Star/Galaxy Separation and Mapping the Milky Way Halo in HSC

José Garmilla

Princeton University

November 16, 2015

Outline

- Motivations
- Building Training Catalog
- Measuring Extendedness
- Extendedness Cuts
- Colors
- Colors & Extendedness Combined
- Mapping the MW Halo with MSTO Stars

Motivations

Reliable star/galaxy classification at faint magnitudes is crucial for MW and LG studies.

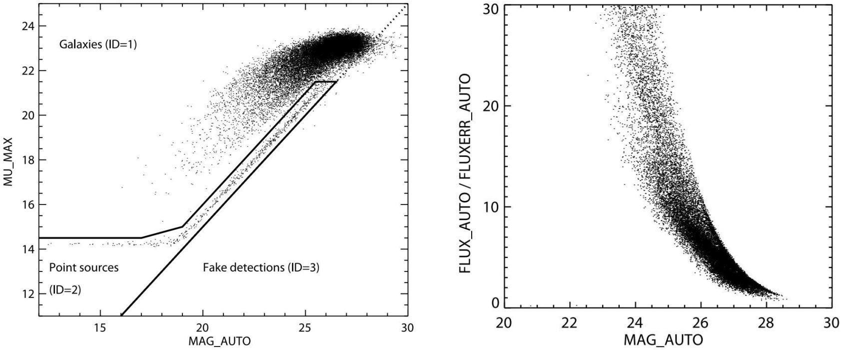
It's also useful for finding supernovae (Misclassified stars can yield false positives).

The PSF estimation code in the HSC pipeline uses samples of stars (The more the better).

The current method in the pipeline is unable to produce pure samples of stars at faint magnitudes.

Building Training Catalog

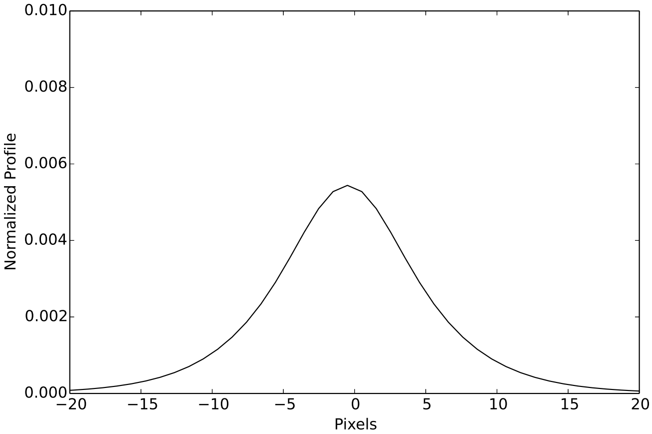
I'm using a public HST/COSMOS catalog (Leauthaud et al. 2007) contiguous 1.64 $\deg^2$ with 0.12" seeing. The labels are obtained from SExtractor parameters MU_MAX and MAG_AUTO.



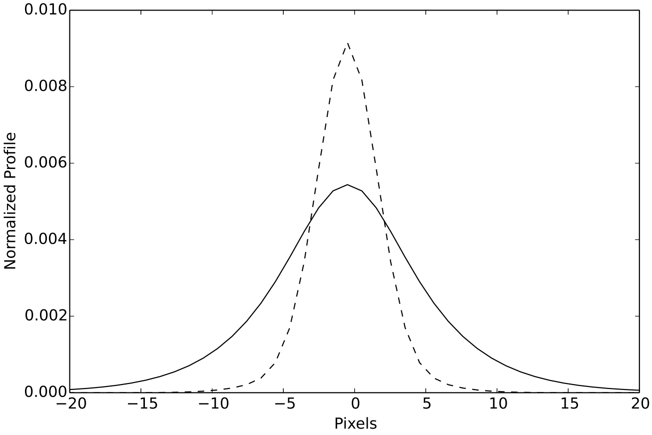
Leauthaud et al. 2007

Measuring Extendedness
Current HSC Pipeline Extendedness Measure

Start with an exponential profile with $R_e = 0.5''$ convolved with the PSF.



