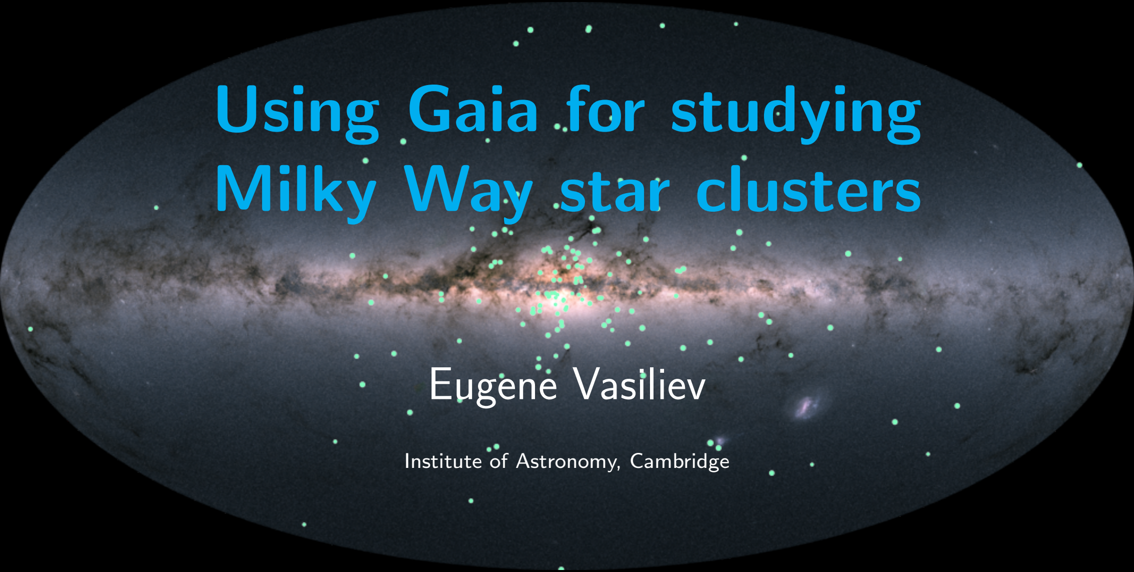
Using Gaia for studying Milky Way star clusters



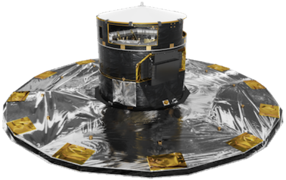
Eugene Vasiliev

Institute of Astronomy, Cambridge

Synopsis

- Overview of the Gaia mission and DR2:

  scientific instruments, catalogue contents, measurement uncertainties, caveats and limitations.



- Using Gaia astrometry for studying internal kinematics of globular clusters:

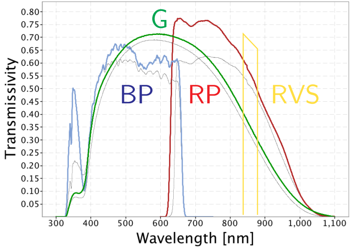
  sample cleaning, membership determination, measurement of velocity dispersion and rotation in the presence of correlated systematic errors.

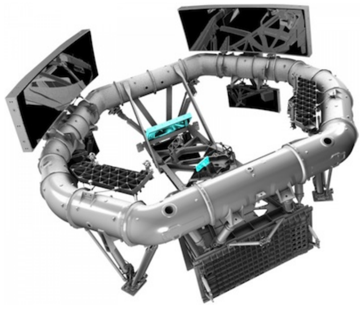
- Galactic population of globular clusters:

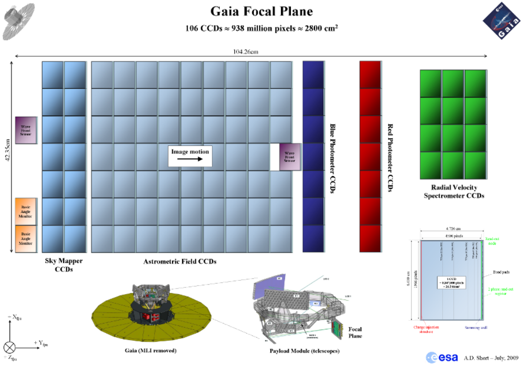
  distribution of clusters in 6d phase space, orbits, inference on the Milky Way potential.

Overview of Gaia mission

- Scanning the entire sky every couple of weeks
- Astrometry for sources down to 21 mag
- Broad-band photometry/low-res spectra
- Radial velocity down to $\sim$ 15 mag (end-of-mission)







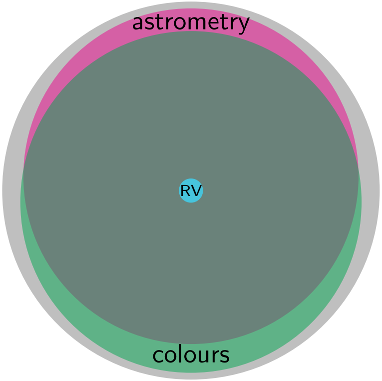
[Source: ESA]

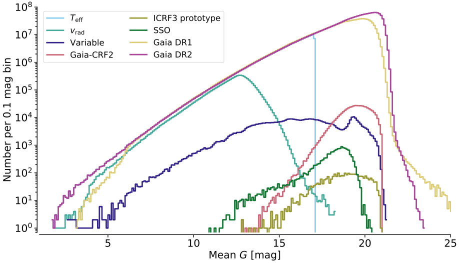
Overview of Data Release 2

- Based on 22 months of data collection

- Total number of sources: 1.69×10^9
- Sources with full astrometry (parallax ϖ, proper motions $\mu_{\alpha*}, \mu_\delta$): 1.33×10^9
- Colours (G_{BP}, G_{RP}): 1.38×10^9
- Radial velocities: 7.2×10^6

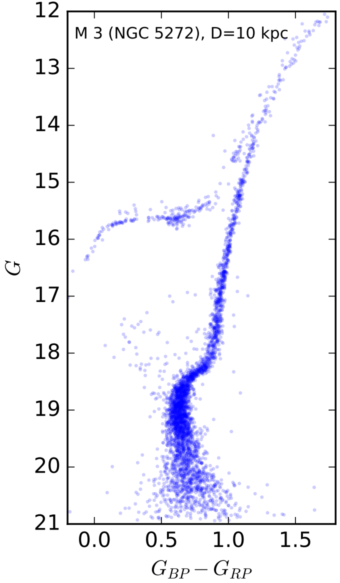
- Effective temperature: 160×10^6
- Stellar parameters $(R_\odot, L_\odot)$: 77×10^6
- Extinction and reddening: 88×10^6
- Variable sources: 0.55×10^6

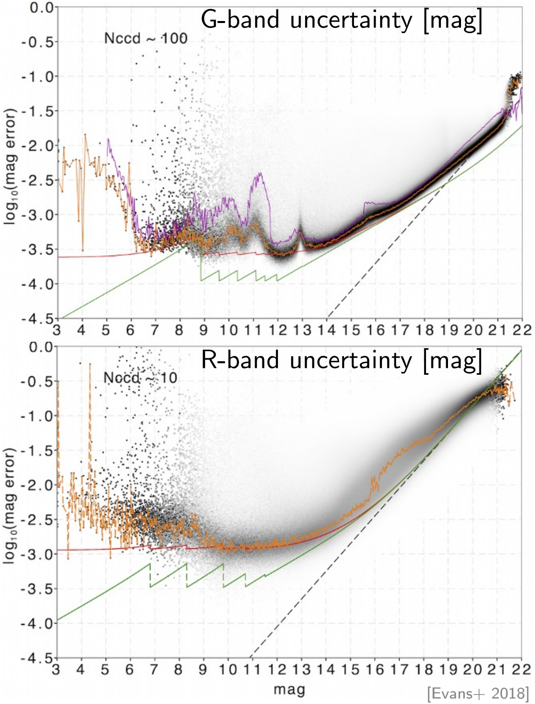




[Brown+ 2018]

