GalaPy, a really fast Python API for modelling galaxy SEDs and other cool stuff

Tommaso Ronconi

@INAF USCVIII - Calcolo Critico, June 2023

Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

tronconi@sissa.it

Collaborators:
Andrea Lapi
Martina Torsello
Alessandro Bressan
+others!

Why a new SED-fitting library?

SUBMITTED!

Soon public on GitHub and PyPI

New modelling

- **Analytical solutions for Galaxy Evolution**:
  realistic modelling of dust absorption/emission
  Gas, Stars, Metals and Dust
  (Lapi et al., 2018; 2020, Pantoni et al., 2019)
- **2-components dust model**:
  as seen in GRASIL (Silva et al., 1998)
  but AGE DEPENDENT ENERGY BALANCE
  (no radiative transfer + physical temperature)

Performance

- **Hybrid C++/CPython** implementation (performance + user friendly)
- First release: **shared memory parallelization** from Python

Bayesian framework

First release:

- **Markov Chain Monte Carlo** with `emcee` Foreman-Mackey et al., 2013
- **Dynamic Nested Sampling** with `dynesty` Speagle, 2019

GalaPy modelling

“In-Situ” Star Formation History

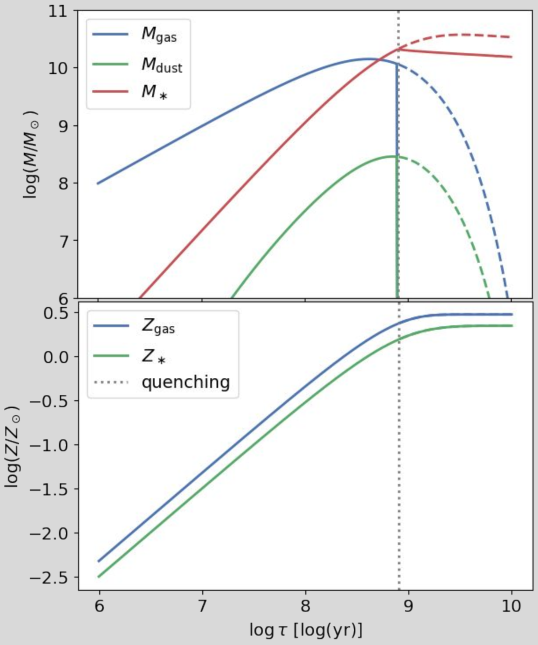
from Lapi et al., 2020 (+ Pantoni et al., 2019)

Analytical derivation:

- ➔ self-consistent
- ➔ age-dependent
- ➔ reduce the parameter space

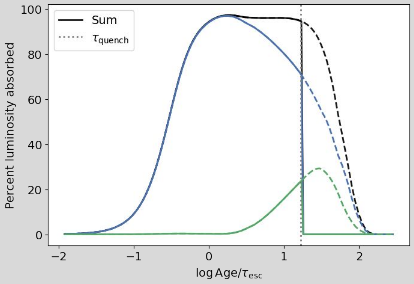
Component Masses

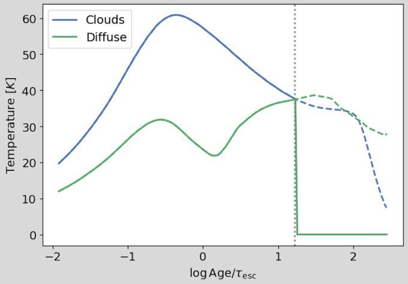
Component Metallicities



Molecular Clouds + Diffuse Dust

ENERGY CONSERVATION (age dependent) ⟹ Temperature of the **Grey-Body**

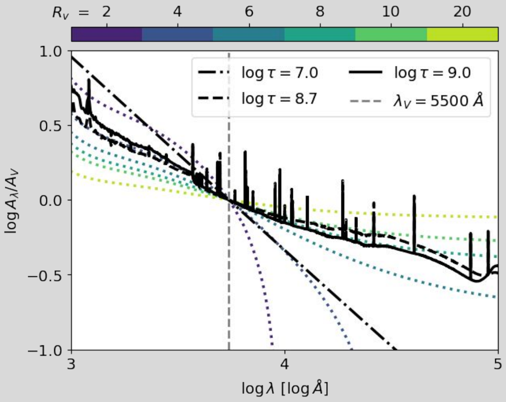




Combined attenuation depends on both:

