, the uncertainty given by the program is $\sigma(\chi_i)/N_{\rm MC}$ (`modelling/stellar.py 3085`), which obviously underestimates the real uncertainty.





Then, either the $1/N_{\rm MC}$ factor is right and uncertainties do not make sense in the context of SFH calculation, because only information about how many MC iterations you are doing is being reported, or all derived uncertainties are ill-calculated. In the second case, the age and metallicity averages means and STDs could be computed directly from the MC iterations, but a linear propagation is done, which is always a worse approximation (`modelling/stellar.py 3013-3018`).

The uncertainty in the extinction is calculated twice, first in the non-linear fit, and then as $0.43 \sum \sigma(\chi_i) A_v^2$. Where does this formula come from? (`modelling/stellar.py 3018`)

Doubts pyPipe3D

- Have there been any changes in the pipeline since Lacerda+2022?
- What is the proper way to determine the center of the galaxy (x_0, y_0)? Just using the pseudo-V band mean flux as in the code?
- Why does the instrumental dispersion change between the different steps of the procedure? What sigma should I use in each one?

- What to do with extremely large (bad fitting) values for EL fitting? These are probably being filtered directly by Sebastián (show cube). These appear because the fitting wavelength range given by the config and the one provided or derived in the non-linear fit are different **This I probably could fix on my own**
- The instrumental sigmas are not being corrected for redshift, right? See Cappellari 2017 MNRAS 466, 798–811

The dust-corrected stellar mass density is calculated as:

$$\log(M_{\text{dust}}) = \log(M) + \log_{10}(e)A_v \tag{1}$$

but in the formula the preceding factor for A_v should be 0.4, from the correction $F_0 = F \cdot 10^{0.4A_v}$

Uncertainties

Since all the derived values come from pyFIT3D, not many doubts here:

- Why are there no uncertainties in the dust-corrected mass? Neither they are given in CAVITY data sometimes.

What else?

High order dataproducts? Comment about optimization? Comment about SFHs comparison with different SPs problems?