# Probabilistic Approach to Finding Geometric Objects in Spatial Datasets of the Milky Way

Jon Purnell
Heidi Jo Newberg
Malik Magdon-Ismail

Rensselaer Polytechnic Institute

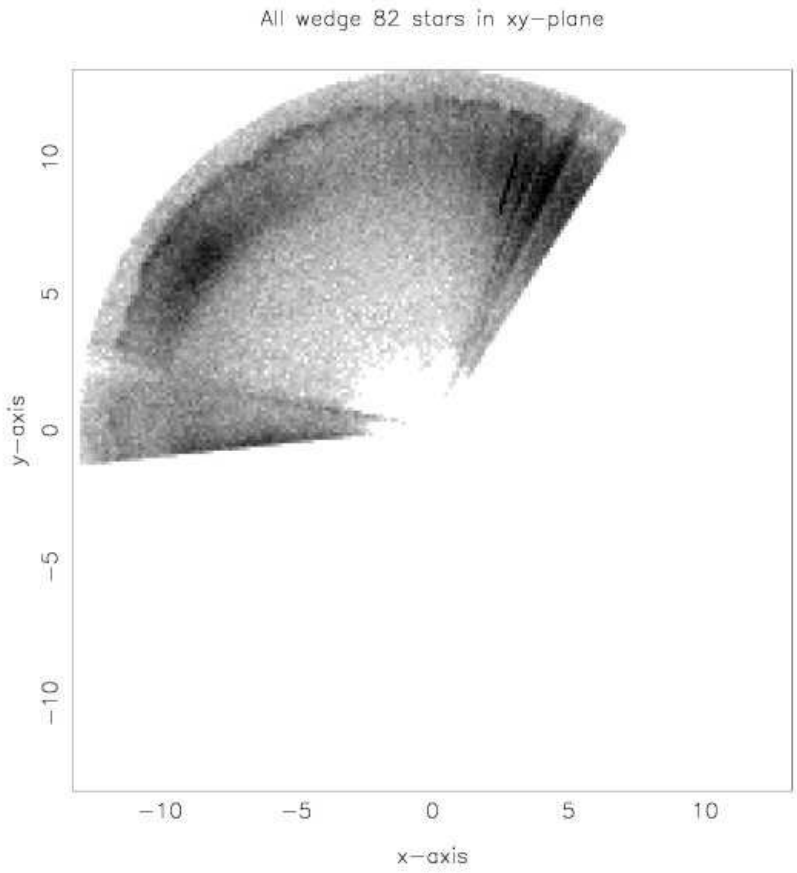
ISMIS 2005

# Motivation

- Simultaneously fit multiple geometric distributions
- Use distributions with varying complexity
- Automatic parameter optimization
- No need to filter data

# Real Data

- Applied algorithm to 2.5 degree wide wedge along the celestial equator from SDSS dataset.