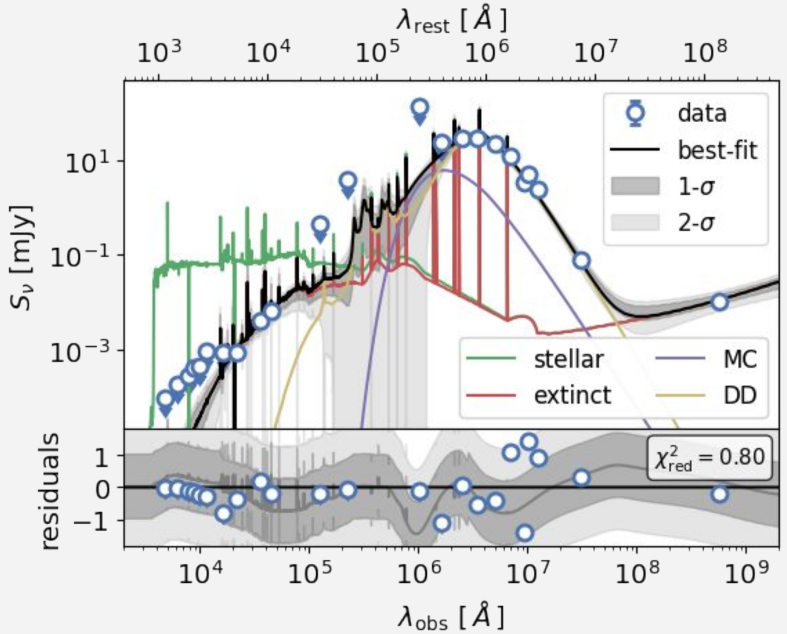
- ➔ wavelength
- ➔ **age of the gxy**

no prescription on grain physics



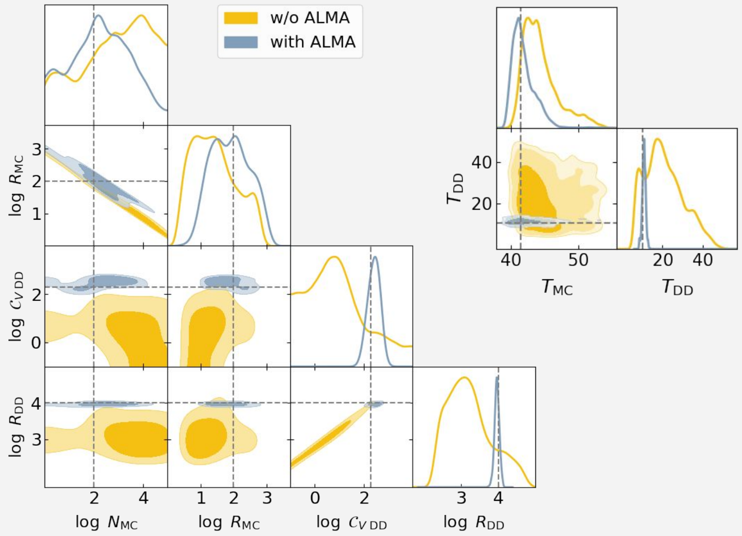
Some results: fitting

- **SED fitting:**
  - panchromatic photometric data
  - includes upper limits
  - large library of models available



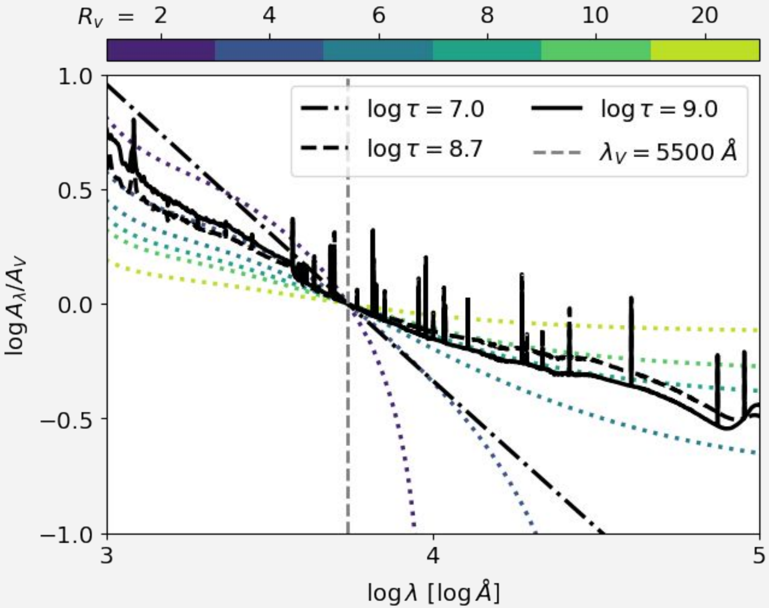
Data from Giulietti et al., 2022

- **Bayesian approach:**
  - large physical parameter space
  - systematic errors accounted
  - derived parameters parameter space sampled

