The Velocity Distribution of Nearby Stars from Hipparcos Data: The Significance of the Moving Groups

Jo Bovy
Center for Cosmology and Particle Physics
New York University

Bovy, Hogg, & Roweis (2009), ApJ, 700, 1794

Introduction

>>3D distribution of velocities of stars near the Sun (<~ 100 pc)

>>Long history:

>First account Mädler 1846: Stars moving together in the Pleiades

Introduction

>>3D distribution of velocities of stars near the Sun (<~ 100 pc)

>>Long history:

>First account Mädler 1846: Stars moving together in the Pleiades

ASTRONOMISCHE NACHRICHTEN.

№. 566.

Die Centralsonne.

Von Herrn Hofrath *Mädler*, Director der Sternwarte in Dorpat.

Vorwort des Herausgebers.

Introduction

>>3D distribution of velocities of stars near the Sun (<~ 100 pc)

>>Long history:

>First account Mädler 1846: Stars moving together in the Pleiades

>Proctor 1869: several groups of comoving stars (Hyades & Ursa Major)

>20th century: moving groups confirmed; new, but often spurious, groups discovered (e.g., Eggen)

>*Hipparcos*: complete samples --> first detailed studies of f(**v**) (e.g., Chen et al. 1997; Figueras et al. 1997; Dehnen 1998; Chereul et al. 1998, 1999; Asiain et al. 1999; Bienayme 1999; Skuljan et al. 1999)

>>Question remained: - How much of the structure is significant

>>We answer this question by modeling the underlying f(**v**) with varying degrees of complexity and finding the "best" model

Relation to *Gaia*

>>This is an exercise in *model selection*: What do we mean by "best"?

>>*Gaia* dynamical modeling questions:

- How complex does the DF need to be

- How complex does the potential need to be

>>Can we the 3D-kinematics from 2D-observations? <-> no RV at the faint end of *Gaia*

Modeling the underlying distribution function I

>>We deconvolve the velocity distribution by fitting for an underlying distribution

>>We model this underlying distribution as a mixture (sum) of *K* multivariate Gaussian components

$$f(\mathbf{v}) = \Sigma\, a_j\, N(\mathbf{v} \mid \mathbf{m}_j, \mathbf{V}_j) \quad ; \mathbf{m} = \text{mean}, \mathbf{V} = \text{variance tensor}$$

>>Fitting consists of comparing the model predictions with the data: Convolve the model with the observational uncertainties

$$\text{Likelihood} = \int d\, \mathbf{v}_{\text{true}}\, f(\mathbf{v}_{\text{true}})\, p(\mathbf{v}_{\text{obs}} \mid \mathbf{v}_{\text{true}})$$

obs. uncertainty

$$= \Sigma\, a_j\, N(\mathbf{v}_{\text{obs}} \mid \mathbf{R}_t\, \mathbf{m}_j, \mathbf{R}_t\, \mathbf{V}_j\, \mathbf{R}_t^T + \mathbf{S}_{\text{obs}})$$

Projection onto tangential comp.

Modeling the underlying distribution function II

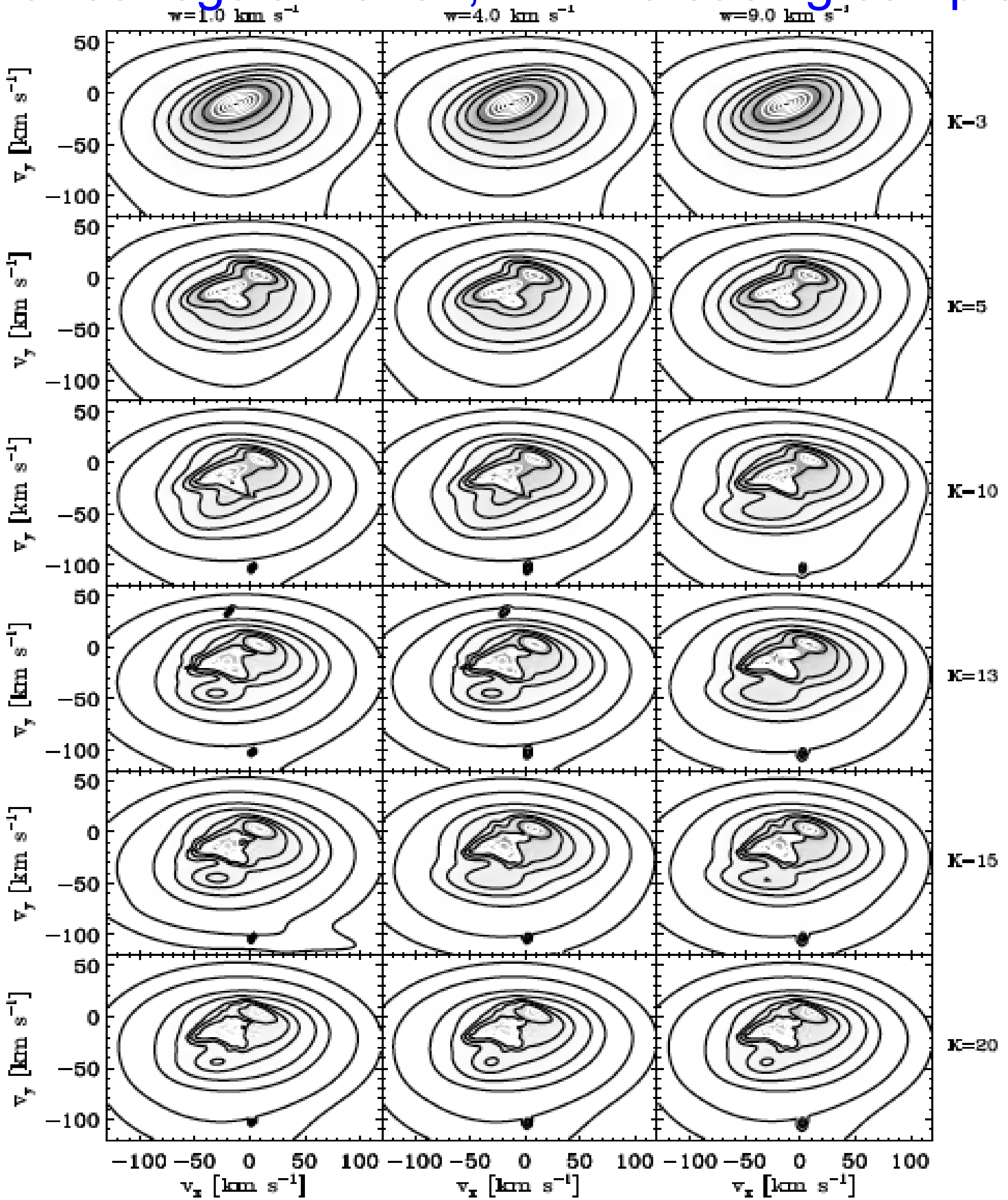
>>Fitting = optimize the total likelihood = Π (individual likelihoods)