# Galactic (Background) Distribution

- Power Law

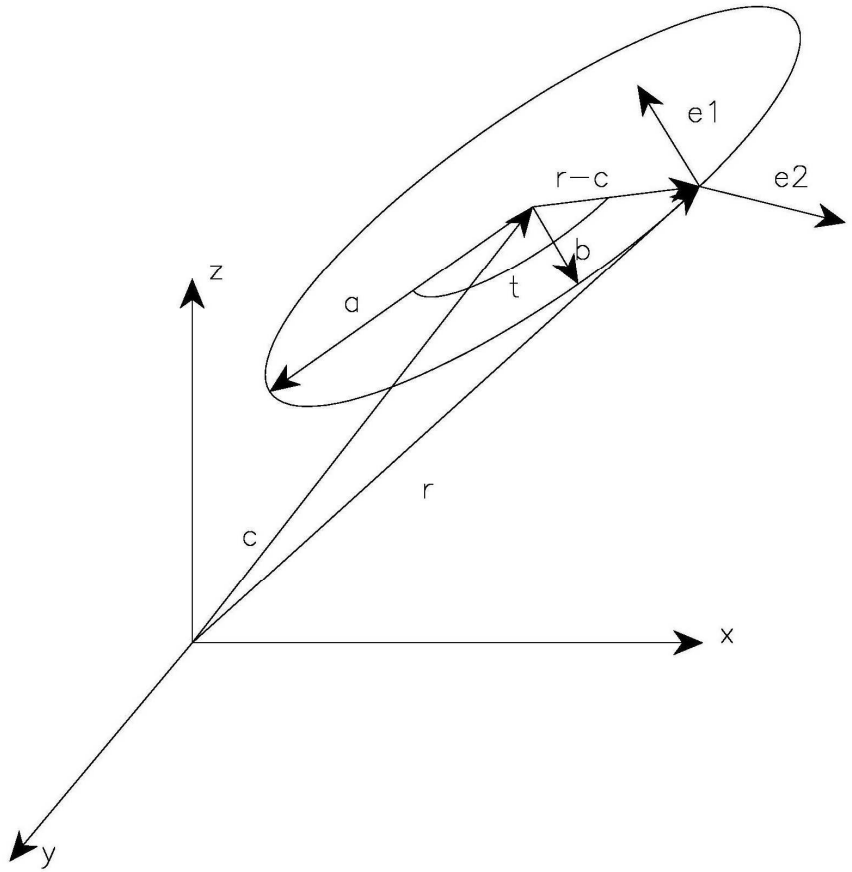
$$P_b(x,y,z) = \frac{1}{r^\alpha} \qquad r = \sqrt{x^2 + y^2 + (z/q)^2}$$

- Hernquist Equation

$$P_b(x,y,z) = \frac{1}{r^\alpha (r + r_o)^{3-\alpha+\delta}}$$

# Tidal Stream Distribution

- An ellipse with a 2-d Gaussian cross-section



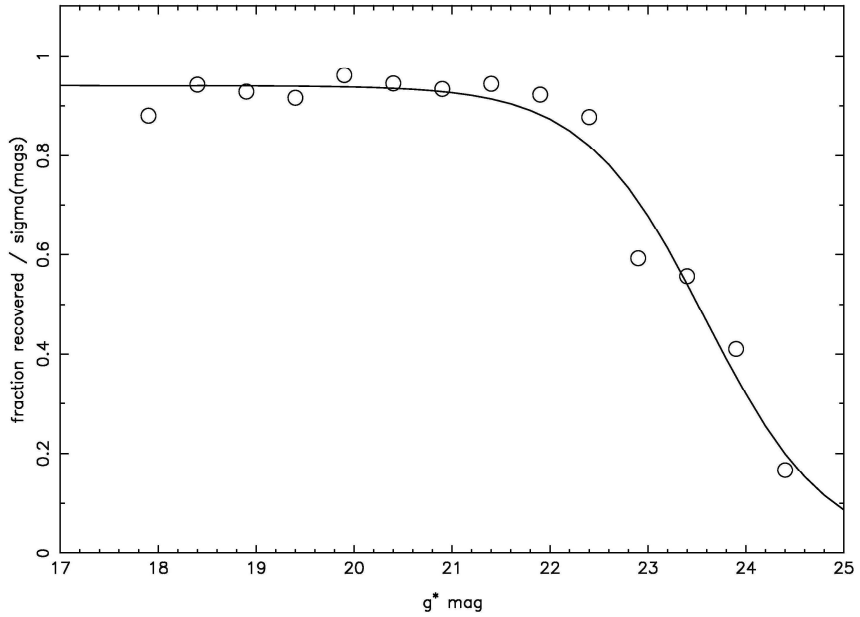
# Model Distribution

- Mixture of Background and Stream distributions
- Numerical Integration over wedge

$$P_m(\mathbf{x}) = \epsilon \frac{P_b'(\mathbf{x})}{\int P_b'} + (1-\epsilon)\frac{P_s'(\mathbf{x})}{\int P_s'}$$