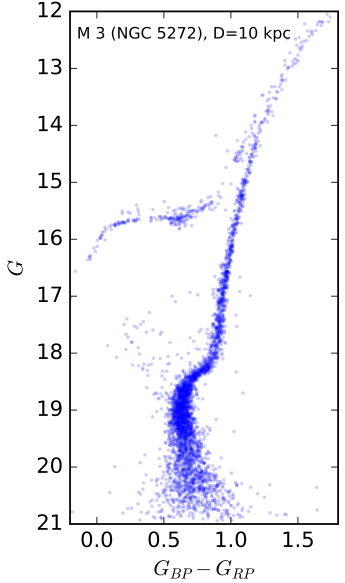
Photometry

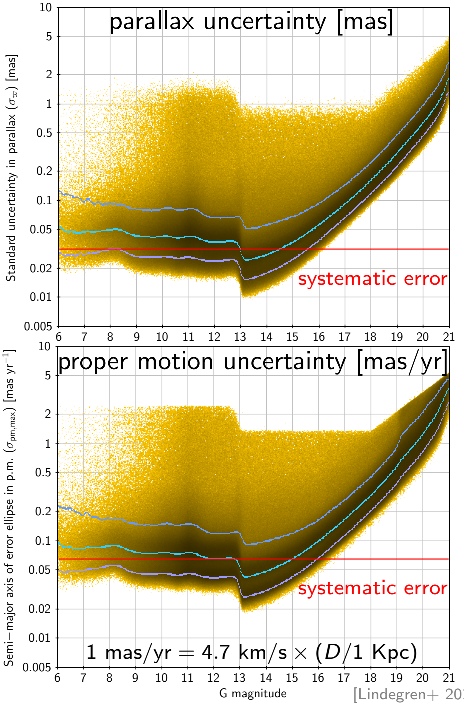




[Evans+ 2018]

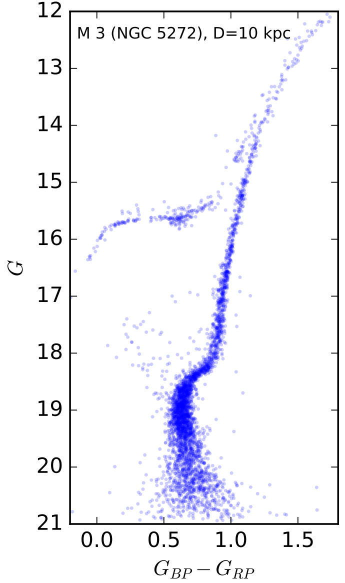
Astrometry



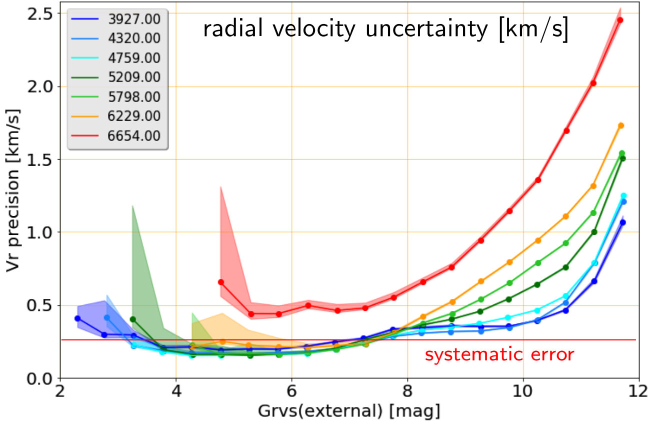


[Lindegren+ 2018]

Spectroscopy



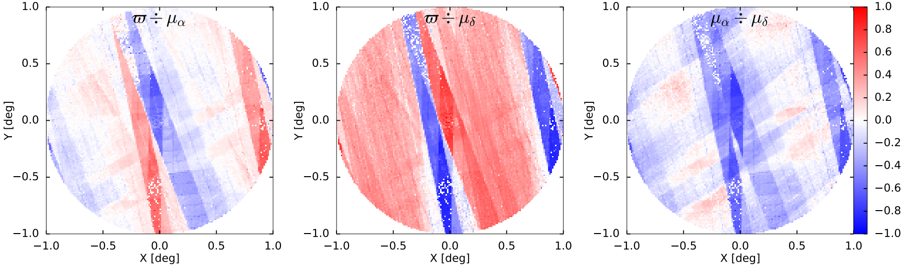
RV measurements only for stars with
$T_{\rm eff} \in [3500 \div 6900]$ K and $G_{\rm RVS} \leq 12$



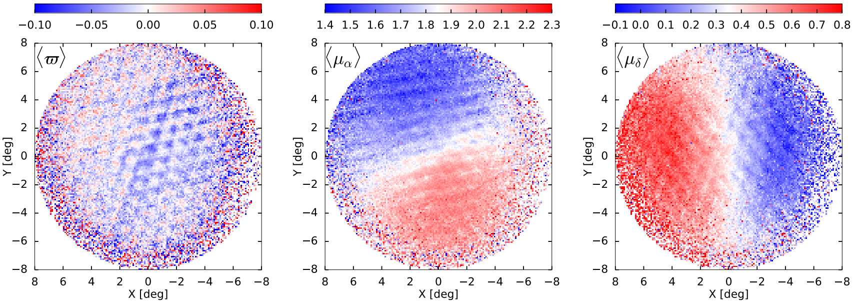
[Katz+ 2018]

Correlations and systematic errors

correlations between parallax and PM (ω Cen region)



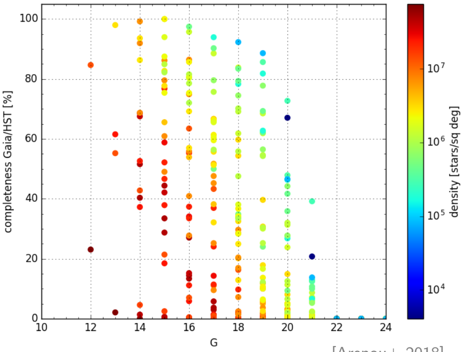
mean parallax and PM (Large Magellanic Cloud region)



Limitations

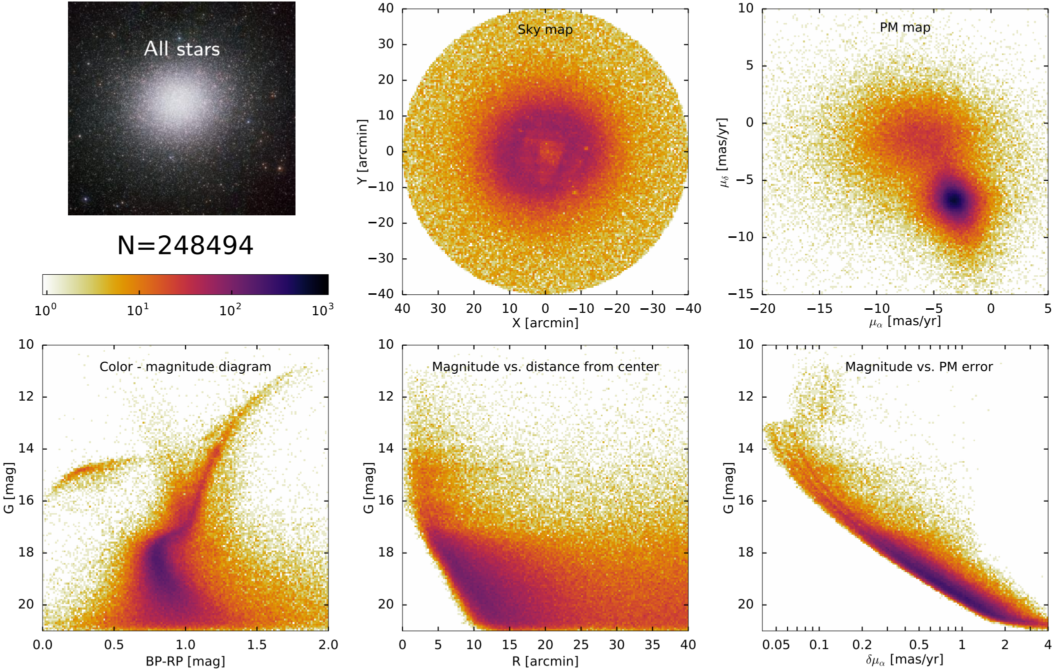
- No special processing for binary stars
  (but a poor astrometric solution is marked clearly – `astrometric_excess_noise`)
- Colour photometry has lower spatial resolution
  (a quality control flag is provided – `phot_bp_rp_excess_factor`)
- Poor completeness at faint magnitudes in crowded regions

- Need to apply various filters to clean up the sample
  (but do it wisely, e.g., don't just throw away stars with negative parallaxes $\varpi < 0$ [see Luri+ 2018 for a discussion])
- Should not generally use $1/\varpi$ as a proxy for distance (unless $\epsilon_\varpi/\varpi \lesssim 0.2$)



[Arenou+ 2018]

Example: NGC 5139 (ω Cen)