>>We developed a general optimization technique to optimize this objective function, see Bovy, Hogg, & Roweis (2009) arXiv:0905:2979; GPLv2 code at http://code.google.com/p/extreme-deconvolution/

>>We can incorporate prior information, e.g., regularization of the variance tensors

Variance regularization, <--- increasing complexity

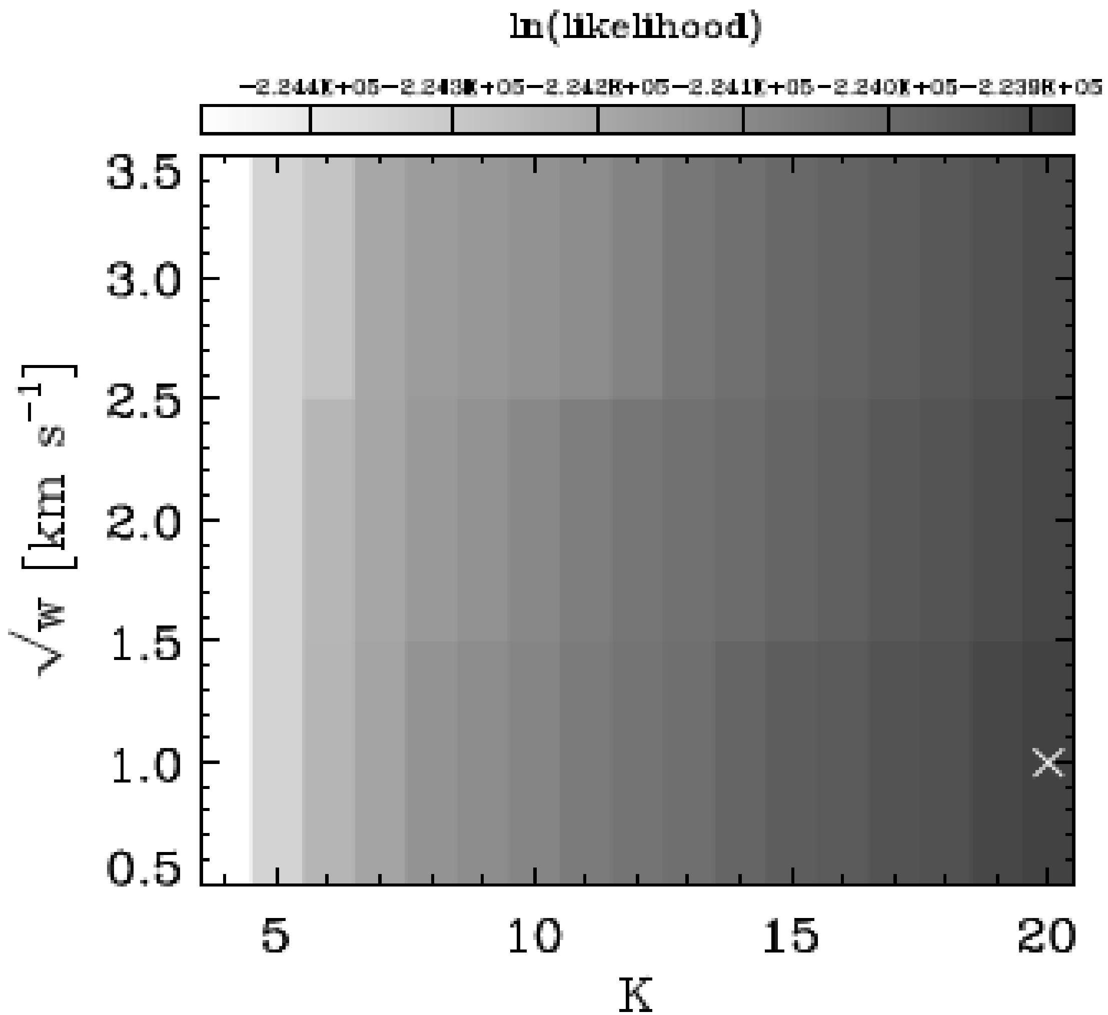
Direction of Gal. rotation



Number of Gaussian components
<---- Increasing complexity

Towards the GC

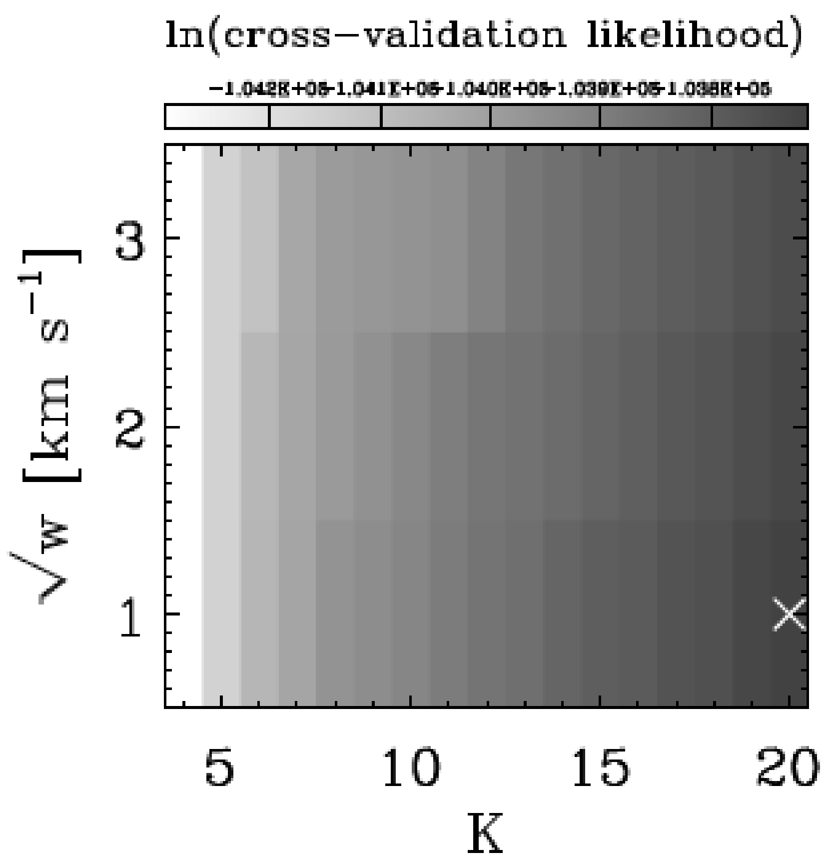
More complexity, better fit



But aren't we *overfitting* the data?

Choosing K (and w)

>>Cross-validation

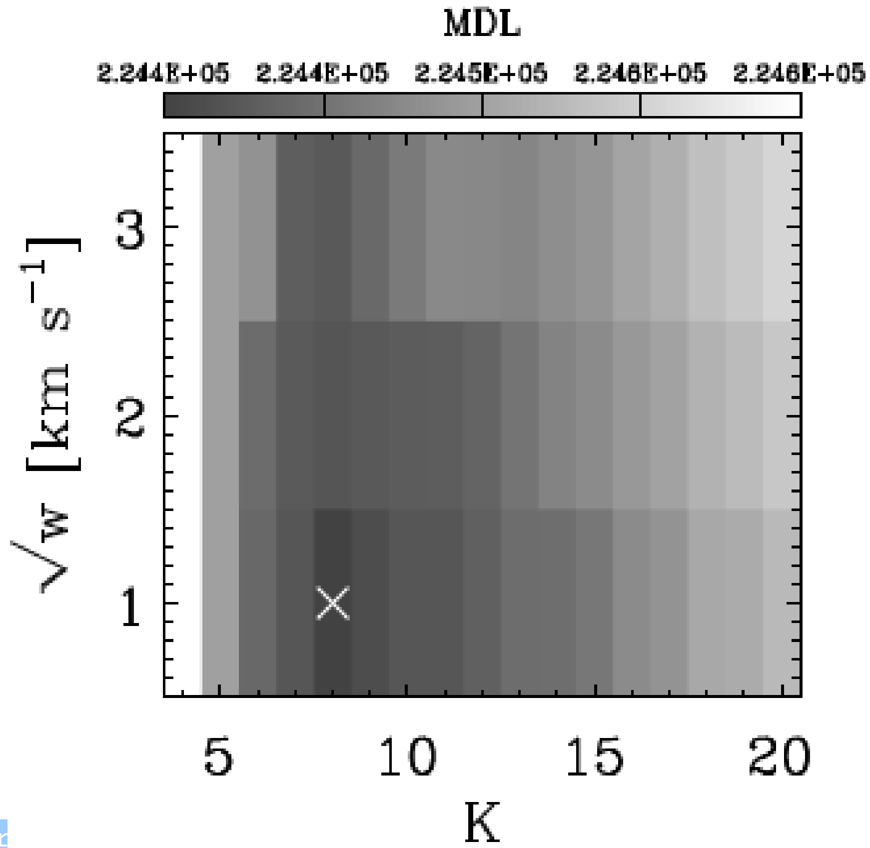


Choosing K (and w)