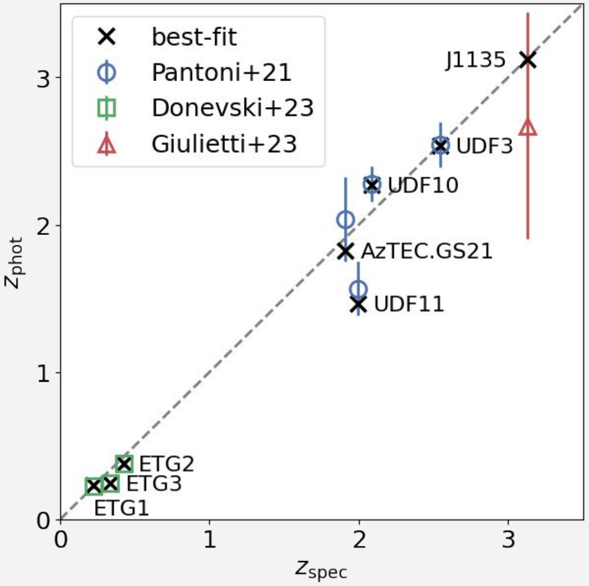


Some results: derived quantities

- **ISM & Dust properties:**
  - metal abundances
  - gas and dust masses
  - **attenuation curves:**



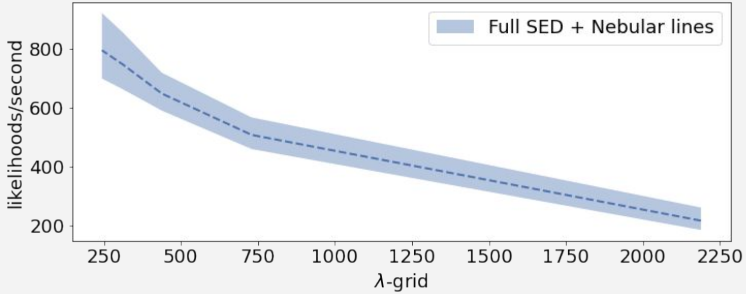
- **Photometric redshift:**

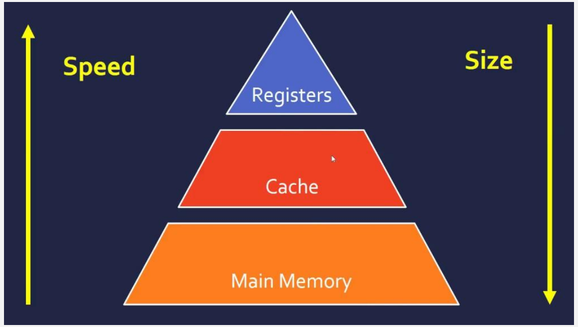


... could use some extension
(i.e. larger samples, higher redshifts)

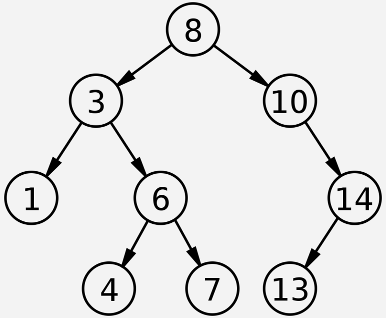
(some of the) Algorithms used

- Lag in interfacing the 2 languages
  - C++ → Pybind11 → Python
- Performance
  - Optimised Object-oriented implementation
  - Beating main competitors
    - CIGALE /100
    - Prospector /10



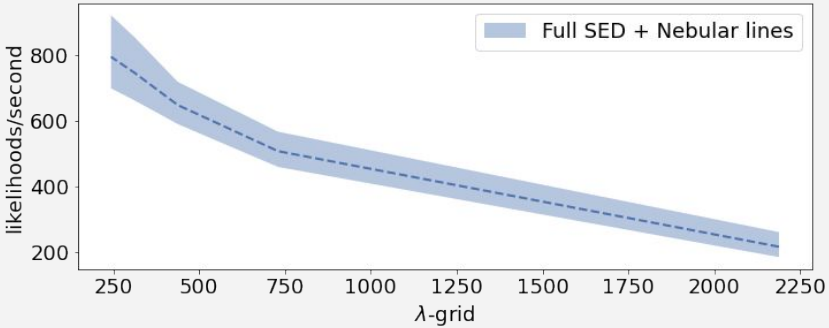


- Architecture-oriented **caching** of templates and globals
- Total number of **operations reduced** to the bare minimum
- Optimisation of **serial model computation**
  - Shared memory **parallelism for sampling**
  - Distributed memory **parallelism for scaling** on large catalogues



