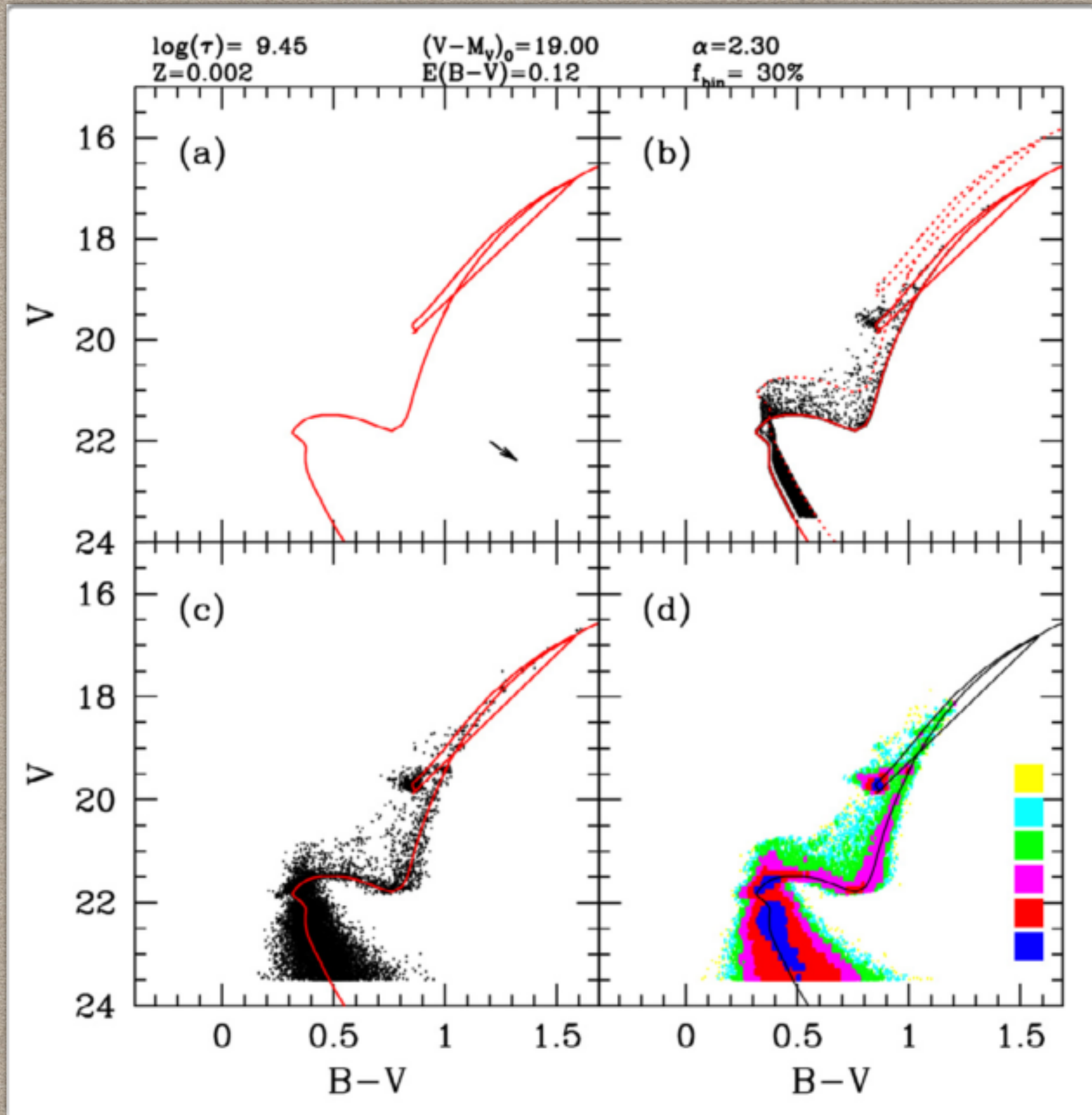
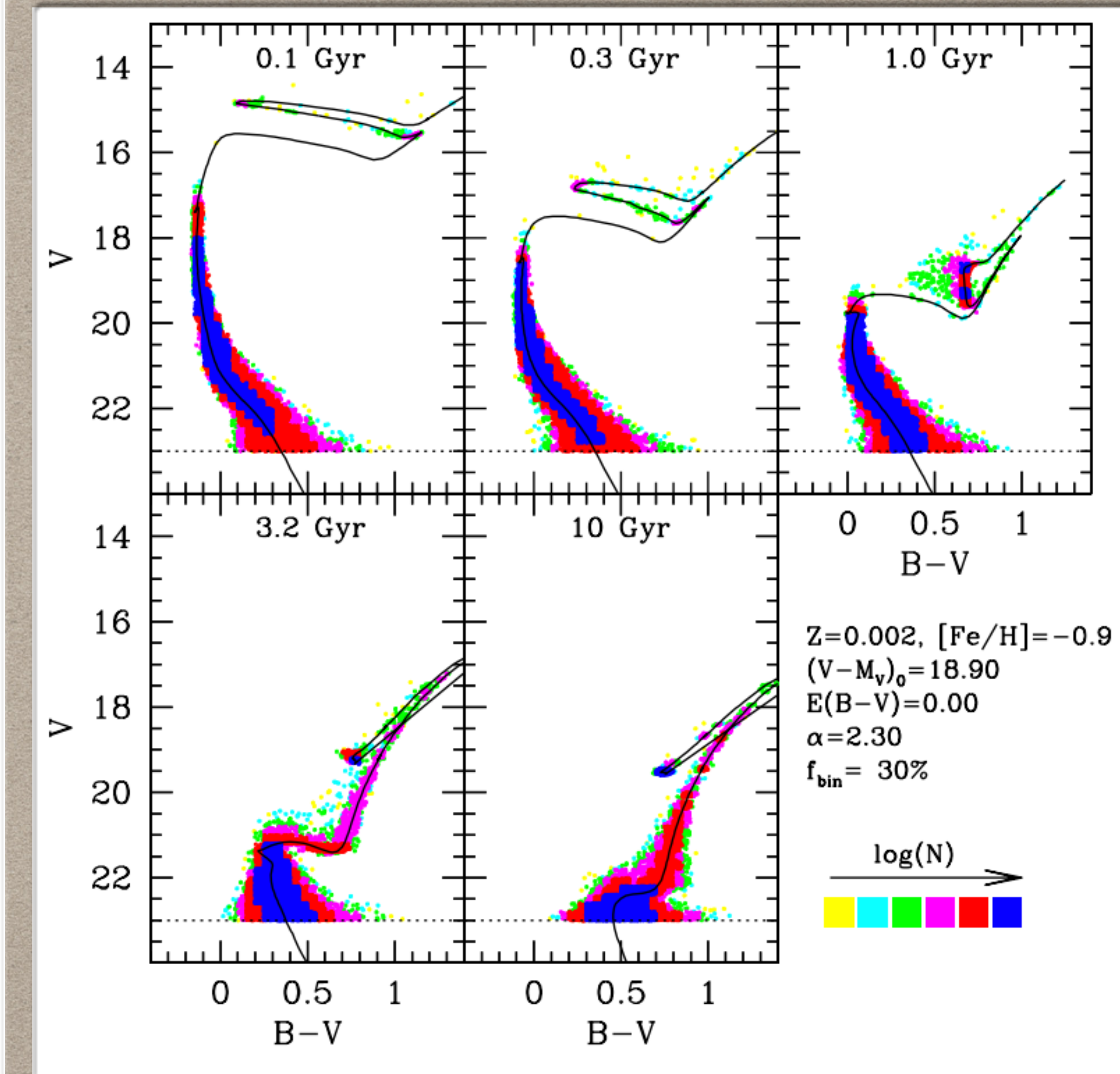




