Report structural plan:

1. Introduction

* Context and aims
* Previous studies
* Open clusters and their importance

1. Data and target selection

* Focus on a range of cluster with different ages, Fe/H and helium fraction that have been relatively extensively studied
* Range = 1Gyr – 12Gyr
* Describe data collection process: membership -> get Teff (describe sources, appendix for table of all stars and method of Teff selection) and Gband -> membership cut (>99%) -> luminosity calculation -> RGB cut -> star removal (blue stragglers and binaries and other weird stars)
  + Luminosity calculation involves: Distances from Jones determined from Gaia parallaxes, Green’s dustmap (Bayestar), Gband extinction coefficient from redenning assuming a 4500K star from xx, BC from Gband to Mbol, with Teff dependence fitted by xx, ignoring redenning
* Note on individual clusters and why they are interesting

1. Stellar models

* Used MESA code (site paper)
* Tanda’s description about the MESA settings
* We have not considered rotating stars
* Initial conditions in a grid of:
  + initialMass = [0.8, 1.7, 0.04]
  + Age = 1 to 12 Gyr
  + initialFe/H = [-0.8, 0.6, 0.2]
  + initialY = [0.24, 0.32, 0.02]
  + MLT = [1.7, 2.3, 0.2]
  + Terminated at beginning of core helium burning
  + In later parts of the study, further cut off RGB stars with equation: 10\*log10(Teff)-log10(luminosity)-35.5>0, take stars that are within range.
* Address drawbacks of using the grid: large Feh, Y and MLT steps…

1. Part A: Machine learning on stellar evolution – Neural networks
   1. Theory
   * Basic mechanisms of machine learning and NNs: current applications, example NN, the activation function, back propagation, training and validation, underfitting and overfitting
   * Do not go into too much detail!!! Pass explanation to other papers to save words (eg. back propagation in verma 2016)
   1. Past studies on the topic of using ML on stellar evolution
   * Bellinger 2016: random forest and backward modelling with asteroseismic and spectroscopic observables
   * Verma 2016: NN to model low mass MS stars, backward modelling with asterseismic observables
   * Hendriks & Aerts 2018: NN forward modelling with a genetic algorithm to backwards sample fundamentals. Uses actual mode frequencies as observable output. MS + subgiant + RGB, no rotational stars
   1. Training process
   * Nadam, monitor and train on MAE, activation function elu, bn after input
   * Training happen in legs, where incrementally increase batch size and learning rate -> explain/cite Adabatch -> show history/loss plots
   * Start by training on a subset of stellar models that are +-1step in Feh and Y around solar values, and increase architecture till overfitting/isochrone messiness happens while trying to minimize loss -> show isochrones
   * After architecture is set, introduce regularization -> explain how reg works
   * Tried dropout -> leads to dropping of low MS -> abandoned
   1. Overfitting in the RGB
   * Trying to train an “one size fits all” NN for MS, subgiant and RGB proved to be rather challenging
   * RGB suffer from servere overfitting across track while NN is able to smoothly interpolate the behaviour of stellar evolution between tracks in MS and subgiant -> show isochrones and in between track plots
   * Introducing and increasing regularization although did largely smooth out the RGB overfitting, the results is still less than satisfactory -> show more in betweens and isochrones
   * Attempts in training NN on RGB alone was unable to solve the problem
   * Propose possible reasons:
     + Data density in age space>> in any other space
     + A very large amount of NN flexibility is needed to fully approximate the bahaviour along tracks due to RGB bump, which calls for a large architecture, allowing massive overfitting from track to track
     + More training seems to worsen the overfitting due to the lack of in-between track datapoints
   1. Resulting best fit NN
   * Abandoned training on RGB
   * Time took to train it and spec of google
   * Final MAE: At a level of around 2x10^-3
   * Show HR, isochrone and other comparisons
   * show typical errors (error plots)
   * persistent problems:
     + overfitting in feh -> show final in between plots
     + overfitting along track -> show plotTrends plot
2. Part B: Hierarchical Bayesian Modelling for population-wide properties
   1. Theory
   * Describe HBMs and why it is so good -> pool together information from individual datapoints to give population-wide estimates, and being Bayesian at the same time
   1. Previous studies???
   2. Our HBM model
   * Show pgm
   * Explain hyper-priors: Beta function of literature value-step to literature value+step
   * Explain how NN serves as a conversion function from fundamentals to observables
   1. NN validation
   * Modelling the sun test
   * Modelling a grid-picked pseudo-cluster: recovery of fundamentals test
3. Results and discussion

* Overall estimated properties table
* All cluster fitted isochrone plots
* General time needed to run HBM sampling and bluebear specs
  1. Compare to other methods:
  + Accuracy
  + Ability to get age, feh and Y spread in the cluster
  1. Cluster by cluster discussion
     1. M67/NGC2682
     + Talk about results on Dstello RGB stars???
     + Compare to literature value determinations
     1. NGC6791
     + Compare to literature value determinations
     1. NGC6819
     + Compare to literature value determinations
  2. Some science implications in here!
  3. Possible improvemements to framework
     1. NN
     + Fix all problems mentioned in section 4
     + Look below
     1. HBM
     + Look below

1. Conclusions

Acknowledgements

Stuff we can talk about improvements on the process:

Better data for NN training:

1. Finer steps in feh and Y, to “use up” flexibility of the NN that is needed to approximate stellar evolution behavior along track
2. Hold up certain different initial condition tracks to be used only as evaluation of the performance of the NN, that is within the training boundaries, so the in between track behavior of the NN can be tested effectively
3. Potentially can just combine a bunch of random initial condition tracks that is off the tracks in the training grid, but still within the training boundaries, into the training grid, to provide much needed “in between tracks anchor points” to prevent overfitting

Improvements to HBM sampling:

1. Can incorporate the conversion between luminosity+Teff and G band mag+Gbp-Grp into the HBM to allow more stars as samples. Currently we are only relying on Teff determined by previous work, which a. uses varying methods and assumptions and b. is very few compared to the full number of available stars observed with gaia G band. It is possible to take the output luminosity and Teff of the NN in the HBM and calculate gaia G band absolute mag and Gbp-Grp absolute colour and use these as comparison points, allowing more stars and a standardized Teff->colour conversion. To take this even further, one can add in elements like extinction, distance and BC into HBM to turn absolute magnitudes to apparent magnitude and use apparent mags as comparison points. This allows to include extinction, distance and BC as free parameters in the HBM that can be estimated for at the same time. But this will also increase the complexity of the HBM and lengthen its computational time requirement.
2. A better trained NN without drastic overfitting almost always ensure faster HBM sampling
3. NNs with smaller architectures speeds up HBM sampling
4. A second layer of hierarchical thing can be put on top on a cluster to cluster level, by assuming a rather constant age, feh and Y spread across clusters (due to them being different versions of the same thing), can get typical behaviours of OCs in the entire galaxy given enough clusters, and the pooling of even more data, which gives even tighter age constraints