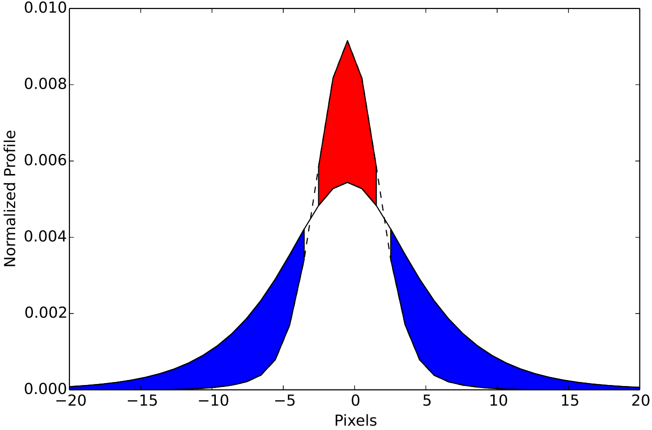
Fit the PSF to the profile (Double Gaussian with $\mathrm{FWHM} = 0.8''$, $\sigma_{\rm out}/\sigma_{\rm in} = 2$, $\mathrm{peak}_{\rm out}/\mathrm{peak}_{\rm in} = 1/10$).



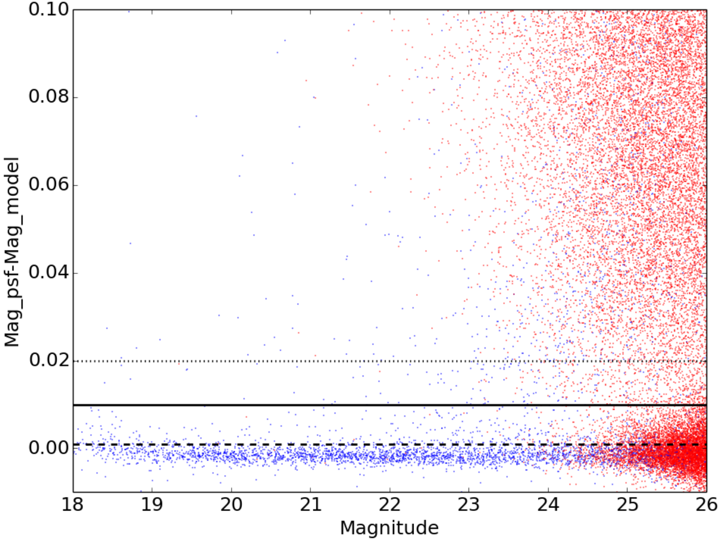
Measuring Extendedness

Current HSC Pipeline Extendedness Measure

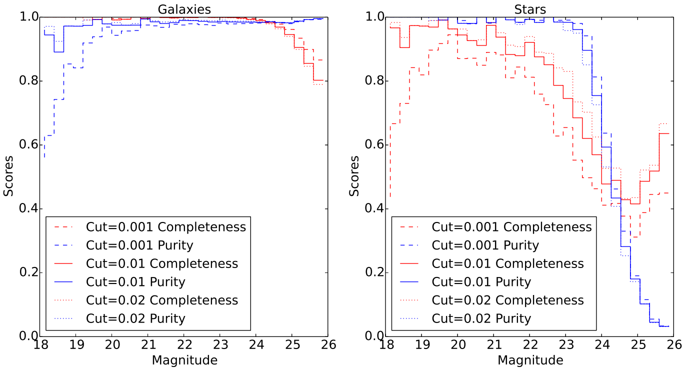
$$\frac{\text{Flux}-\text{psfFlux}}{\text{Flux}} = 1 - \frac{\text{psfFlux}}{\text{Flux}};\ \text{Extendedness} = -2.5\log\left(\frac{\text{psfFlux}}{\text{Flux}}\right).$$



Extendedness Cuts



Extendedness Cuts (Scores)



Colors

If we had the probability distributions of stars and galaxies in color-color diagrams.