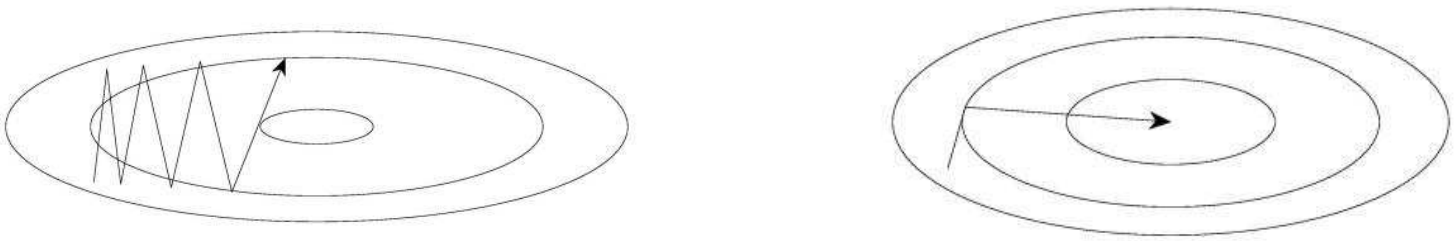
# Data Efficiency

- Efficiency – ratio of the number of stars in dataset to the actual number of stars in the galaxy



