ML algorithm for automatic determination of a dynamical state of GCs based on CMC-SURVEY

Arkadiusz Hypki, Valletta, ML4Astro3

2026.09.03

[1] Faculty of Mathematics and Computer Science of Adam Mickiewicz University

Outline

Globular clusters

CMC code

CMC-SURVEY

CMC-ML

Globular clusters

Globular clusters

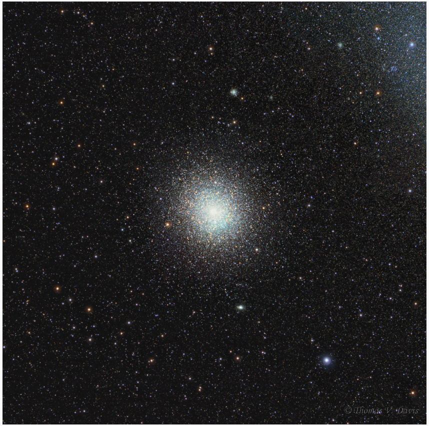

Figure 1: 47Tuc globular star cluster, one of the biggest and oldest in the Milky Way.

- very old (age comparable to the age of the Universe)
- size up to around 100 ly
- a core is clearly visible – best place for creating of many exotic objects: cataclysmic variables, X-ray binaries, black holes, intermediate-mass black holes, blue stragglers interactions between stars
- Milky Way GCs: 50% GC within 5 kpc, the most distant 130 kpc
- **great laboratories for studying stellar evolution and dynamical**
    - $\rightarrow$ **we need to include as many physical processes as possible**

Dynamical modelling – importance

- may provide basic information to understand the **formation and then the evolution of exotic objects** within star clusters (e.g. hard binaries)
- dynamical interactions between stars may lead to **perturbations, disruptions, collisions and mass transfers** between stars
  - e.g. this may lead to decrease the semi-major axes and allow mergers which would now happen otherwise
  - may lead to formation of exotic binaries, supernova explosions (especially in the initial phase when many of massive stars are present), formation of black holes
- dynamical interactions in GCs may eject a lot of binaries that could be potential **sources of GWs**

CMC code

CMC – features

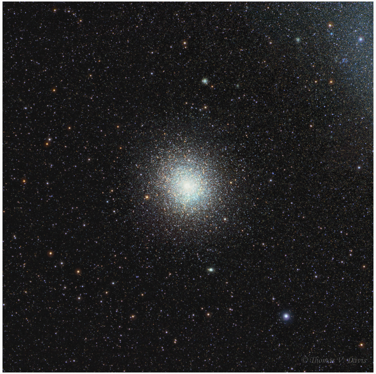

Figure 2: CMC code, `https://clustermontecarlo.github.io/CMC-COSMIC/`

- one of the most advanced codes for simulations of real-size star clusters
- based on Monte Carlo method (a few simplifications in comparison to N-body codes, e.g. one radial position)
- agrees very well with N-body codes
- provides almost as much details about stars as N-body codes
- model exotic objects: BH-BH, blue stragglers, IMBHs, CVs...
- provides “observations” of simulations vs. real observations
- very fast, which allows to test **whole range of possible initial conditions** (CMC-SURVEY)
- data analysis with BEANS

CMC-SURVEY

CMC-SURVEY

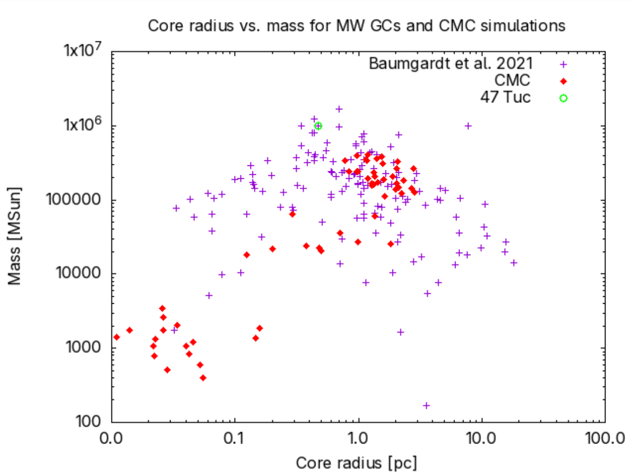

Figure 3: Current Milky Way GCs coverage by CMC-SURVEY

CMC-SURVEY – currently over 210 CMC models for various initial conditions:

- N: from 40k up to 1M
- binary fraction: from 0.0 to 0.95
- different r_{vir}, r_{tid}, metallicities
- very soon freely available through `https://beans.hypki.net` – **public, open-access, fully downloadable, extensive database of detailed globular cluster models**

CMC-ML

CMC-ML

The initial conditions for GCs are highly degenerated, we don't know their parameters.

$\Rightarrow$ **The goal is to narrow down (and some day predict) the initial conditions for all Milky Way GCs based on trained ML model**

Predicting initial GC mass based on r_c, r_h, m_{tot}, f_b from all CMC simulations

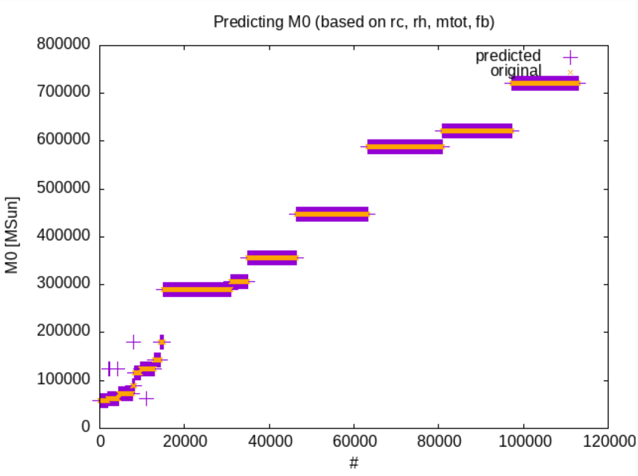

Figure 4: Predicting initial GC mass (scikit-learn, XGBClassifier). Similarly we can get initial densities, number of stars, binary fractions etc. **The predictions are remarkable**.

Next steps

Figure 5: Work in progress

Next steps:

- extend CMC-SURVEY
- fine-tune ML parameters
- apply for real MW GCs

⇒ **Check soon** `http://beans.hypki.net` **for public, open-access, fully downloadable, extensive database of detailed globular cluster models made with CMC code**