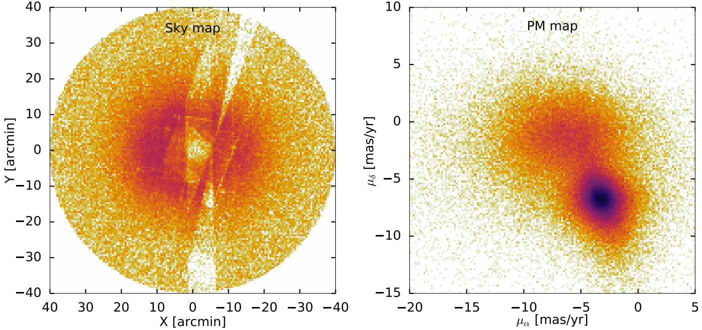


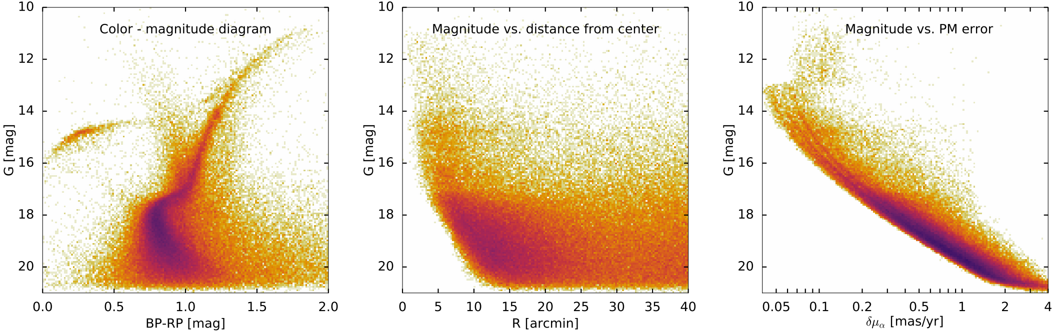
Example: NGC 5139 (ω Cen)

Stars with full astrometry

$(\varpi, \mu_\alpha, \mu_\delta)$

N=156227





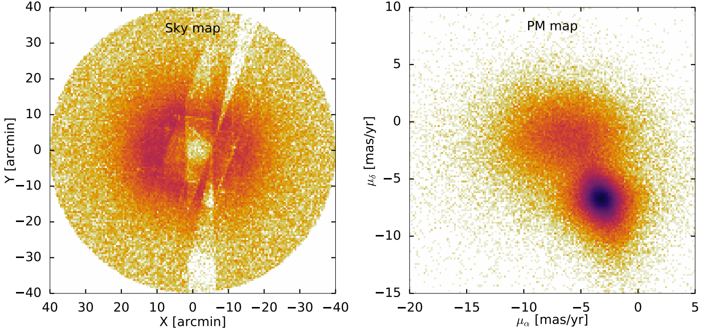
Example: NGC 5139 (ω Cen)

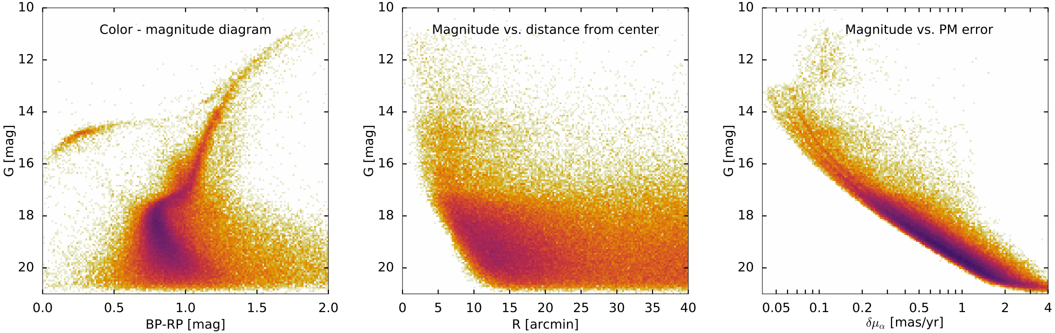
Parallax cut:

$$\varpi - \varpi_0 < 3\,\delta\varpi,$$

$$\varpi_0 = 0.2 \text{ mas}$$

N=146313





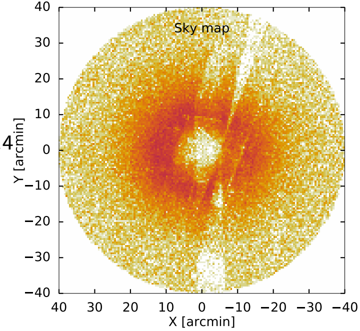
Example: NGC 5139 (ω Cen)

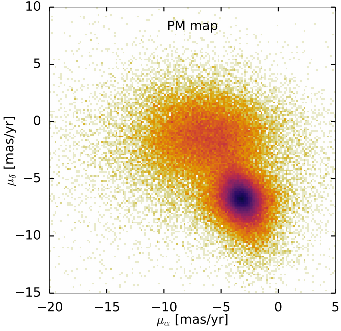
Cut on astrometric quality:

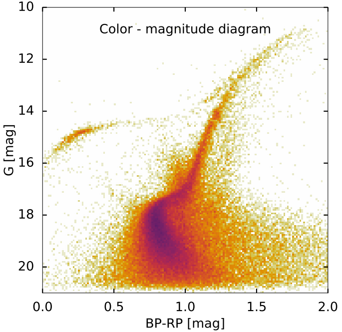
`astrometric_excess_noise` < 1 mas

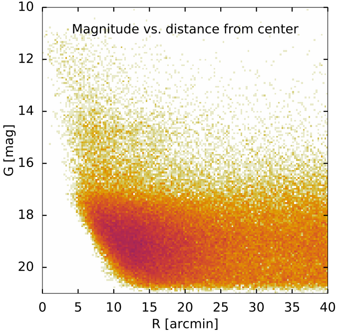
`renormalized_unit_weight_error` < 1.4

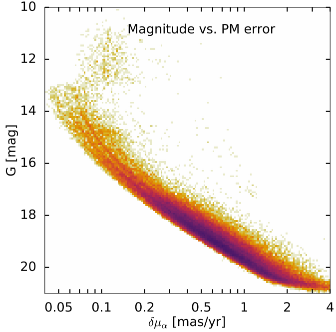
N=111336











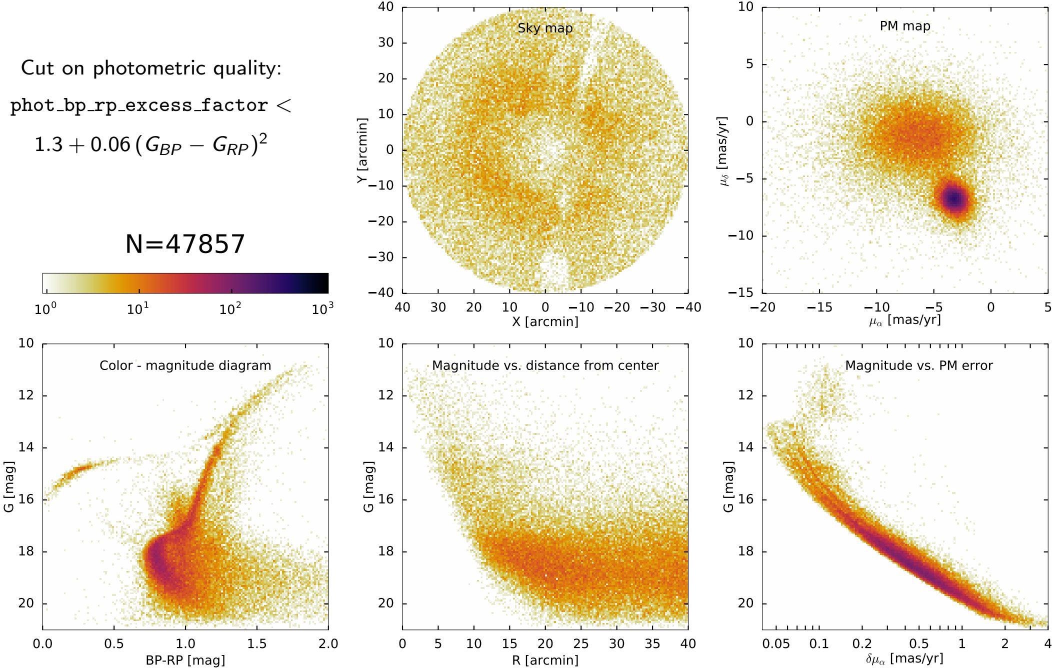
Example: NGC 5139 (ω Cen)

Cut on photometric quality:

`phot_bp_rp_excess_factor` $<$

$$1.3 + 0.06\,(G_{BP} - G_{RP})^2$$

N=47857



Example: NGC 5139 (ω Cen)

Cut on proper motions:

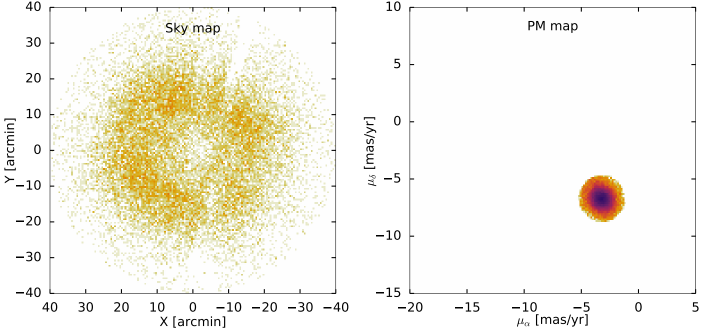
$$(\mu_\alpha - \mu_{\alpha,0})^2 + (\mu_\delta - \mu_{\delta,0})^2 < \Delta\mu^2,$$

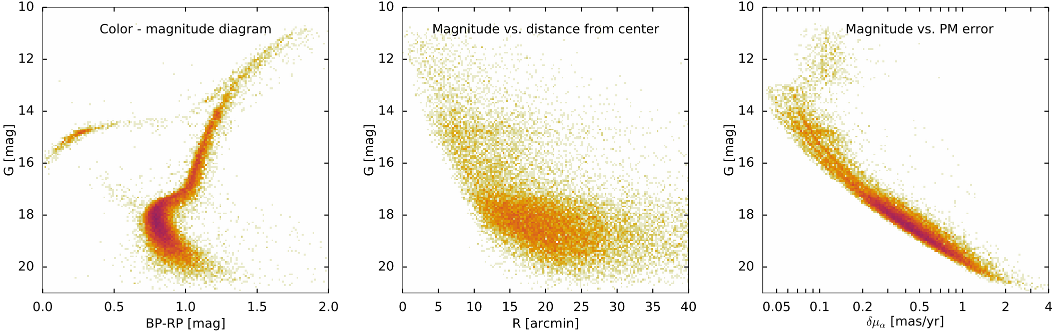
$$\mu_{\alpha,0} = -3.2 \text{ mas/yr},$$

$$\mu_{\delta,0} = -6.7 \text{ mas/yr},$$

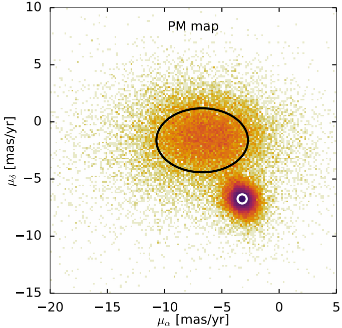
$$\Delta\mu = 2 \text{ mas/yr}$$

N=22865





Membership determination by proper motions



A hard cutoff in PM space is not always possible and is conceptually unsatisfactory.

A more mathematically well-grounded alternative: gaussian mixture modelling.

$$f(\boldsymbol{\mu}_i) = q_{\rm cl}\,\mathcal{N}(\boldsymbol{\mu}_i \mid \overline{\boldsymbol{\mu}_{\rm cl}}, \Sigma_{{\rm cl};i}) + (1-q_{\rm cl})\,\mathcal{N}(\boldsymbol{\mu}_i \mid \overline{\boldsymbol{\mu}_{\rm fg}}, \Sigma_{{\rm fg};i})$$

$$\mathcal{N}(\boldsymbol{\mu} \mid \overline{\boldsymbol{\mu}}, \Sigma) \equiv \frac{\exp\left[-\frac{1}{2}(\boldsymbol{\mu}-\overline{\boldsymbol{\mu}})^T\,\Sigma^{-1}\,(\boldsymbol{\mu}-\overline{\boldsymbol{\mu}})\right]}{2\pi\,\sqrt{\det\Sigma}},$$

where the mean PMs $\overline{\boldsymbol{\mu}}$ and dispersions Σ of the cluster and foreground distributions, and the fraction of cluster members $q_{\rm cl}$, are all inferred by maximizing the likelihood of the observed stellar PMs.

Membership determination by proper motions

Take into account the measurement errors $\epsilon_{\mu_\alpha}, \epsilon_{\mu_\delta}, \rho_{\mu_\alpha\mu_\delta}$ for each star i:

$$\Sigma_{\mathrm{cl};i} = \begin{pmatrix} \sigma_{\mathrm{cl}}^2(\mathbf{r}_i) + \epsilon_{\mu_\alpha}^2 & \rho_{\mu_\alpha\mu_\delta}\,\epsilon_{\mu_\alpha}\,\epsilon_{\mu_\delta} \\ \rho_{\mu_\alpha\mu_\delta}\,\epsilon_{\mu_\alpha}\,\epsilon_{\mu_\delta} & \sigma_{\mathrm{cl}}^2(\mathbf{r}_i) + \epsilon_{\mu_\delta}^2 \end{pmatrix}$$

$$\Sigma_{\mathrm{fg};i} = \begin{pmatrix} S_{\alpha\alpha} + \epsilon_{\mu_\alpha}^2 & S_{\alpha\delta} + \rho_{\mu_\alpha\mu_\delta}\,\epsilon_{\mu_\alpha}\,\epsilon_{\mu_\delta} \\ S_{\alpha\delta} + \rho_{\mu_\alpha\mu_\delta}\,\epsilon_{\mu_\alpha}\,\epsilon_{\mu_\delta} & S_{\delta\delta} + \epsilon_{\mu_\delta}^2 \end{pmatrix}$$

And allow for a spatially-dependent density of cluster members: $q_{\mathrm{cl}}(\mathbf{r}_i)$.

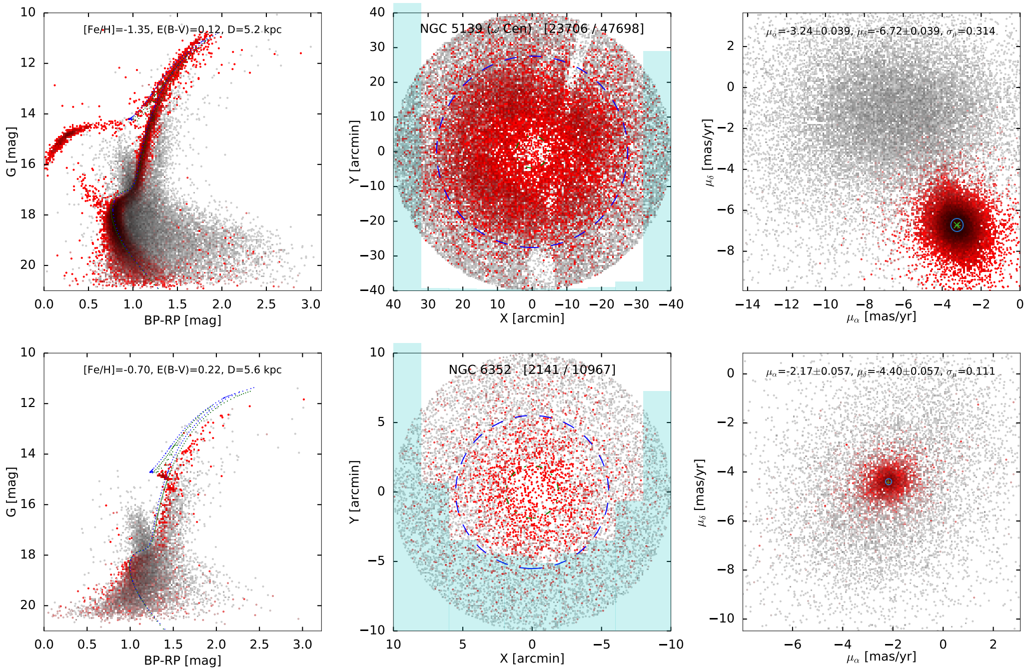
Maximize $\ln\mathcal{L} \equiv \sum_{i=1}^{N_{\mathrm{stars}}} \ln f(\boldsymbol{\mu}_i)$ by adjusting free parameters:

$\overline{\boldsymbol{\mu}_{\mathrm{cl}}}$, $\overline{\boldsymbol{\mu}_{\mathrm{fg}}}$, $S_{\alpha\alpha}, S_{\delta\delta}, S_{\alpha\delta}$, radius and normalization of $q_{\mathrm{cl}}(\mathbf{r})$, normalization of $\sigma_{\mathrm{cl}}(\mathbf{r})$.