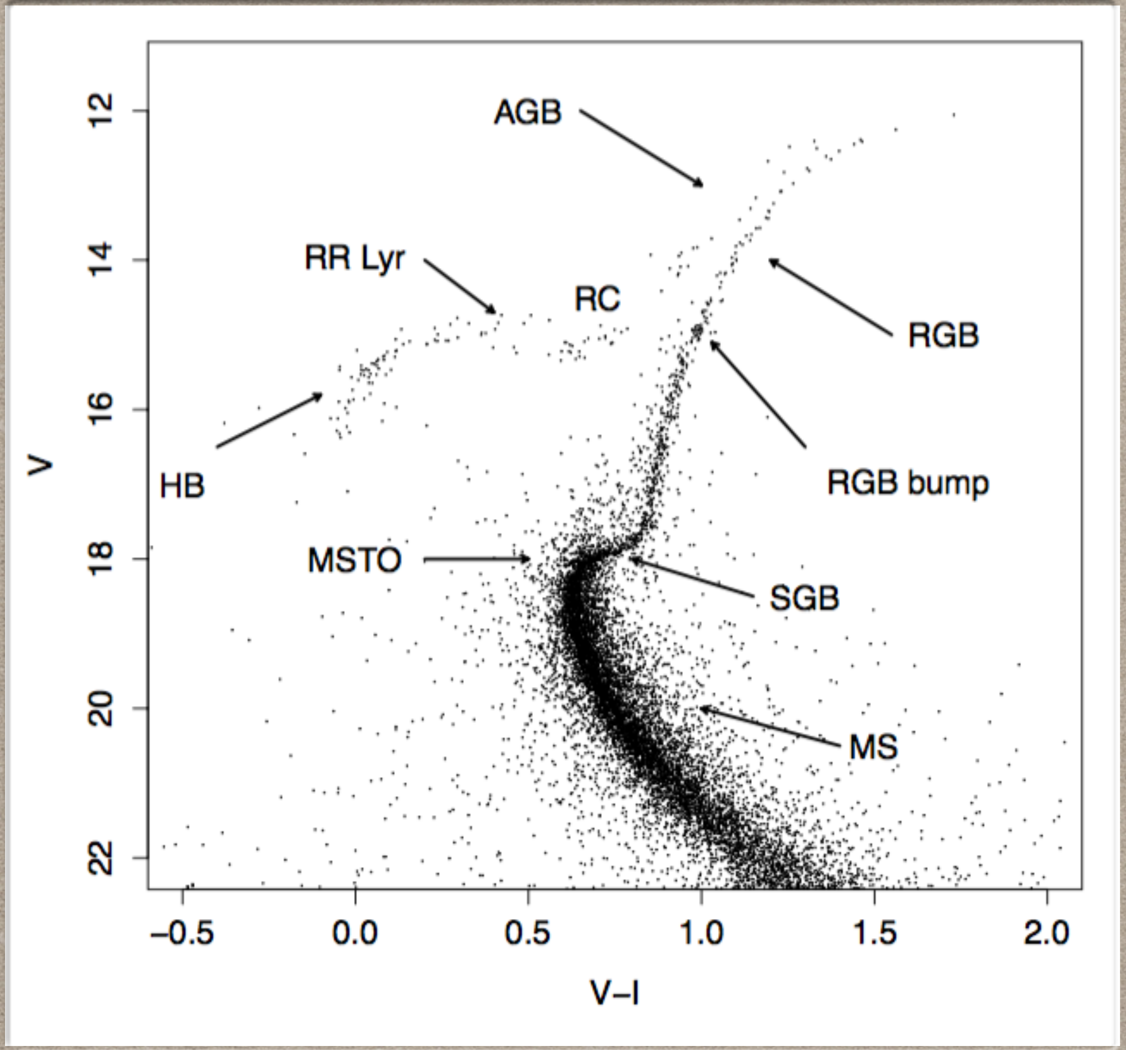
Dias et al. (2014,2016)