>>Cross-validation

>>Information-based criteria: minimum message length:
"Best" model is the model that compresses the data most

Could be part of dynamical modeling (e.g., made-to-measure)



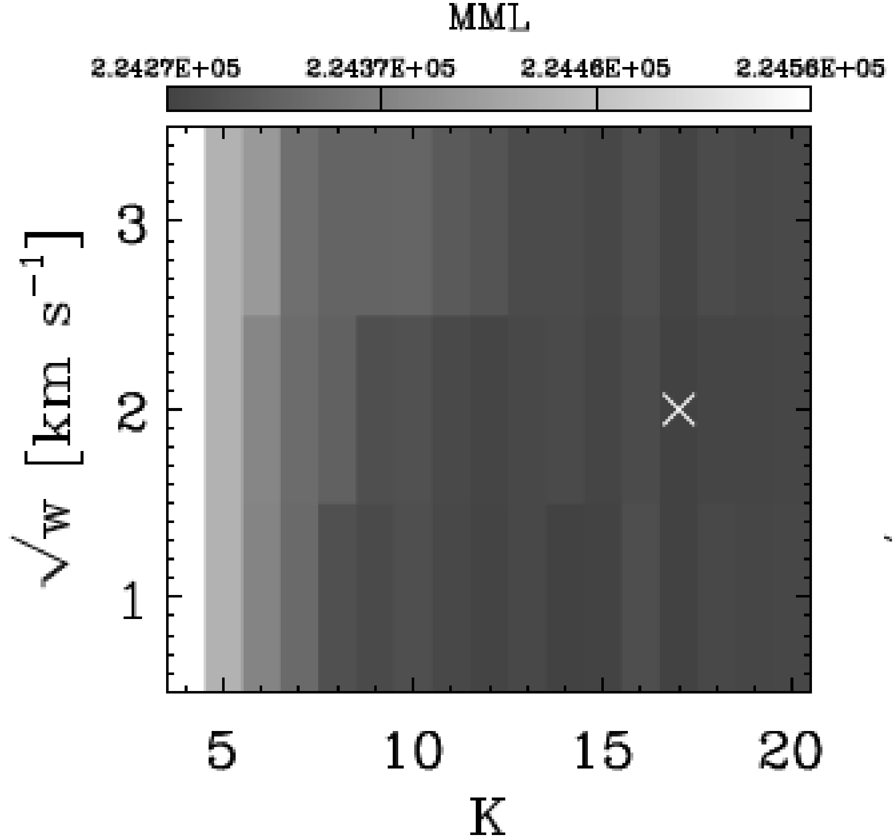
Choosing K (and w)

>>Cross-validation

>>Information-based criteria: minimum message length:

"Best" model is the model that compresses the data most

Could be part of dynamical modeling (e.g., made-to-measure)



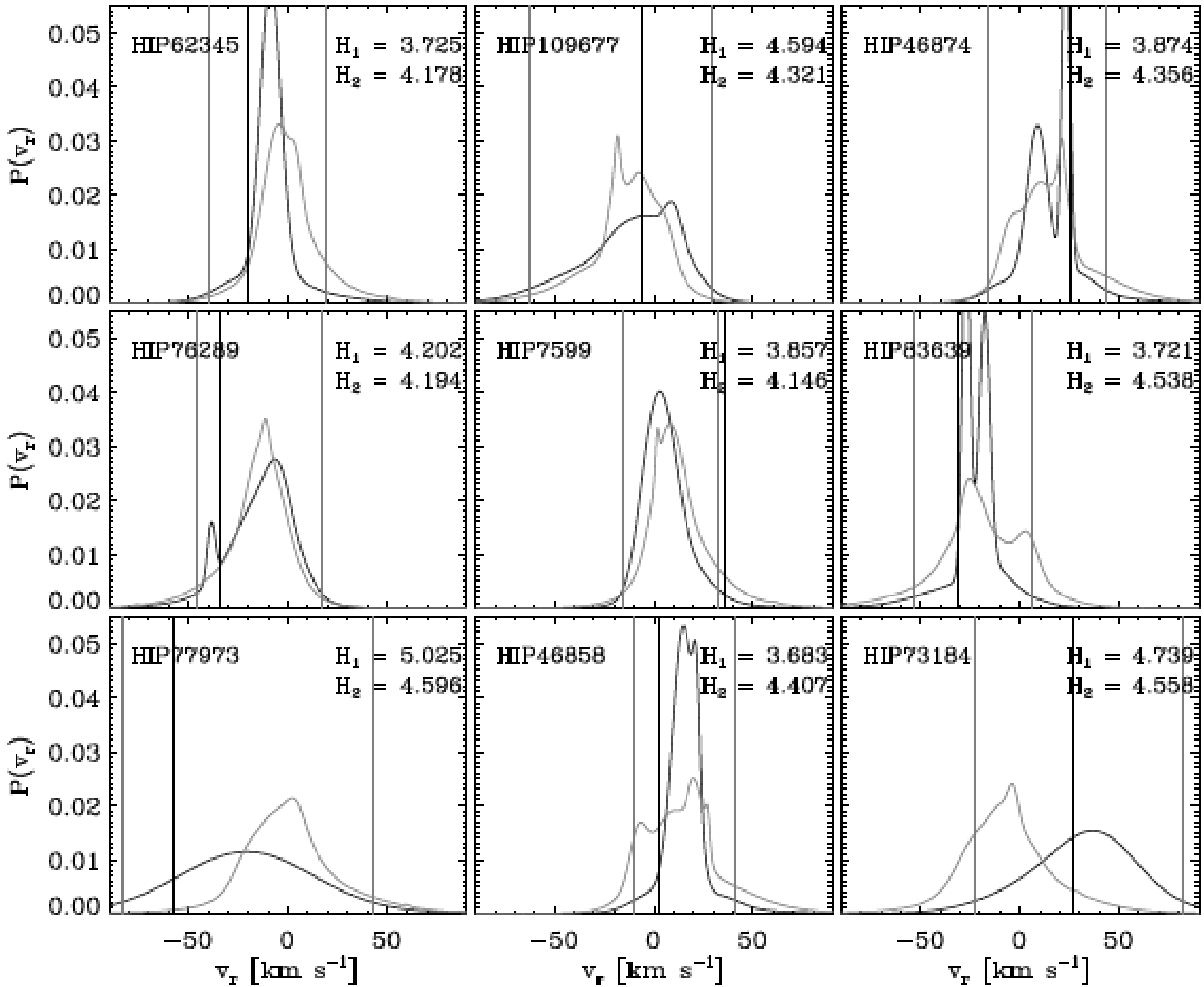
Choosing K (and w)