- **Interpolation** on a non uniform and non-uniformly accessed grid
  - Binary-interval search tree
  - Interpolator objects
    (c++/python, no external library dependence but **faster than scipy and gsl**)

Minimal performance measurements and comparison



Mainly dependent on **spectra resolution**

- low res: up to **800-900 likelihood/sec**
- high res: up to **200-250 likelihood/sec**

Comparison with other codes is not trivial.

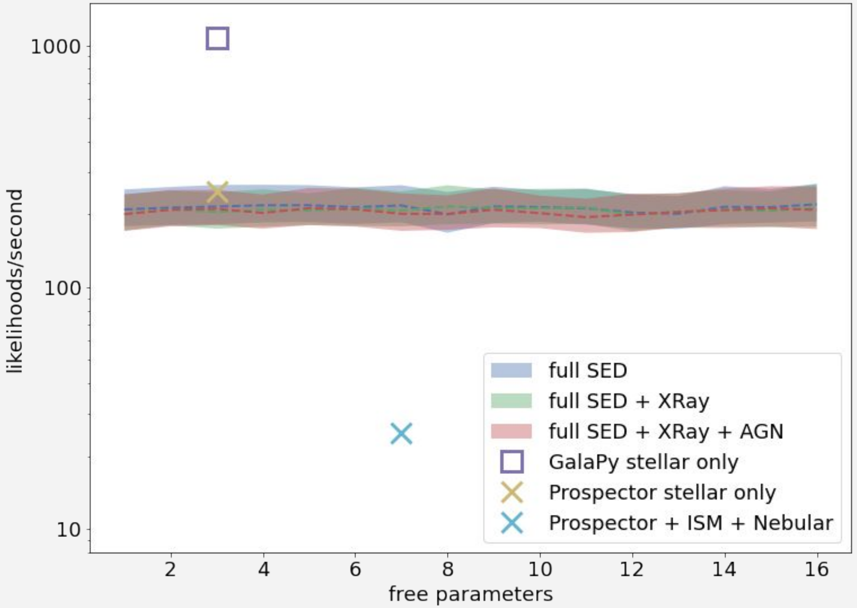
CIGALE: Boquien et al., 2019

- different statistical framework
- **GalaPy** model generation:
  ∼ 100 times faster

Prospector: Johnson et al., 2021

- Bayesian (MCMC & nested sampling)
- **GalaPy** likelihood more than
  ≳ 10 times faster

maybe **MAGPHYS+photo-z:**
Da Cunha et al., 2008 Battisti et al., 2019



Algorithms and skills acquired

Present

- serial part highly-optimised
- **interval-binary search tree**
- **register-friendly fine time-grid integration**

- specialisation on the problem size
- no architecture-specific optimisation
- parallelisation through Python

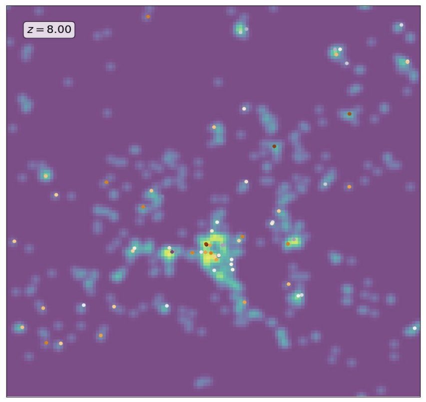
Future

- **Hierarchical Bayesian Sampling**
- **Hamiltonian Sampling**
- **Direct posteriors from ML**

International context

- **ALMA**
  really important for the **dust re-radiation** and to **disentangle photo-z degeneracies**
- **JWST**
  highest redshift sources
- **SKAO (& precursors)**
  e.g. Evolutionary Map of the Universe (EMU)
  - stellar un-attenuated
  - AGN contribute

Some more devs from SISSA-APC



[Ronconi et al., 2020]

- Empirical modelling of the Galaxy/Halo connection
  - **packaged** easy-to-use interface for SCAM
  - some w.i.p. devs include **AI-assisted modelling**

SKAO ➔ radio continuum+HI foreground

- Visibility-plane stacking of ALMA observations
  - **aim:** detect extremely faint molecular lines
  - **target:** high-z DSFGs

Martina Torsello in the audience, talk with her!

... and this is my last slide, thanx!