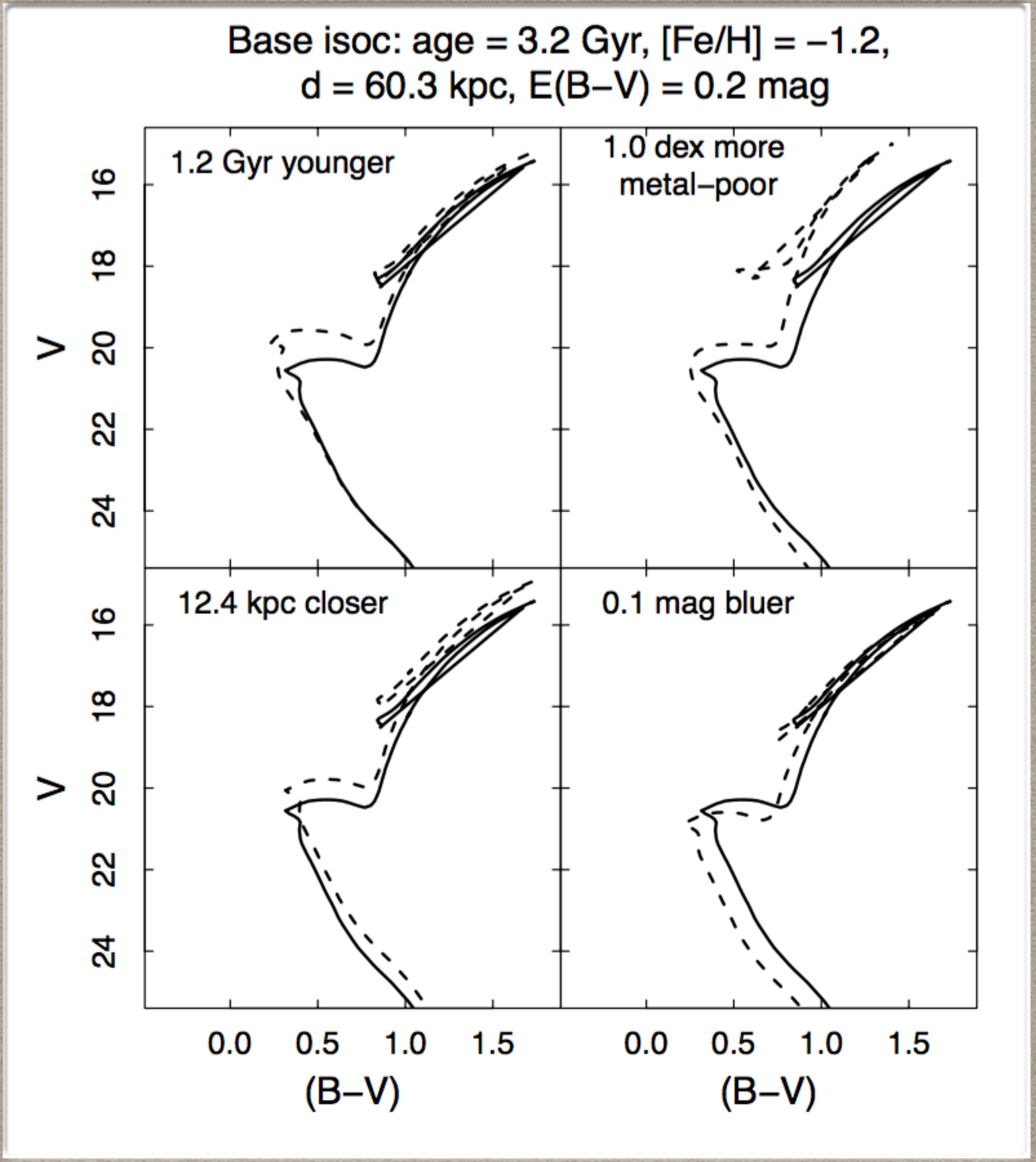
MAXIMUM LIKELIHOOD ESTIMATIONS I: SYNTHETIC COLOUR-MAGNITUDE DIAGRAM FITTING

BRUNO DIAS – MCMC COFFEE, ESO – OCT 5, 2016 (4.2-4.5)