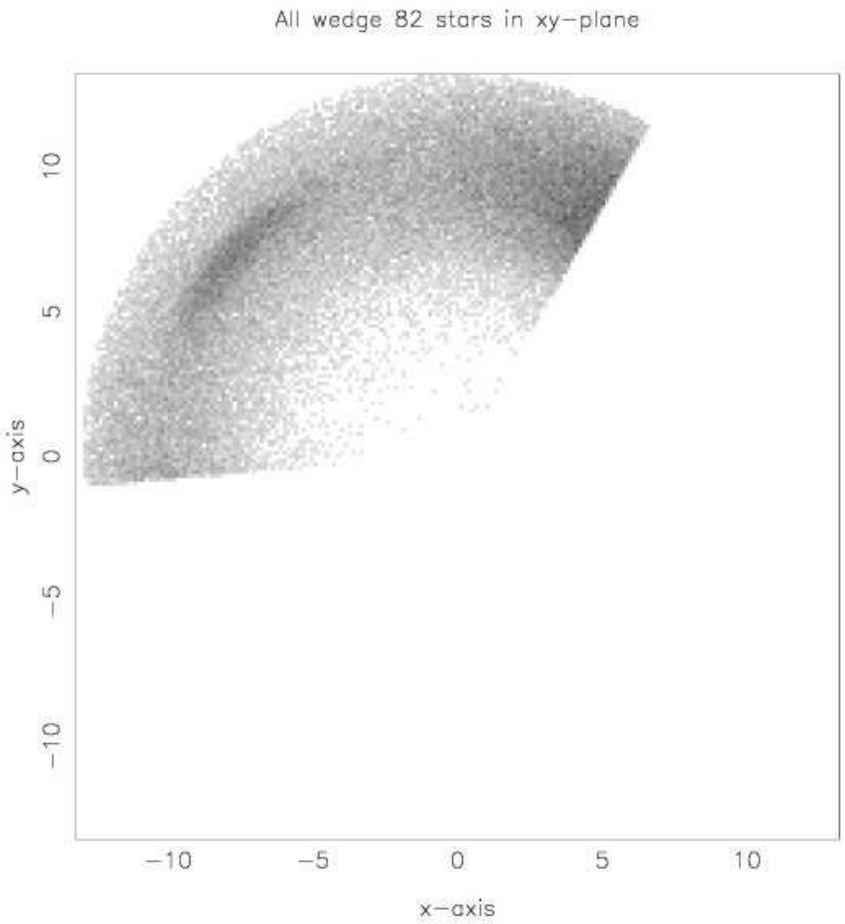
# Parameter Optimization

- Reduce parameter set
- Unconstrain parameters
- Conjugate gradients

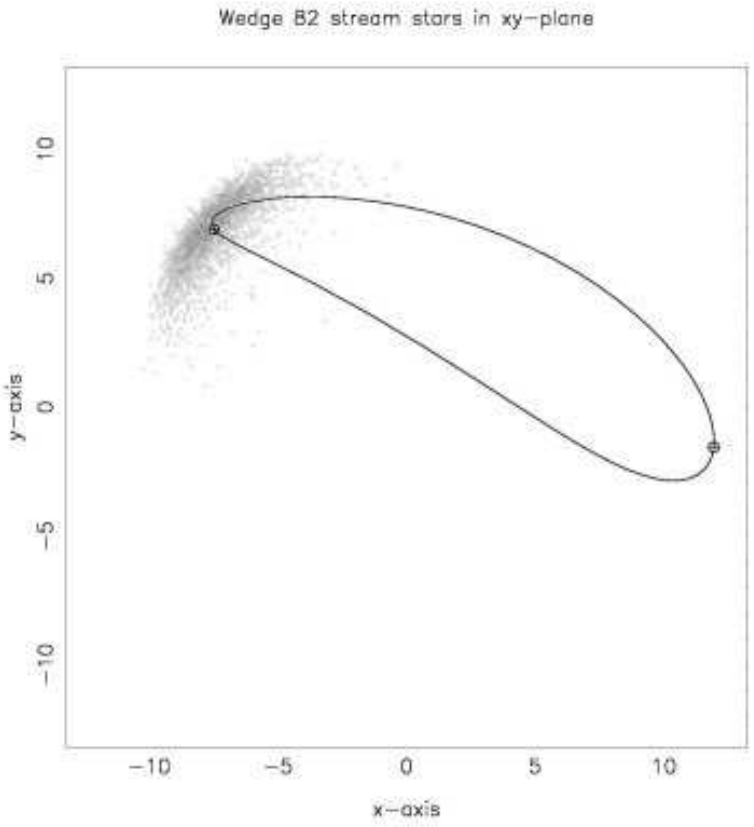