>>Cross-validation

>>Information-based criteria: minimum message length:
"Best" model is the model that compresses the data most

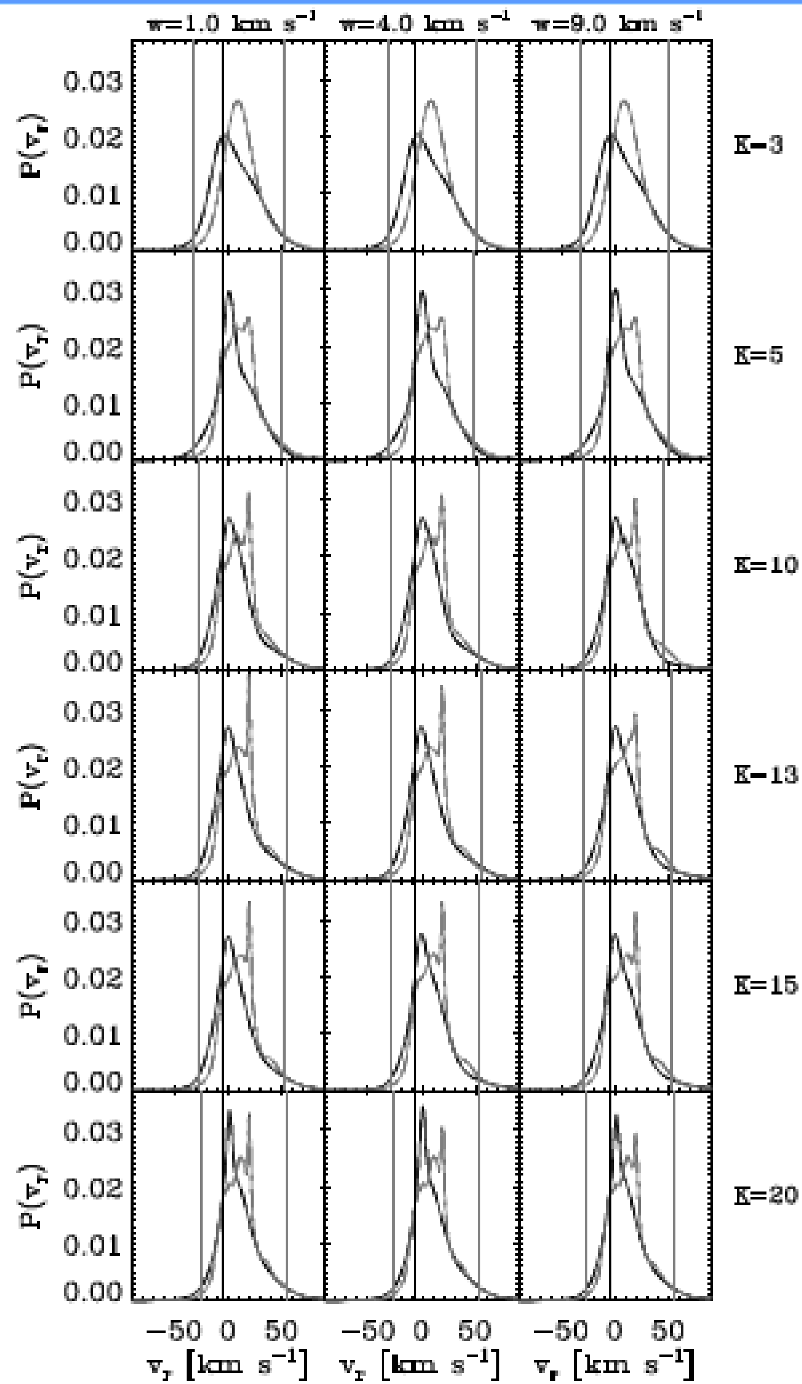
>>External validation: predict GCS radial velocities