THE CASE: FITTING BY EYE IS TRICKY

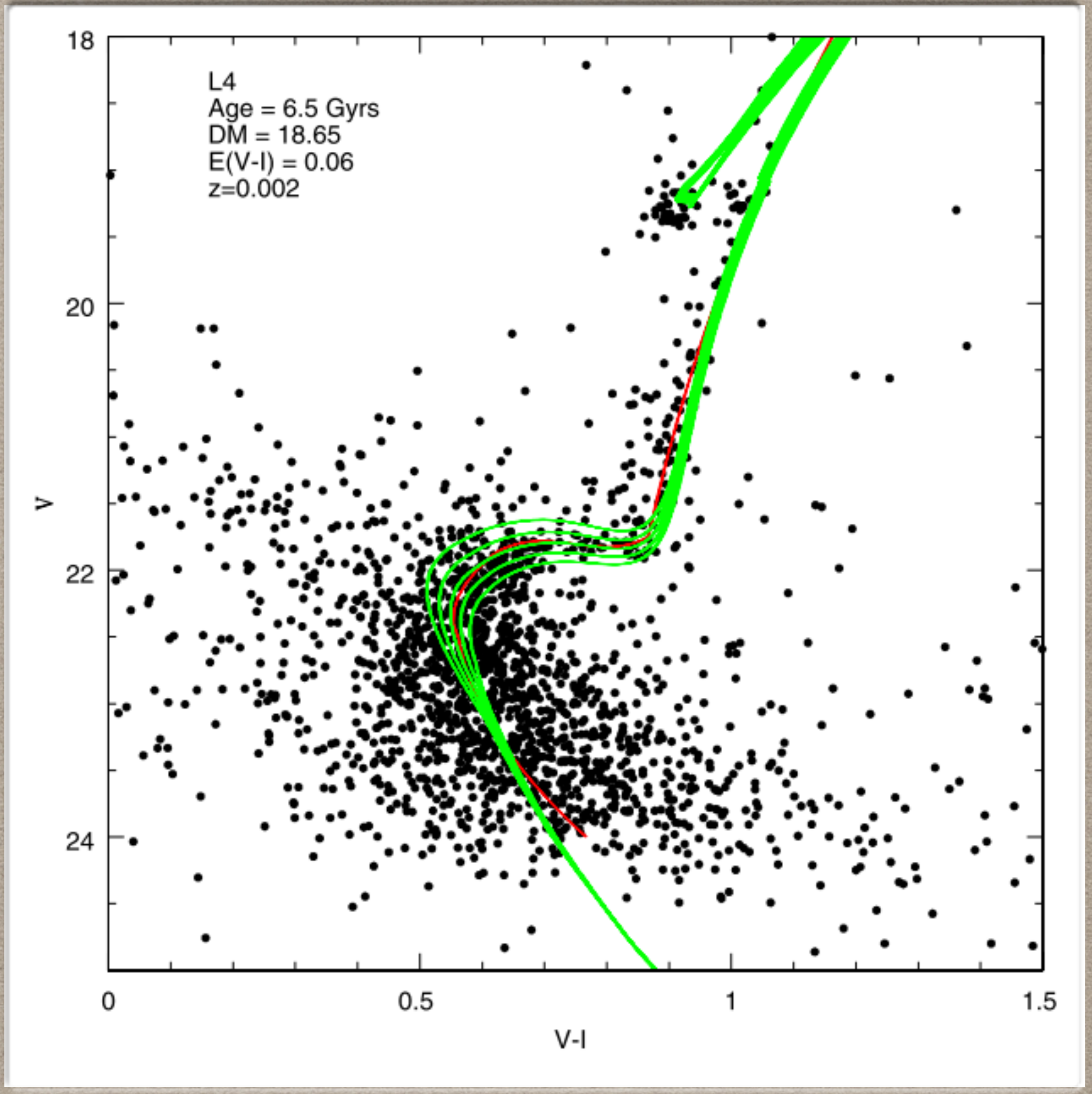
Colour-magnitude diagram = CMD





THE CASE: LINE VS. LINE

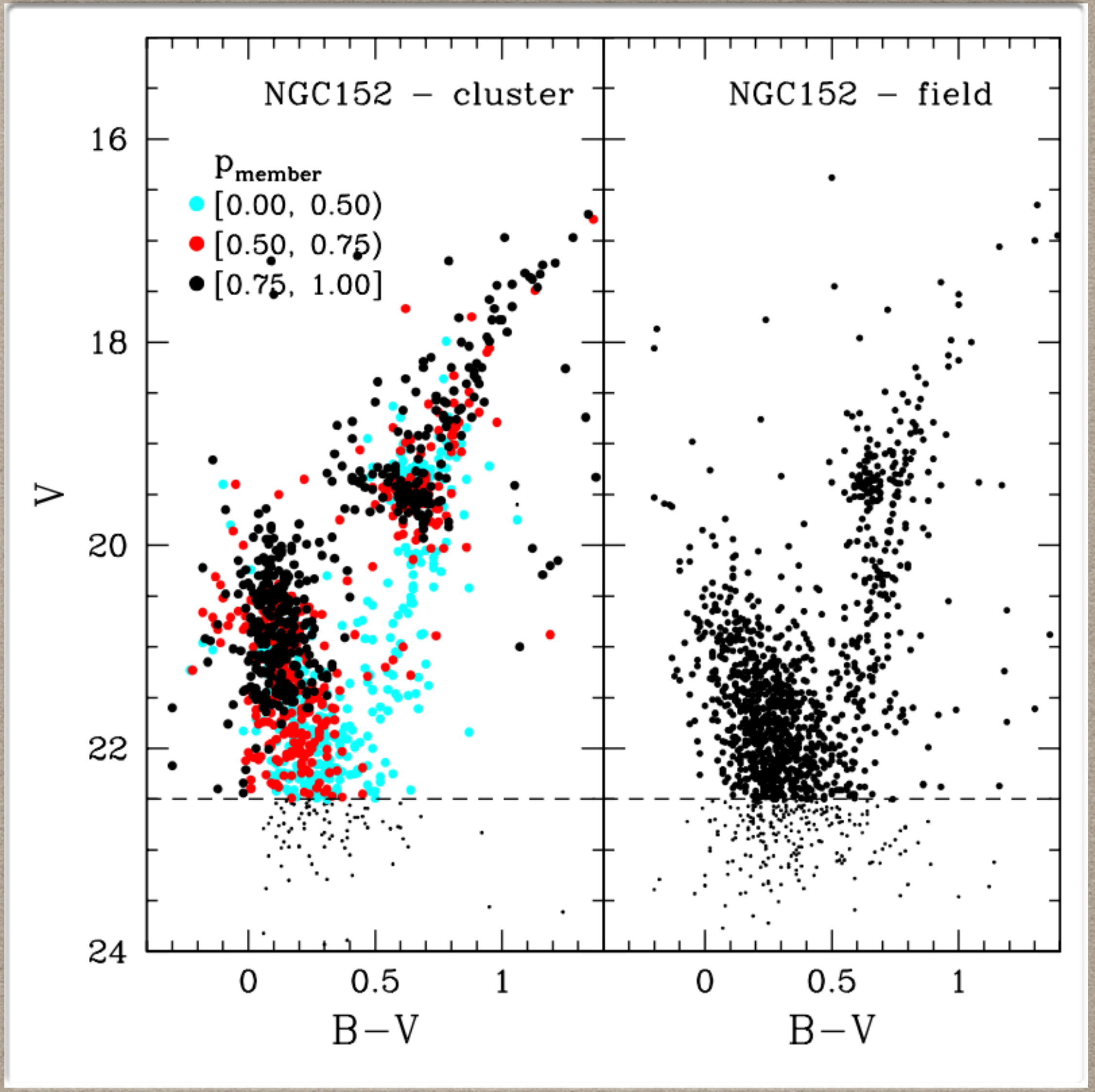
- Red: ridgeline
- Green: BaSTI isochrones (Pietrinferni et al. 2013)
- (m-M) = literature
  Δ = 0.2 ~ 10kpc
- [Fe/H] = literature
- E(V-I) = literature



Parisi et al. (2014)