Posterior membership probability for each star:

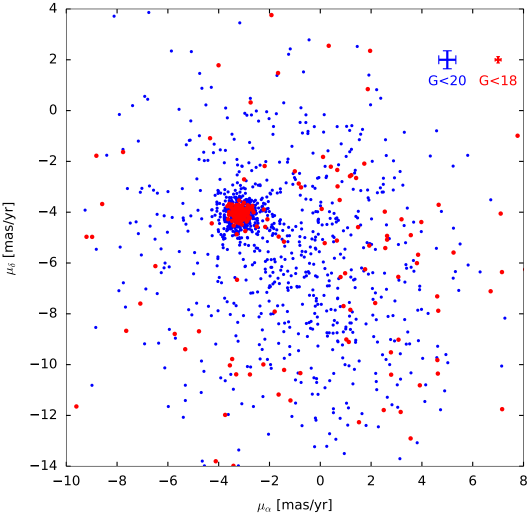
$$p_{\mathrm{cl};i} = \frac{q_{\mathrm{cl}}(\mathbf{r}_i)\,\mathcal{N}(\boldsymbol{\mu}_i \mid \overline{\boldsymbol{\mu}_{\mathrm{cl}}}, \Sigma_{\mathrm{cl};i})}{q_{\mathrm{cl}}(\mathbf{r}_i)\,\mathcal{N}(\boldsymbol{\mu}_i \mid \overline{\boldsymbol{\mu}_{\mathrm{cl}}}, \Sigma_{\mathrm{cl};i}) \;+\; [1 - q_{\mathrm{cl}}(\mathbf{r}_i)]\,\mathcal{N}(\boldsymbol{\mu}_i \mid \overline{\boldsymbol{\mu}_{\mathrm{fg}}}, \Sigma_{\mathrm{fg};i})}$$

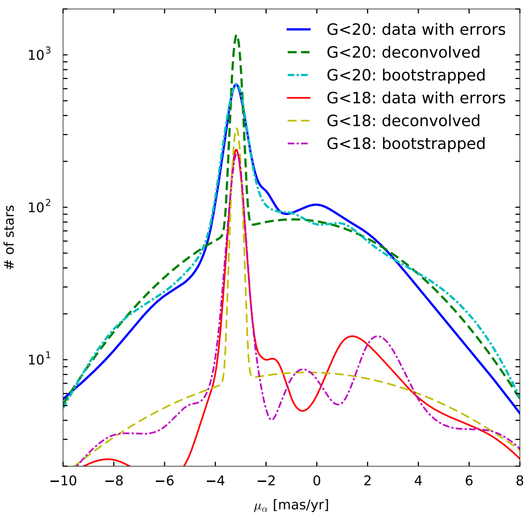
Examples of membership determination



Inferring the internal velocity dispersion

The intrinsic PM distribution is broadened by observational errors, but if they are properly taken into account, the inferred PM dispersion σ_{cl} should be independent of the selected subset of stars (bright or faint).



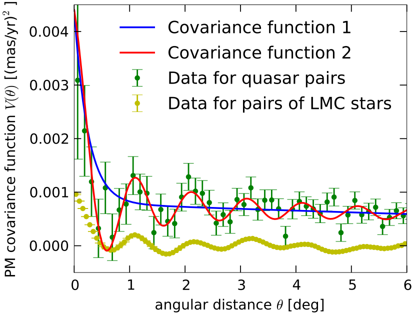


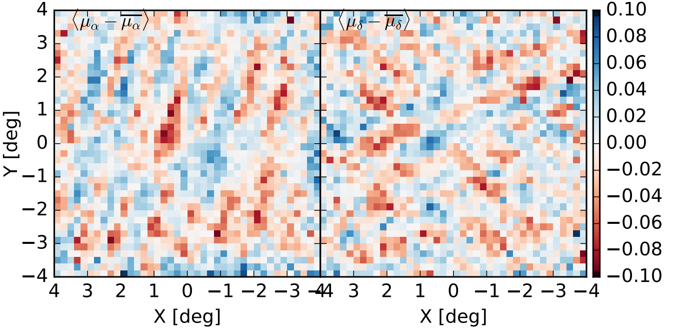
Caveat: spatially correlated systematic errors

$V(\theta_{ij}) = \langle \mu_i \, \mu_j \rangle$, averaged over pairs of sources separated by angular distance θ.
Large scale: quasars ($\sim 0.5 \times 10^6$ sources across the entire sky).
Small scale: stars in LMC ($\sim 0.4 \times 10^6$ sources, subtract the mean PM).

$$V(\theta) = 0.0008 \exp(-\theta/20^\circ) + \left[\begin{array}{ll} 0.0036 \exp(-\theta/0.25^\circ) & \text{(cov.fnc.1)} \\ 0.004 \,\mathrm{sinc}(\theta/0.5^\circ + 0.25) & \text{(cov.fnc.2)} \end{array} \right.$$

On small scales, the typical systematic error is $\sqrt{V(0)} \sim 0.06$ mas/yr;
on scales $\gtrsim 0.5^\circ$, it is ~ 0.03 mas/yr.





How to properly account for correlated systematic errors

Likelihood function for the entire dataset ($\boldsymbol{\mu} \equiv \{\mu_i\}_{i=1}^{N}$):

$$\mathcal{L} = \mathcal{N}(\boldsymbol{\mu} \mid \mathbf{1}\,\overline{\mu}, \Sigma)$$

$$\Sigma = \begin{pmatrix} V(0)+\epsilon_1^2 & V(\theta_{12}) & V(\theta_{13}) & \cdots & V(\theta_{1N}) \\ V(\theta_{21}) & V(0)+\epsilon_2^2 & V(\theta_{23}) & \cdots & V(\theta_{2N}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ V(\theta_{N1}) & V(\theta_{N2}) & V(\theta_{N3}) & \cdots & V(0)+\epsilon_N^2 \end{pmatrix}$$

($\epsilon_1 .. \epsilon_N$ are statistical errors of each datapoint).

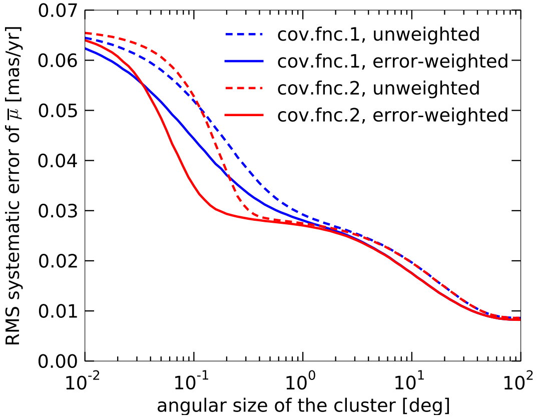
This is easily generalized to 2d case with 2×2 covariance matrices of statistical errors, and allowing for spatially-dependent internal dispersion and mean value of μ.

The downside is that one needs to invert the $N \times N$ covariance matrix Σ for the entire dataset (for the optimal error-weighted estimate).

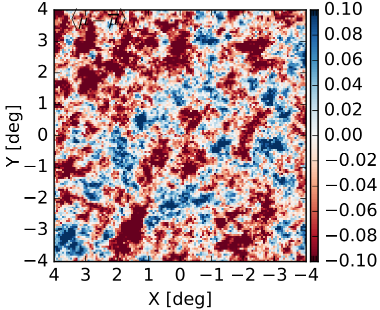
Alternatively, an unweighted estimate of the uncertainty on $\overline{\mu}$ is $\frac{1}{N}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\Sigma_{ij}}$.

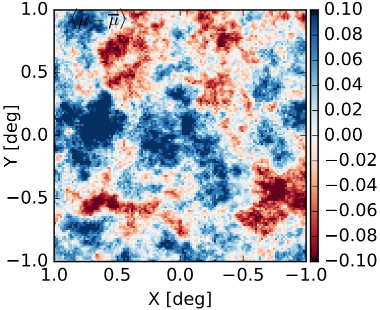
Systematic uncertainty on the mean PM

Generate many realizations of mock PM maps
with the given spatial covariance,
for clusters with different spatial extent,
and estimate the uncertainty in the mean PM.