$P(\mathrm{Star}|\mathbf{x},\mathbf{S}) = \frac{p(\mathbf{x}|\mathbf{S},\mathrm{Star})P(\mathrm{Star})}{p(\mathbf{x}|\mathbf{S})}$

where $\mathbf{x}$ contains the colors, $\mathbf{S}$ is the covariance matrix of $\mathbf{x}$,

$p(\mathbf{x}|\mathbf{S}) = p(\mathbf{x}|\mathbf{S},\mathrm{Star})P(\mathrm{Star}) + p(\mathbf{x}|\mathbf{S},\mathrm{Galaxy})P(\mathrm{Galaxy})$,

$p(\mathbf{x}|\mathbf{S},\mathrm{Star}) = p(\mathbf{v}|\mathrm{Star}) * \mathcal{N}(\mathbf{0},\mathbf{S})$,

$p(\mathbf{x}|\mathbf{S},\mathrm{Galaxy}) = p(\mathbf{v}|\mathrm{Galaxy}) * \mathcal{N}(\mathbf{0},\mathbf{S})$,

and $p(\mathbf{v}|\mathrm{Star})$ is the noiseless distribution of Stars in $\mathbf{x}$'s colors.

Colors

Extreme Deconvolution (Bovy et al. 2011)

Extreme deconvolution solves the problem of estimating the density of a variable with noisy observations, assuming the noise is Gaussian.

The algorithm estimates the density of random variable $\mathbf{v}$ from a set of observations $\mathbf{x}_i$ given by

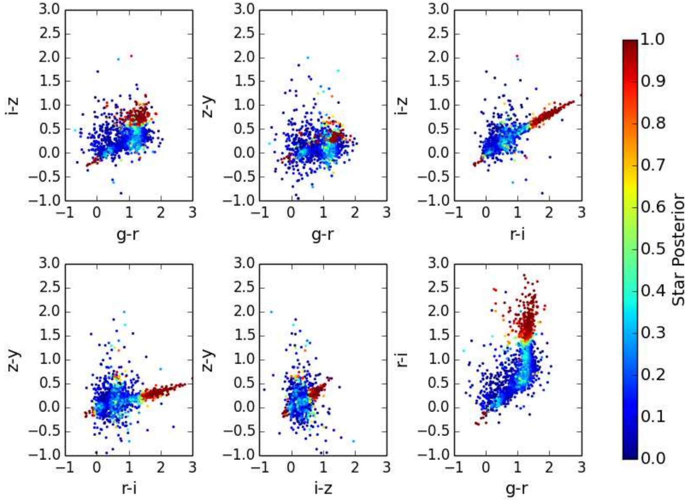
$$\mathbf{x}_i = \mathbf{v}_i + \mathcal{N}(0, \mathbf{S}_i).$$

The density of $\mathbf{v}$ is modeled with a mixture of Gaussians

$$p(\mathbf{v}) = \sum_{k=1}^{K} \alpha_k \mathcal{N}(\mathbf{x}|\mathbf{m}_k, \mathbf{\Sigma}_k).$$

Colors

Posteriors computed with XD fit.



Extendedness & Colors Combined

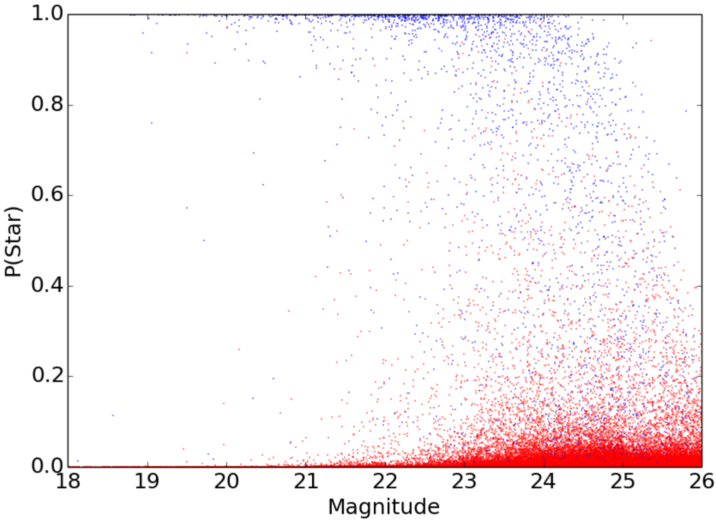
Add $\text{mag}_{psf} - \text{mag}_{cmodel}$ to the vector **x** in the XD inputs, and assume it's not correlated with any colors. To be conservative, assign it the worst-case errorbar $\sqrt{\sigma^2_{psf} + \sigma^2_{cmodel}}$.