THE SOLUTION: 1.CLEANING THE CMD

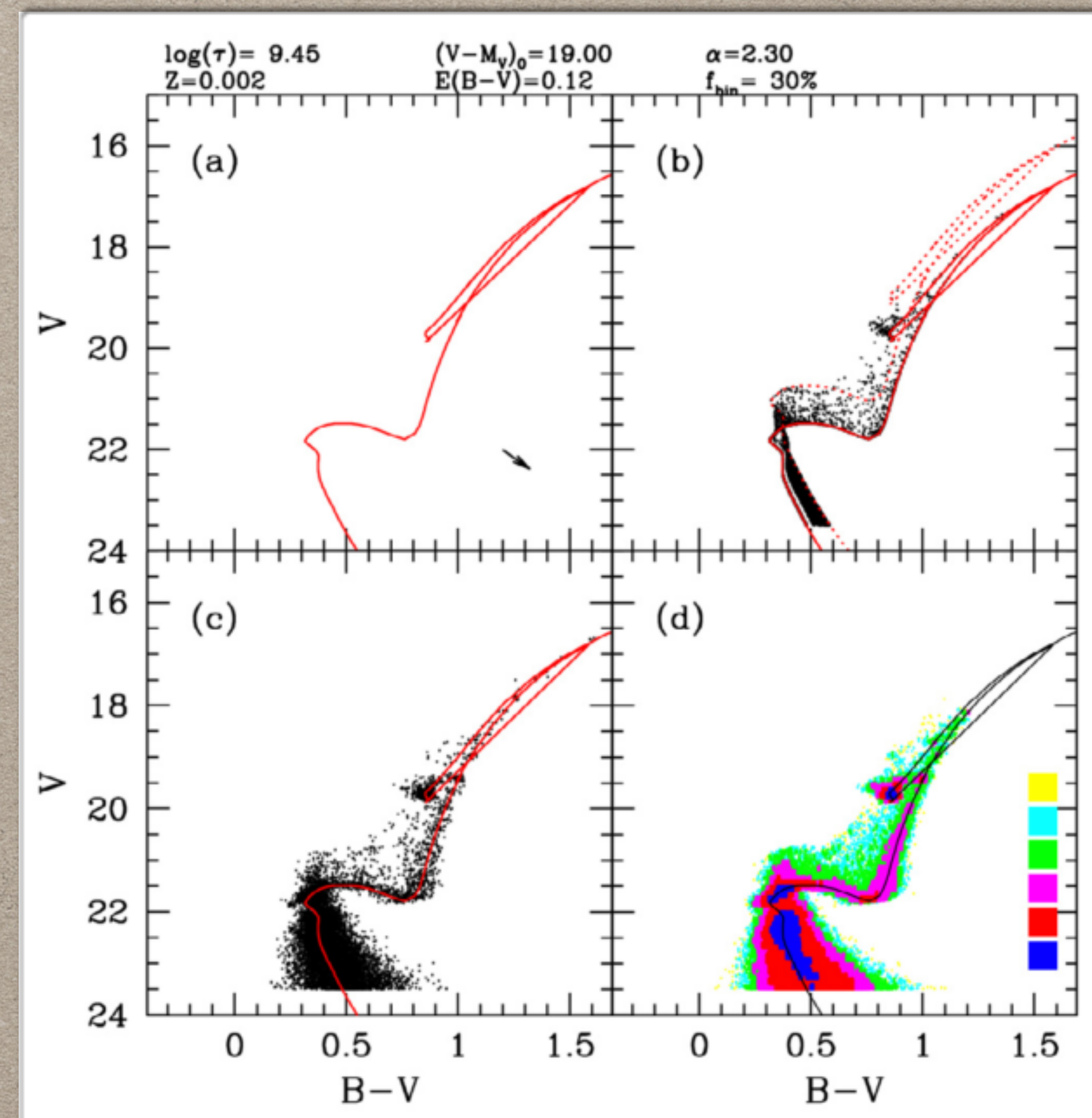
- Control field nearby
- Hypothesis: the positions of the off-cluster stars represent the most likely positions for field stars on any similar CMD.
- p_{member} = probability of a given (observed) star be member of a given cluster



Dias et al. (2016)

THE SOLUTION:
2. CREATING SYNTHETIC CMDs

- Isochrone:
  age, Z, E(B-V), (m-M)
- IMF
- binary fraction
- photometric errors
- photometric completeness
- magnitude range
- Hess diagram: weights!