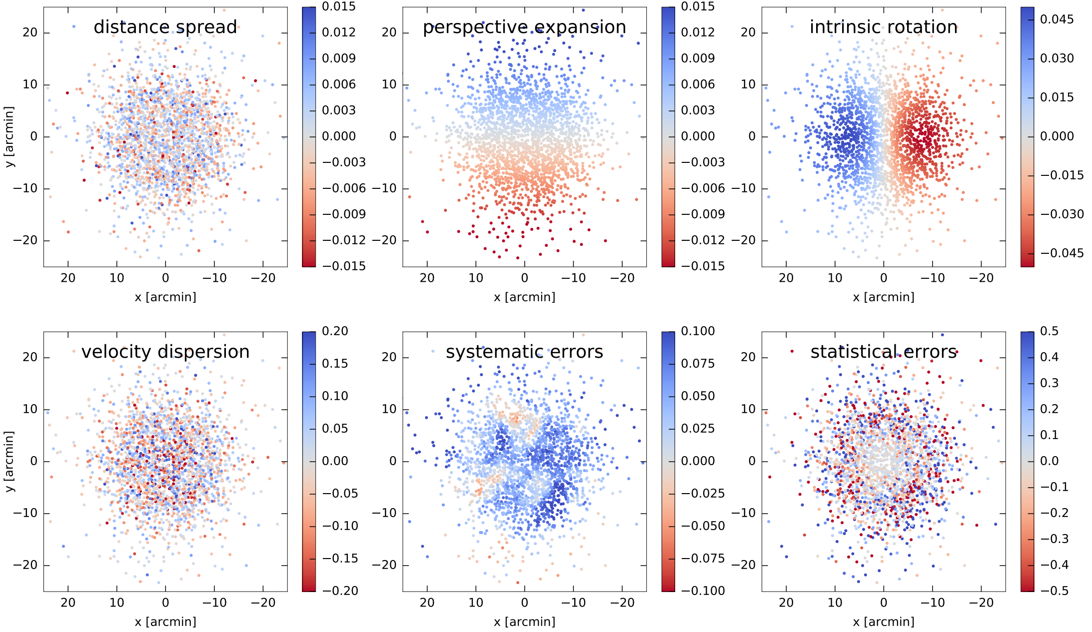
mock systematic error maps





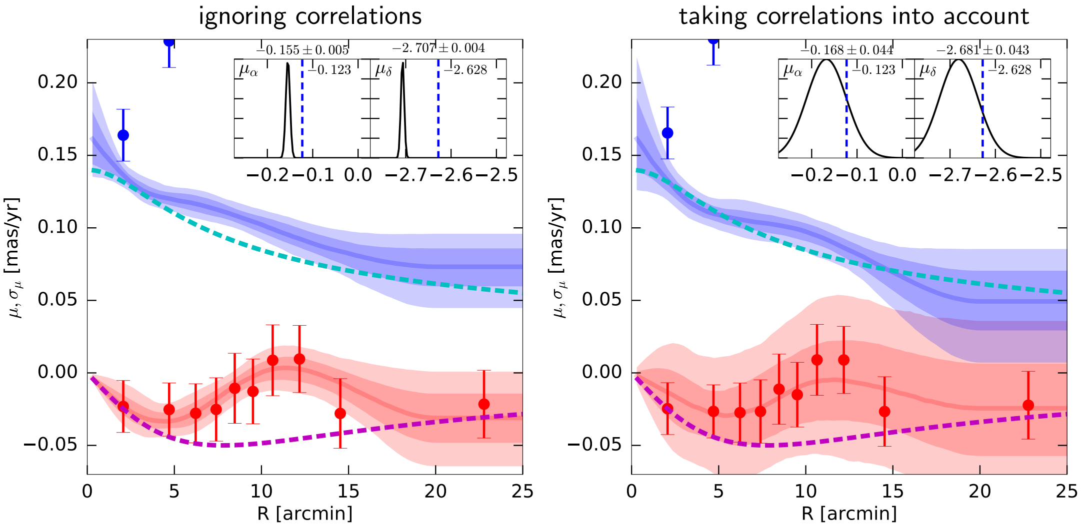
Example: mock PM maps for NGC 5272 (M 3)

distance = 10 kpc; $v_{\rm los} = -150$ km/s; $\sigma = 6$ km/s; $v_{\rm rot} = 2$ km/s; 4000 stars



Spatial correlation should not be ignored!

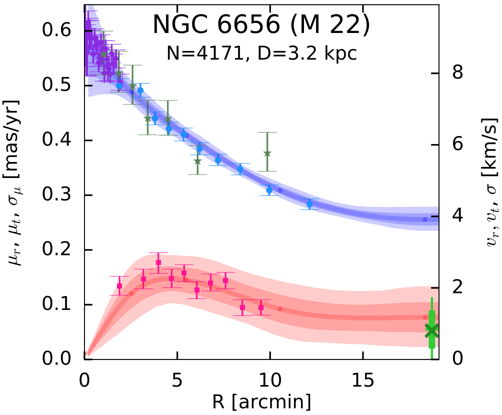
Doing so underestimates the error bars on fit parameters,
even when systematic errors are much smaller than statistical errors.

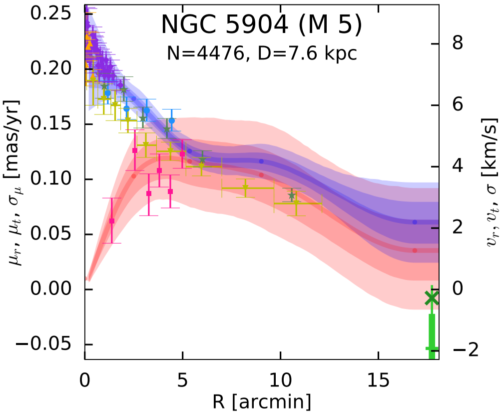


Internal kinematics of globular clusters from Gaia PM

Clear signature of rotation in $\sim$ 10 clusters:

confirming the results of Bianchini+ 2018 based on Gaia data.

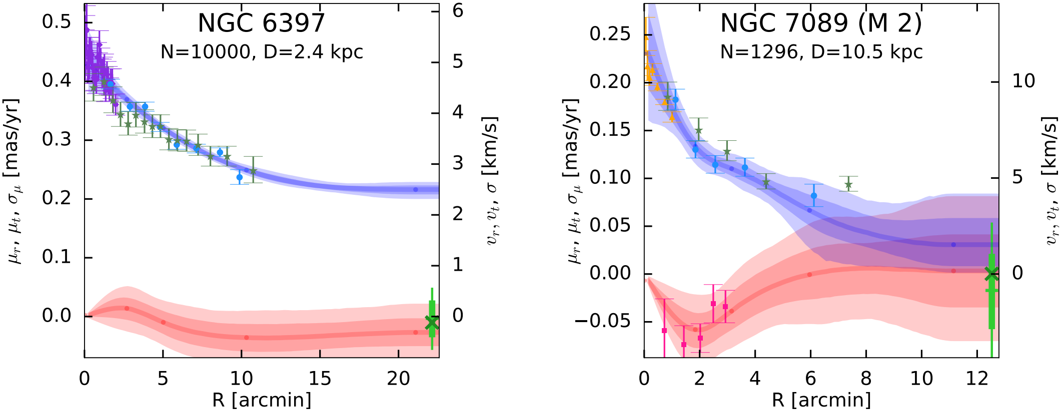




Internal kinematics of globular clusters from Gaia PM

Good match between line-of-sight velocity dispersion and PM dispersion:

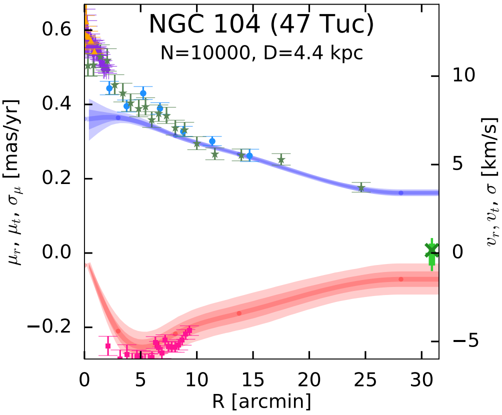
confirming the analysis of Gaia PM dispersion profiles by Baumgardt+ 2018, and complementing the HST measurements in central parts [Bellini+ 2014, Watkins+ 2015].

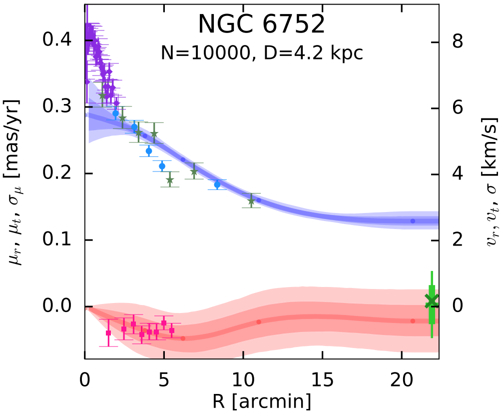


Internal kinematics of globular clusters from Gaia PM

Mismatch between $\sigma_{\rm los}$ and PM dispersion in central parts:

likely due to crowding issues and aggressive sample cleanup.