Divide the data into four magnitude bins: 18–22, 22–24, 24–25, 25–26.

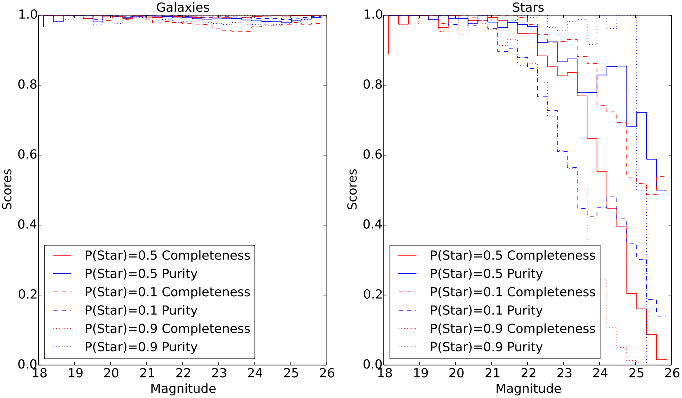
Fit an XD model in each magnitude bin for stars and galaxies separately.

Fix the prior probability of being a star to the fraction of stars we have in the training set in each magnitude bin.

Extendedness & Colors Combined (Posteriors vs Magnitude)

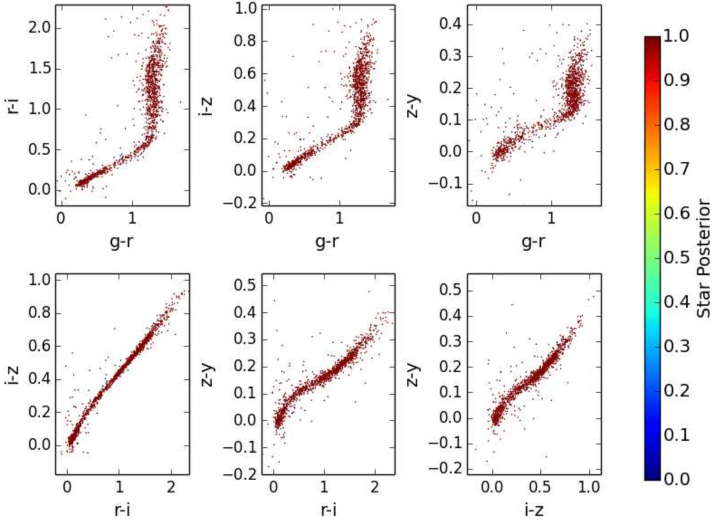


Extendedness & Colors Combined (Scores)