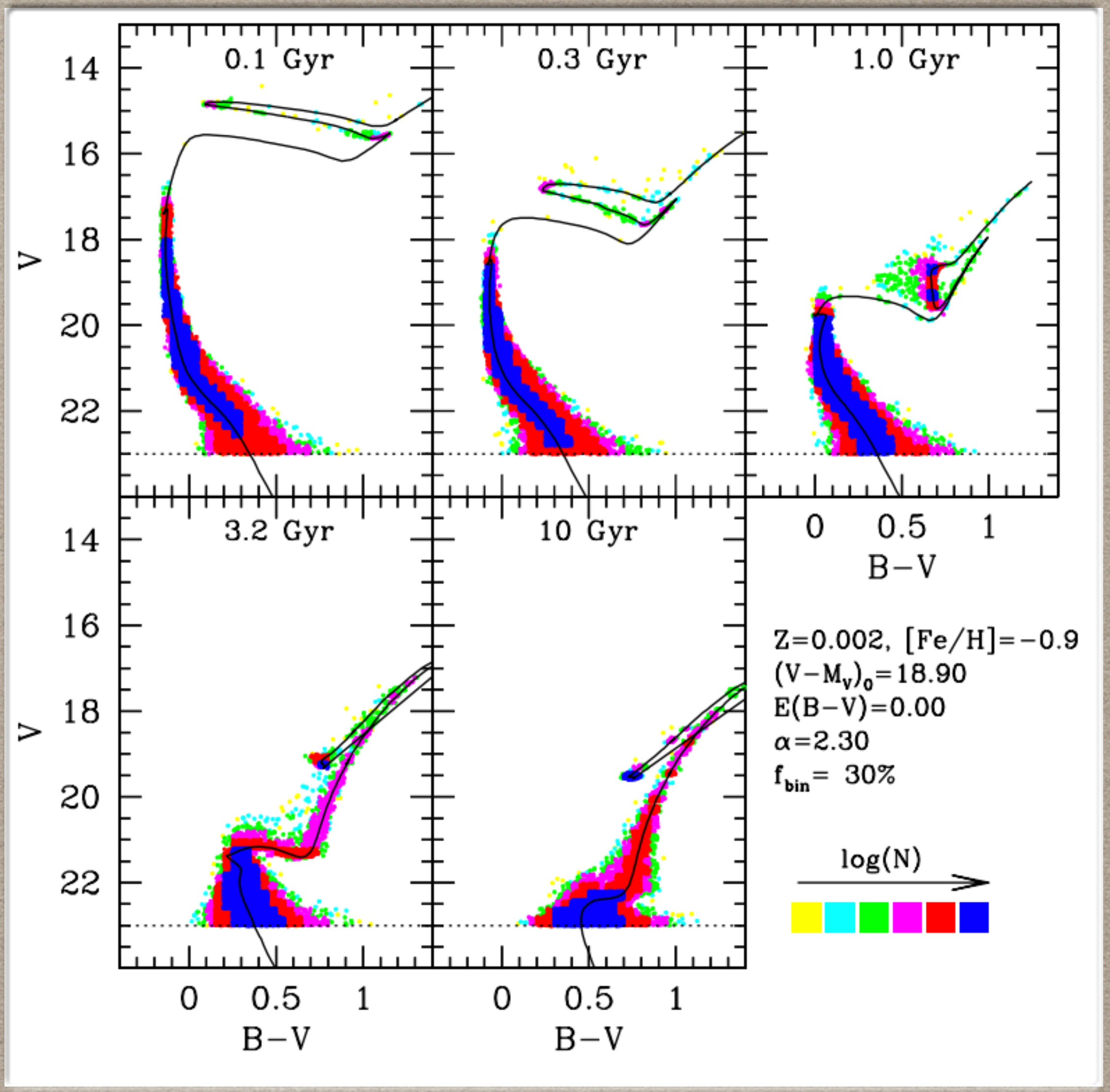


Dias et al. (2014)

THE SOLUTION:
3. FITTING THE CMD

- "make use of as much information provided by the bidimensional CMD plane as possible." (Kerber et al. 2002)

- p_{CMD} = probability of a given (synthetic) star at [(B-V),V] be of a given SSP ~ density in the Hess diagram



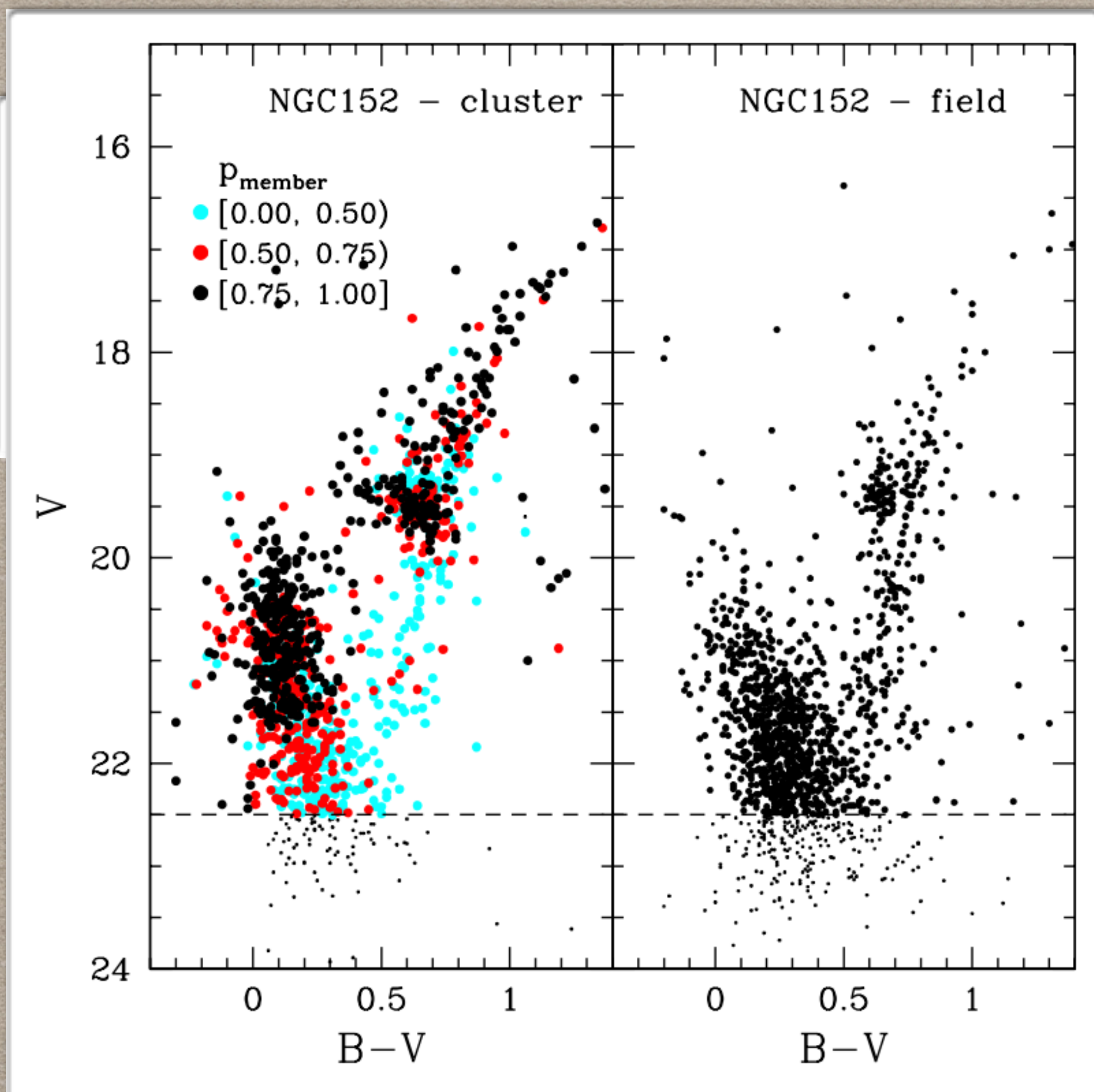
Dias et al. (2016)