Predicting the GCS radial velocities I



10 components
w = 4 (km/s)^2

Predicting the GCS radial velocities II

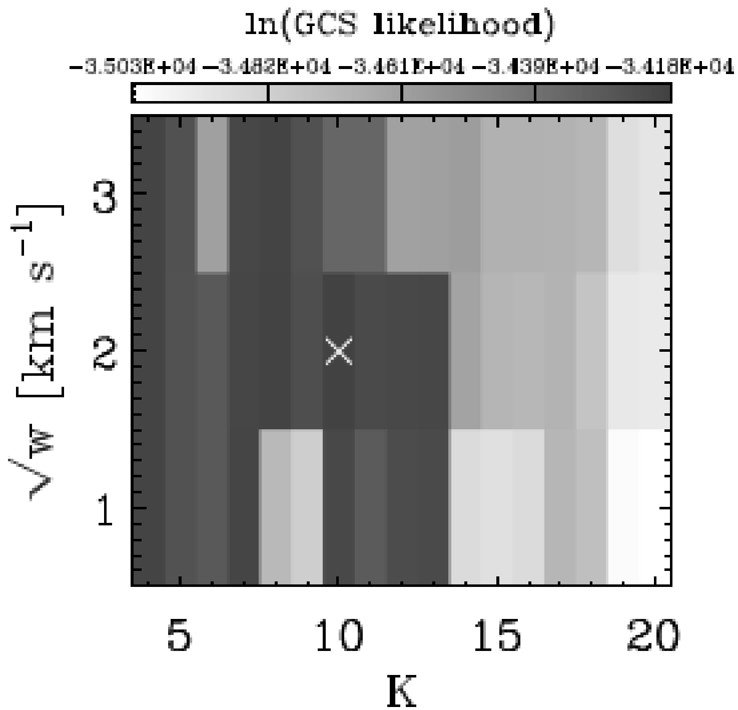


Choosing K (and w)

>>Cross-validation

>>Information-based criteria: minimum message length:
"Best" model is the model that compresses the data most

>>External validation: predict GCS radial velocities:
"Best" model is the model that predicts the RVs best



Choosing K (and w)

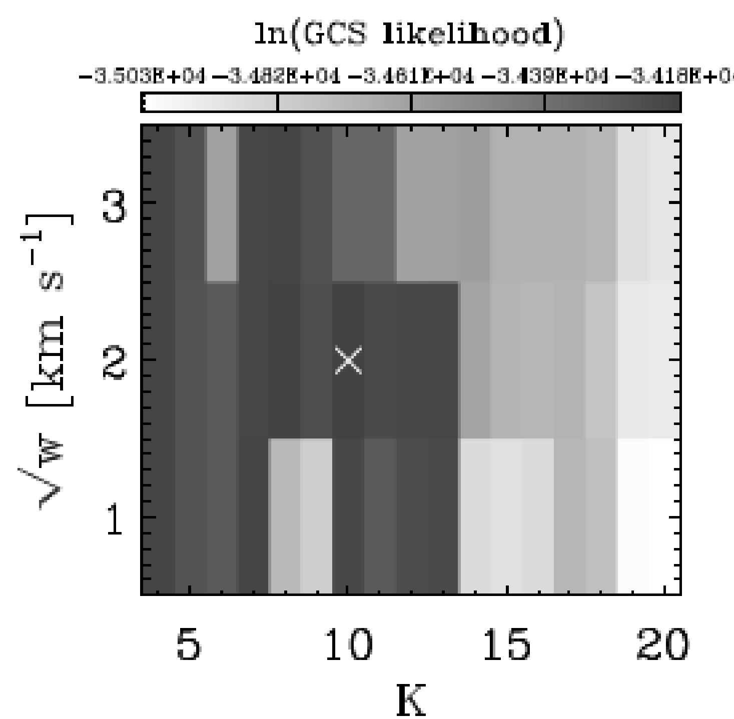
>>Cross-validation

>>Information-based criteria: minimum message length:
"Best" model is the model that compresses the data most

>>External validation: predict GCS radial velocities:
"Best" model is the model that predicts the RVs best