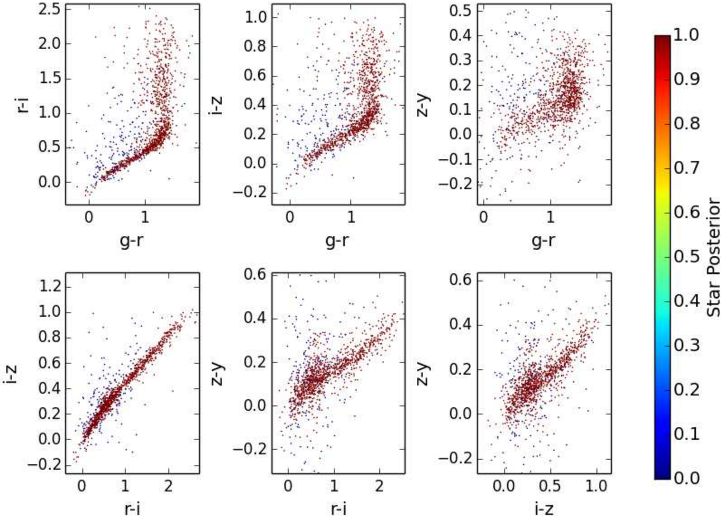
Extendedness & Colors Combined

Bias in Samples of Stars: $18 < i < 22$



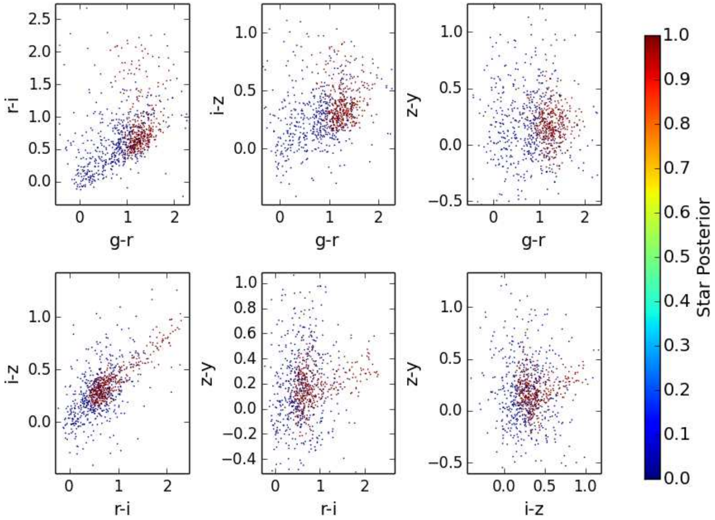
Extendedness & Colors Combined

Bias in Samples of Stars: $22 < i < 24$



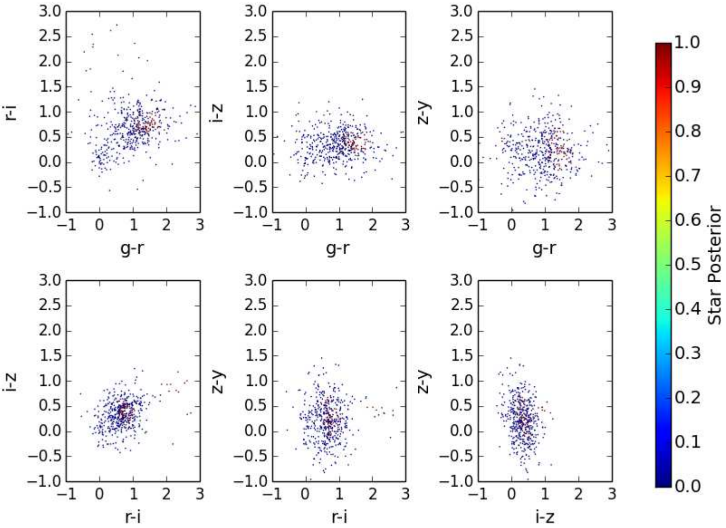
Extendedness & Colors Combined

Bias in Samples of Stars: $24 < i < 25$



Extendedness & Colors Combined

Bias in Samples of Stars: $25 < i < 26$



Additional Information

For more information go to:

http://hscsurvey.pbworks.com/w/page/96040922/Report

The density profile of stars in the halo needs to be known in order to use dynamical tracers to infer the galaxy's mass profile.

This can be easily seen in the radial Jean's equation assuming spherical symmetry and steady-state hydrodynamic equilibrium

$$M(r) = -\frac{r\sigma_{rr}^2}{G}\left[\frac{d\ln\nu}{d\ln r} + \frac{d\ln\sigma_{rr}^2}{d\ln r} + 2\beta(r)\right]$$

where

$$\beta \equiv 1 - \frac{\sigma_{tl}^2}{\sigma_{rr}^2}$$

Mapping the MW Halo with MSTO Stars

Recent Studies

- About 2000 BHB stars from SDSS up to $\sim 100\ \mathrm{kpc}$ (e.g Xue et al. 2008).
- Catalog of 1200 RR Lyrae Stars from the Catalina Sky Survey up to $\sim 100\ \mathrm{kpc}$ (Drake et al. 2013).
- SEGUE K Giant Survey, catalog of 6000 K giants up to 125 kpc (Xue et al. 2014).
- 400 M giant candidates from UKIDSS to about $\sim 200\ \mathrm{kpc}$ (Bochanski et al. 2014).