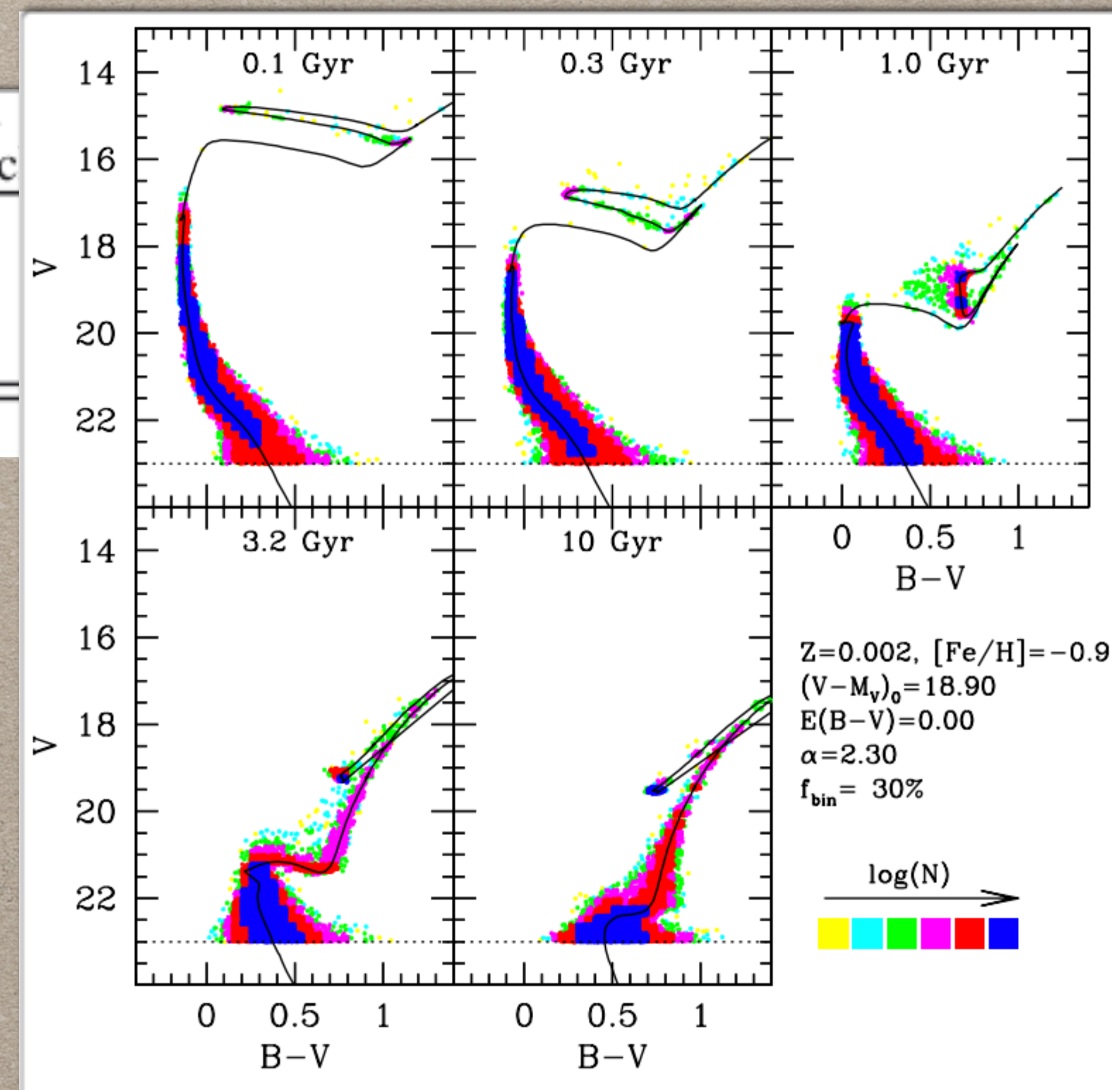
THE SOLUTION:
4. MAXIMUM LIKELIHOOD (EMPIRICAL)

$$L \propto \prod_{i=1}^{N_{\rm clus}} p_{{\rm CMD},i} \times p_{{\rm member},i} \propto \prod_{i=1}^{N_{\rm clus}} N[V_i, (B-V)_i] \times p_{{\rm member},i}$$

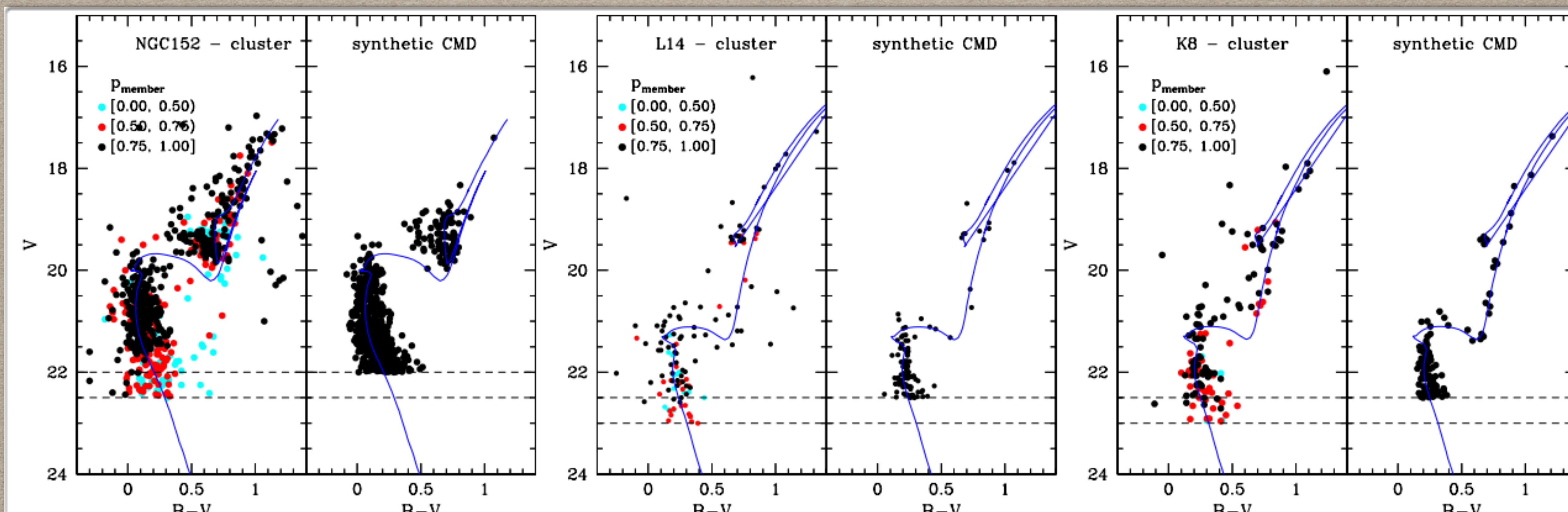
- p_{member} = probability of a given (observed) star be member of a given cluster
- p_{CMD} = probability of a given (synthetic) star at [(B-V),V] be of a given SSP ~ density in the Hess diagram