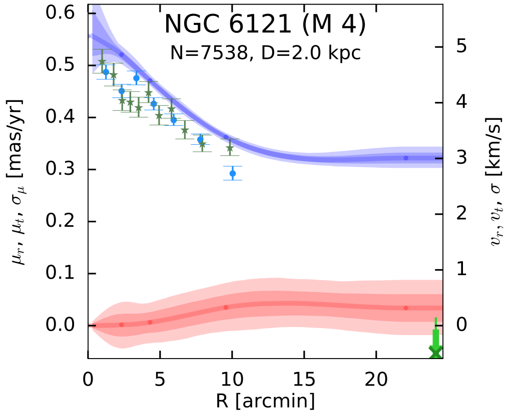


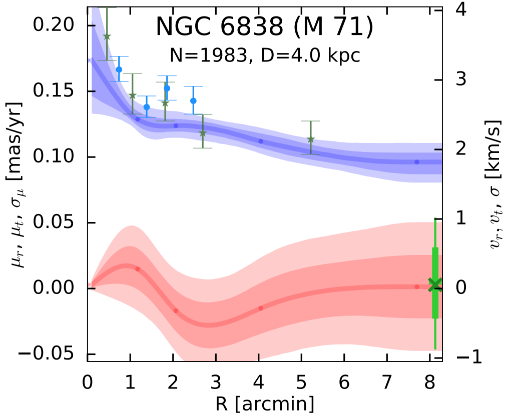


Internal kinematics of globular clusters from Gaia PM

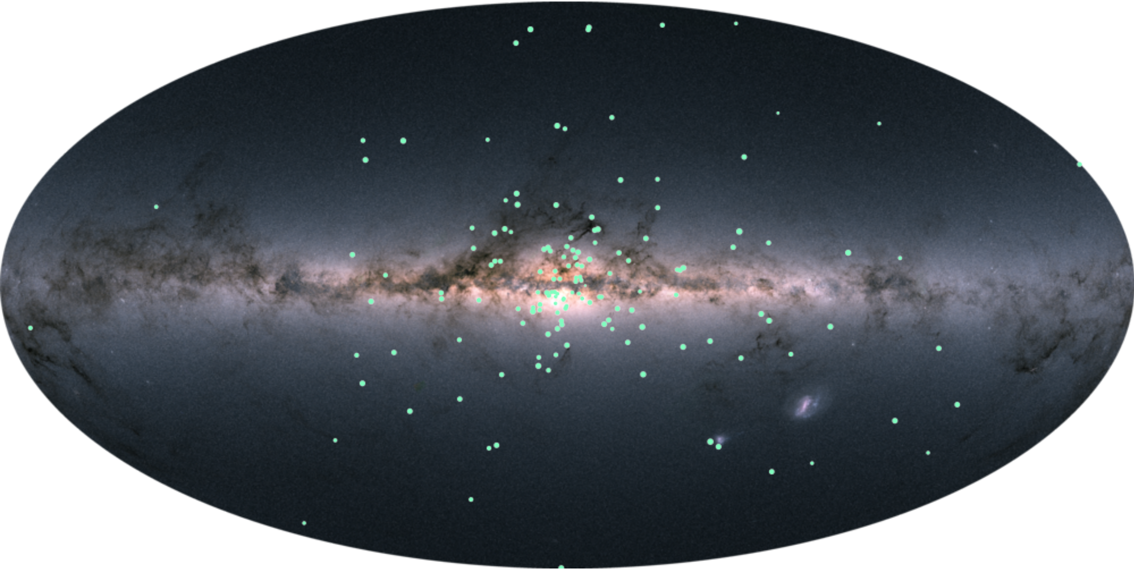
Overall scale mismatch between $\sigma_{\rm los}$ and σ_μ:

a prime method for measuring the distance from kinematic analysis.



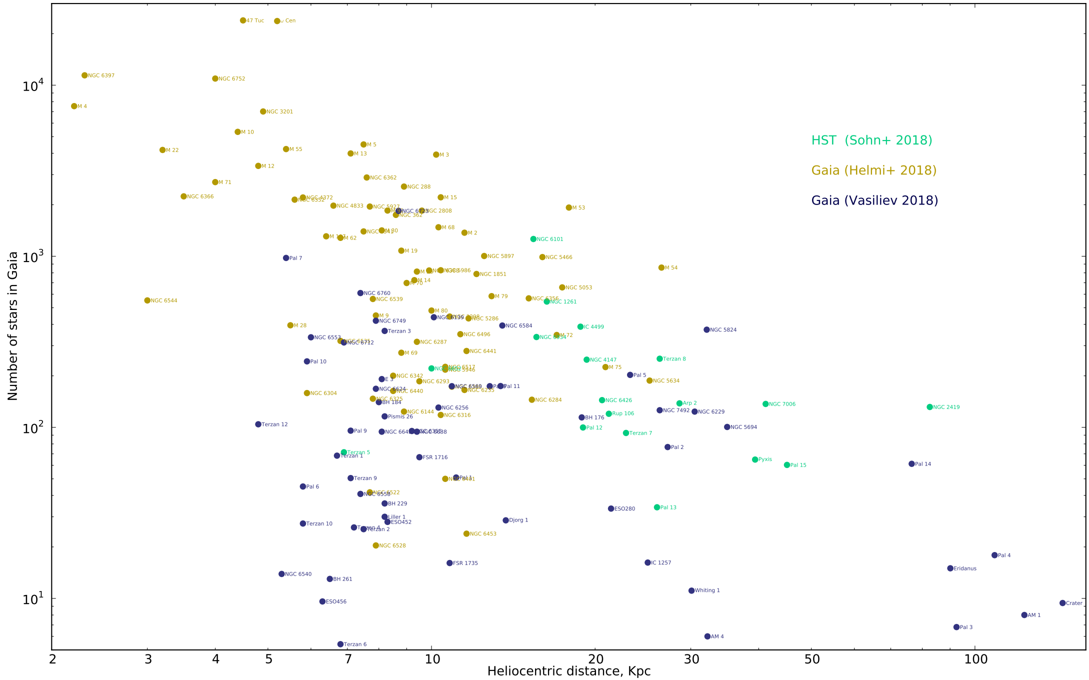


Distribution of globular clusters in galactic coordinates

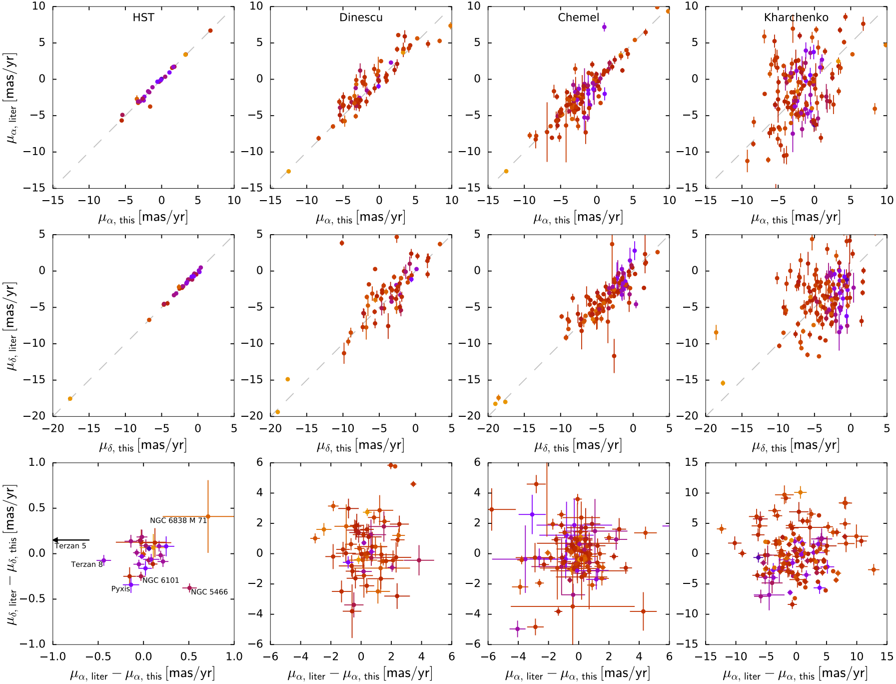


~ 150 globular clusters in the Milky Way

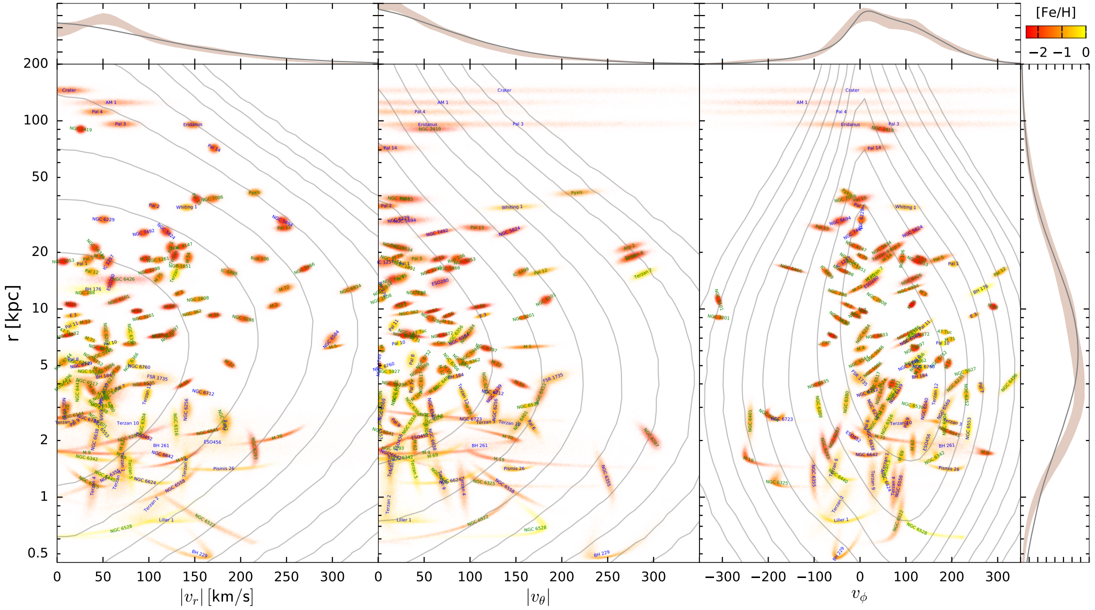
Distribution of globular clusters in distance and # of stars