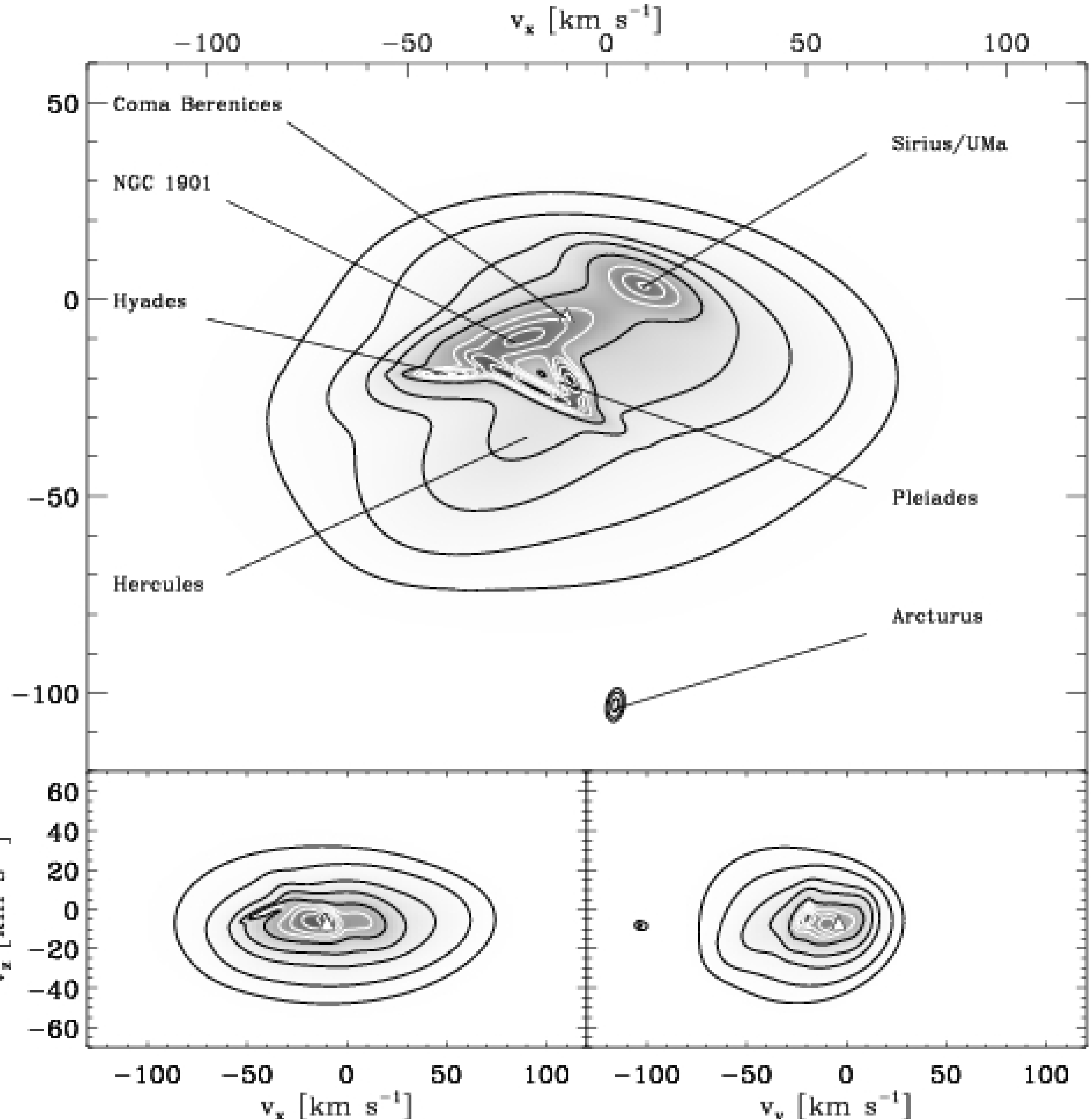
Most conservative

--> K = 10 components



Resulting velocity distribution

Toward the GC



Gal. rotation direction

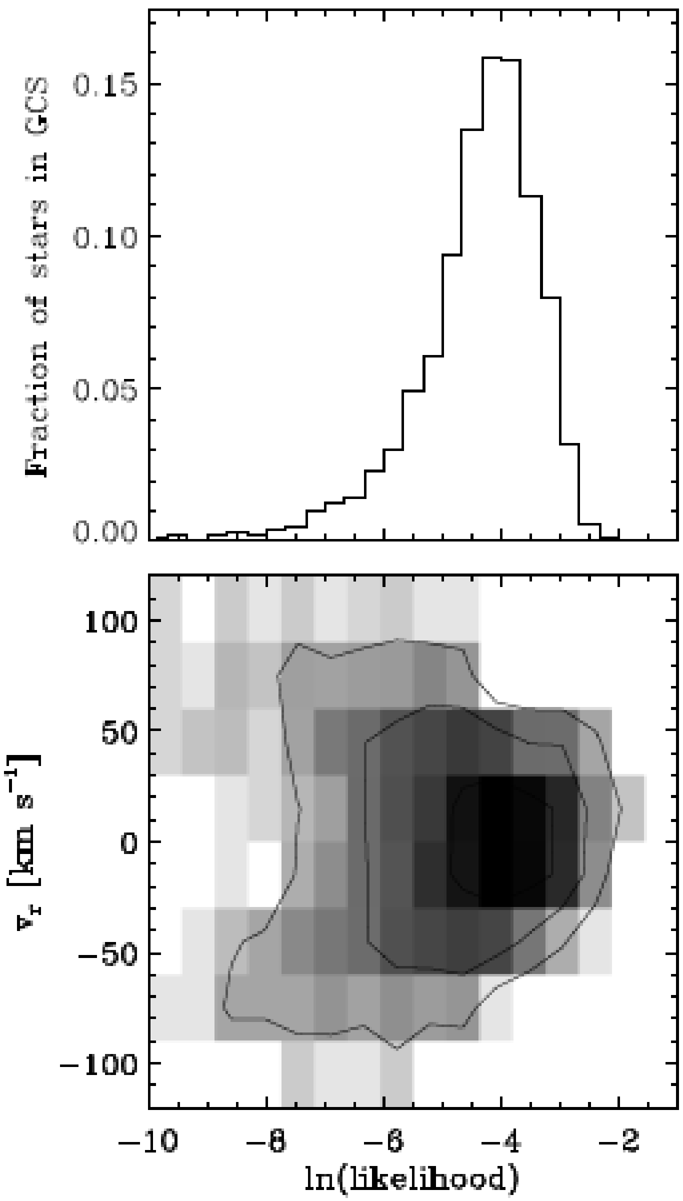
Toward the NGP

Toward the GC

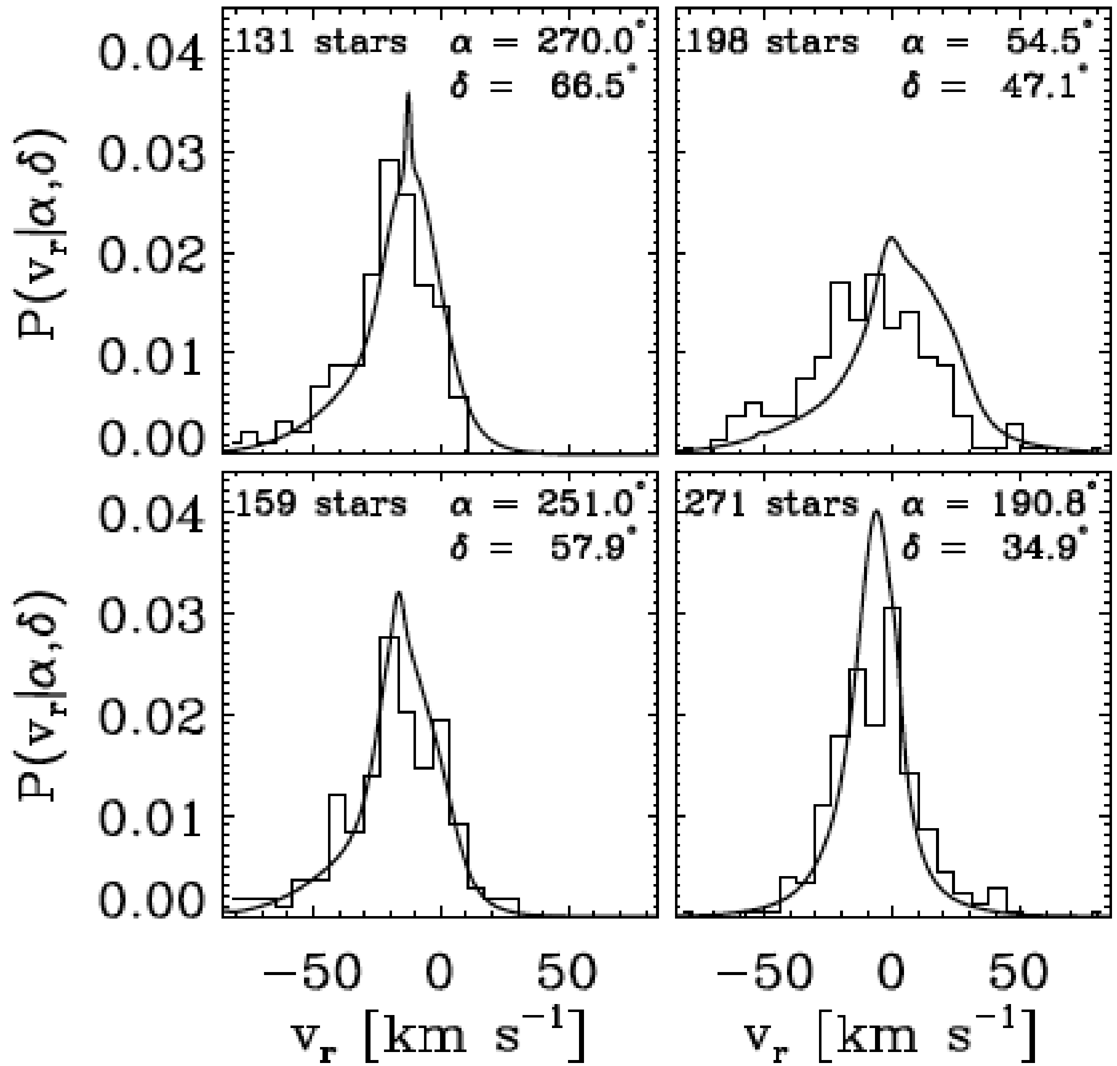
In the direction of Gal. rotation

Good fit?

GCS probabilities:



radial velocity distribution in patches of the sky:



Conclusions

>>Significant structure in the local velocity distribution was all known before *Hipparcos* (but the velocity distribution is nevertheless very complex)

>>We can reconstruct the local velocity distribution from tangential velocities alone

>>Deconvolution (or, more generally, forward modeling) gets the most out of the data (and no more) + it handles arbitrary uncertainties, including missing data

>>But, these approaches should mature to sampling the uncertainty in the DF, especially when modeling the DF in dynamical modeling, in order to marginalize over the DF when inferring the Galactic potential