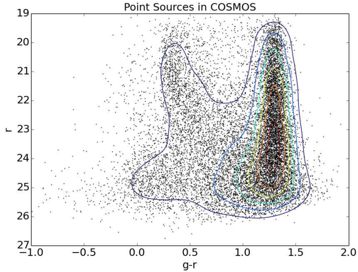
Mapping the MW Halo with MSTO Stars

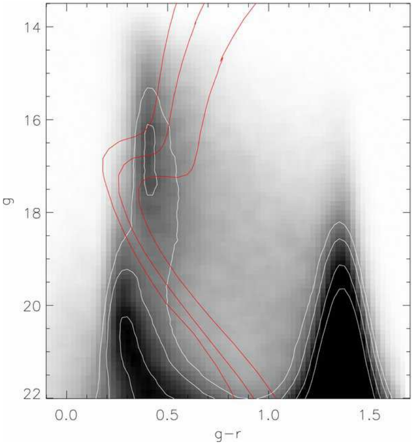
Stellar Populations

Some stellar populations give a biased view of the halo

- BHB and RR Lyrae Stars are most prevalent in old and metal poor populations.
- K Giants are associated with older stellar populations.
- M giants are associated with higher metallicity populations (Higher luminosity accretion events).
- MSTO Stars provide a mostly unbiased picture of the halo.

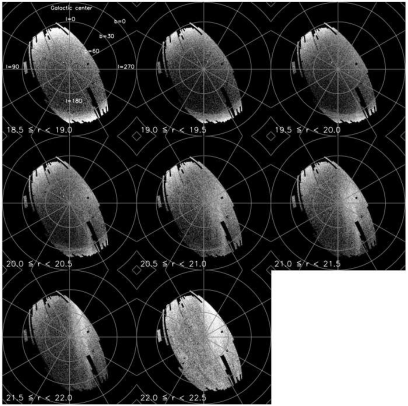
Mapping the MW Halo with MSTO Stars

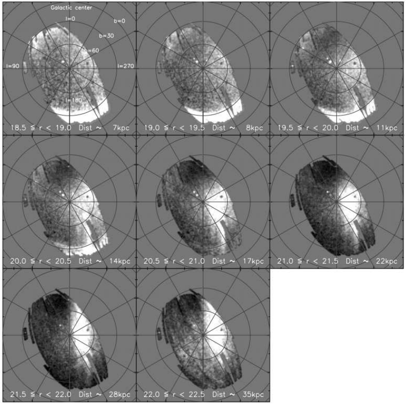




de Jong et al. 2010

Mapping the MW Halo with MSTO Stars





Bell et al. 2008

Mapping the MW Halo with MSTO Stars

If HSC produces a MSTO survey to $i \sim 25$, it will map the halo to ~ 100 kpc with $\sim 10^6$ stars.

This would be by far the most complete and unbiased picture of the halo beyond 30 kpc, and the best estimate of for the smooth radial profile.

Like in Bell et al. 2008, the MSTO survey could be used to estimate the fraction of stars that reside in substructures, and get a handle on the halo oblateness at ~ 100 kpc.

Unkown substructures (dwarf galaxies and/or tidal streams) are likely to emerge.

Conclusions

Star/galaxy separation is very important for MW science.

The current star/galaxy separation technique in the HSC pipeline fails to produce clean samples of stars beyond $i \approx 24$.

HSC colors by themselves do not provide a way of cleanly separating stars from galaxies (Late M type stars being a notable exception).

Combining HSC colors with extendedness adds a lot of information. One can produce clean samples of stars to $i \approx 25$.