THE SOLUTION:
4. MAXIMUM LIKELIHOOD





THE SOLUTION:
5. RESULTS



THE THEORY BEHIND

Cleaned CMD *Synthetic CMD*

$$L \equiv p(\{x_i\}|M(\theta)) = \prod_{i=1}^{n} p(x_i|M(\theta))$$

"Likelihood: the probability of the data given the model"

- *Not normalized, i.e., not P.D.F.*
- *Function of the data or the model*
- *MLE is not probability of parameters θ*
- *Usually log(L) is used: same maximum and Π -> Σ*

THE THEORY BEHIND

- Critical assumption: the data truly come from the specified model class.
  Synthetic CMD = stellar evolution and SSP

- Important property: MLE converges to the true parameter value as the number of data point increases
  Even with low statistics we recover well the parameters

THE THEORY BEHIND

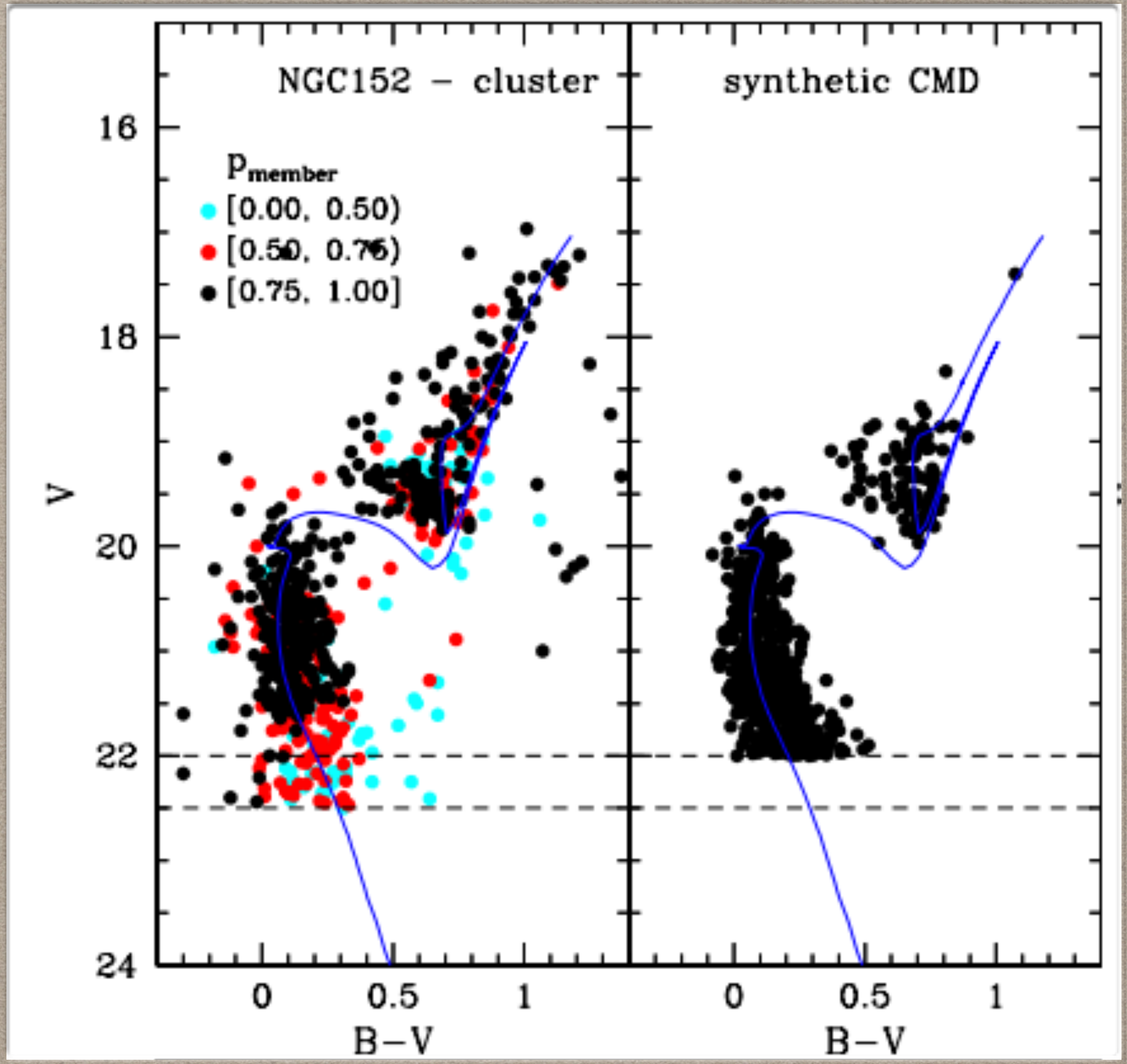
- Goodness of fit: if it is very unlike to obtain L_{max} by randomly drawing data from the best-fit distribution, then the model does not represent well the data
- To compare models and expected L_{max}, we should know the expected distribution of L_{max} (for Gaussian it is the χ^2 distribution)
- To compare L_{max} the models should have the same number of parameters

THE THEORY BEHIND

- Uncertainties: analytically with correlation matrix
  What if we use an empirical model?

THE SOLUTION: 6. UNCERTAINTIES

- 300 synthetic CMDs with same number of stars as observed
- Compute L distribution
- Take the standard deviation



THE THEORY BEHIND: BOOTSTRAP