Previous measurements of mean PM of globular clusters



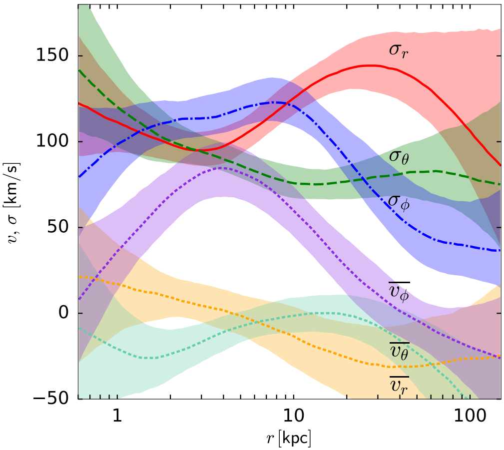
Distribution of globular clusters in position/velocity space



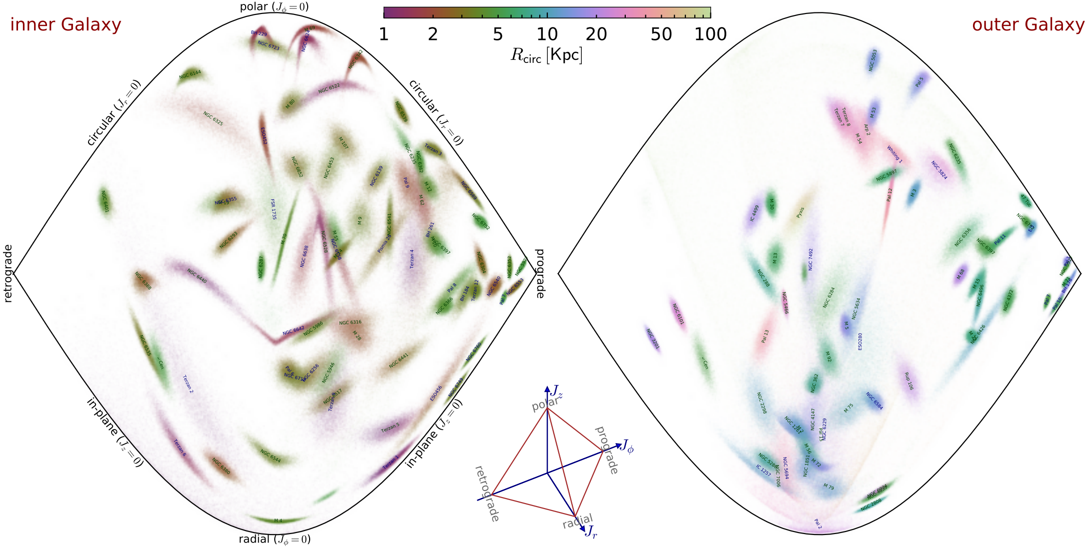
Velocity dispersion and rotation profiles

Main kinematical features of the entire population of globular clusters:

- Significant overall rotation, especially within the central 10 kpc (more prominent for metal-rich clusters).
- Nearly isotropic dispersion at $r < 10$ kpc, more radially anisotropic in outer parts; a population of ~ 10 clusters on eccentric orbits [Myeong+ 2018].
- Correlated orbits (e.g., Sgr stream: M 54, Terzan 7, Terzan 8, Arp 2, Pal 12, Whiting 1).



Distribution of globular clusters in action space



Dynamical modelling of the entire globular cluster population

Assume an equilibrium distribution function (in action space):

$$f(\mathbf{J}) = \frac{M}{(2\pi\, J_0)^3} \left[1 + \left(\frac{J_0}{h(\mathbf{J})}\right)^{\eta}\right]^{\Gamma/\eta} \left[1 + \left(\frac{g(\mathbf{J})}{J_0}\right)^{\eta}\right]^{-\mathrm{B}/\eta} \left(1 + \tanh \frac{\varkappa J_\phi}{J_r + J_z + |J_\phi|}\right),$$

$$g(\mathbf{J}) \equiv g_r J_r + g_z J_z + (3 - g_r - g_z)\,|J_\phi|, \quad h(\mathbf{J}) \equiv h_r J_r + h_z J_z + (3 - h_r - h_z)|J_\phi|,$$

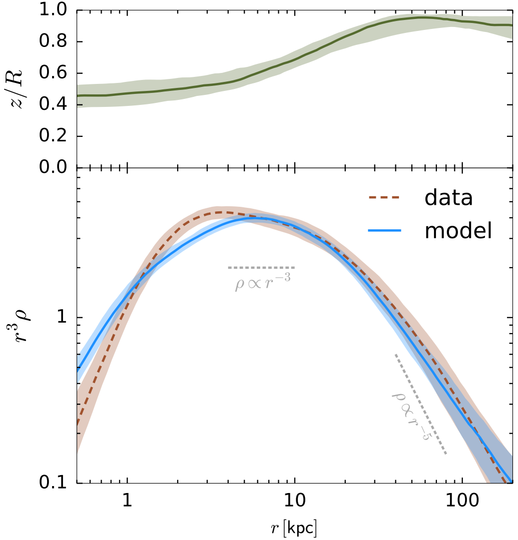
and a potential – bulge, disk and a flexible halo profile:

$$\rho(r) = \rho_\mathsf{h} \left(\frac{r}{r_\mathsf{h}}\right)^{-\gamma} \left[1 + \left(\frac{r}{r_\mathrm{h}}\right)^{\alpha}\right]^{(\gamma-\beta)/\alpha}.$$

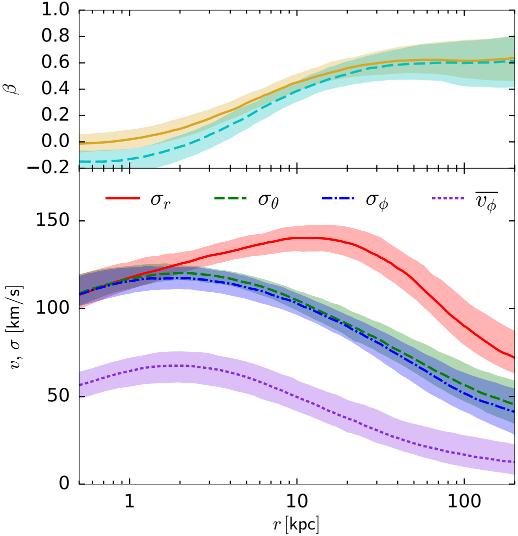
Maximize the likelihood of drawing the observed positions and velocities of clusters (taking into account their uncertainties) by varying the parameters of potential and DF.

Results: distribution of clusters

flattening and density profile



velocity dispersion and anisotropy



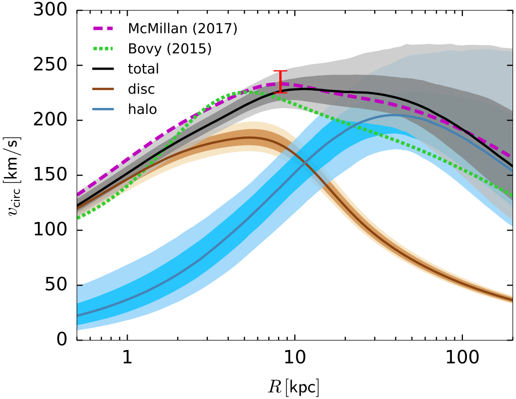
Results: constraints on the Milky Way potential

Results are broadly consistent with other studies based on globular clusters [Binney&Wong 2017, Sohn+ 2018, Watkins+ 2018, Posti&Helmi 2018, Eadie&Juric 2018];

the potential from McMillan(2017) is acceptable, the one from Bovy(2015) has too low rotation curve.

Clusters should be combined with other dynamical tracers (dSph, halo stars) for a more robust inference on the potential.

circular velocity curve



Summary

- *Gaia* is an immense source of valuable data, and is complementary to other surveys
- It can be used to measure the internal kinematics of globular clusters (velocity dispersion, rotation)
- Motion of globular clusters in the Galaxy is an important probe of the Milky Way potential