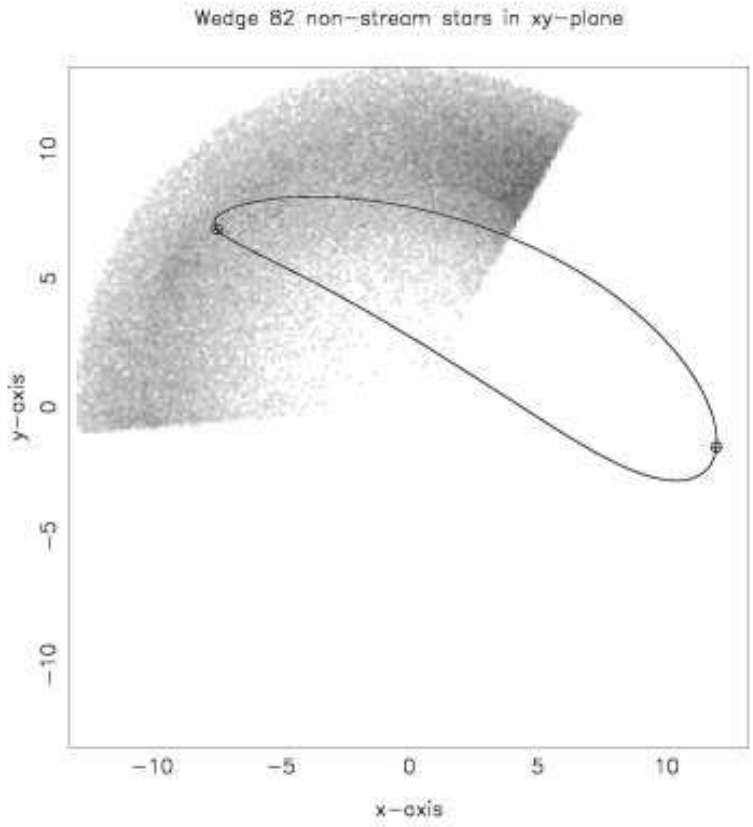
# Synthetic Data

- Generate data using mixture model



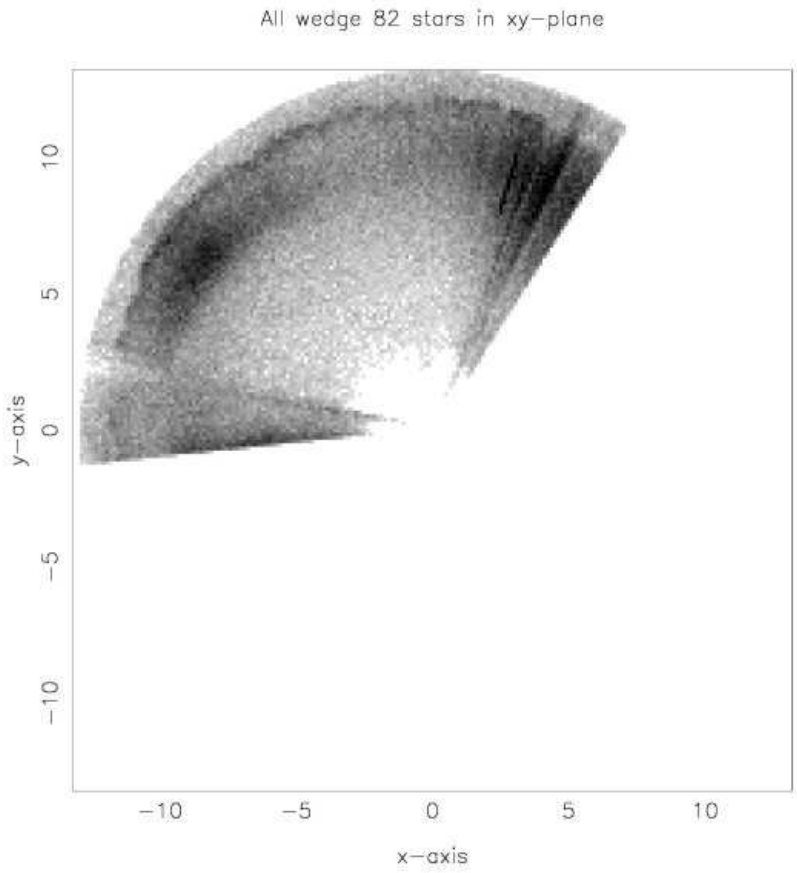
# Synthetic Data Results



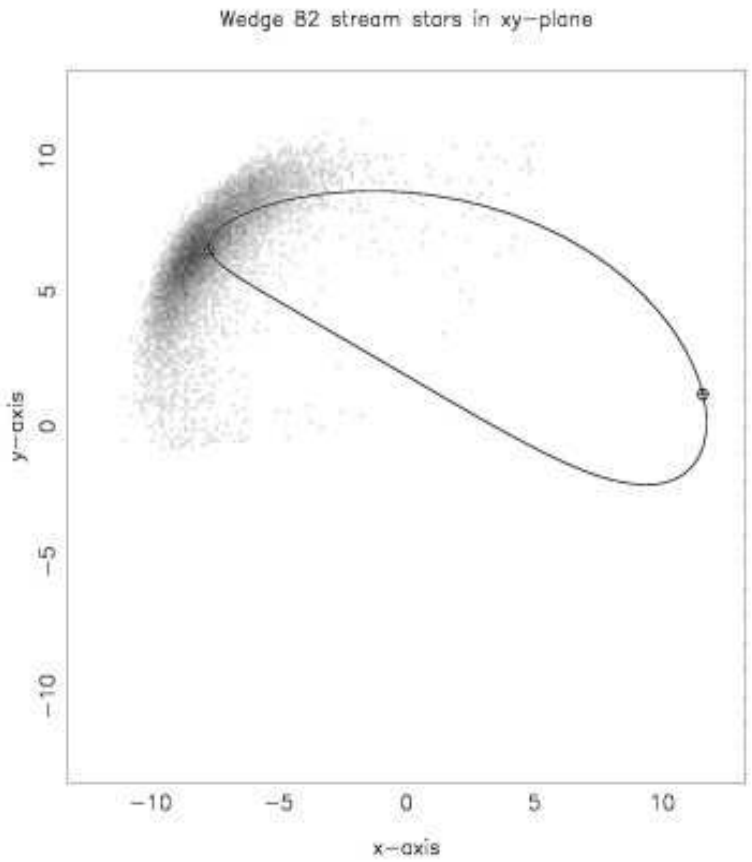



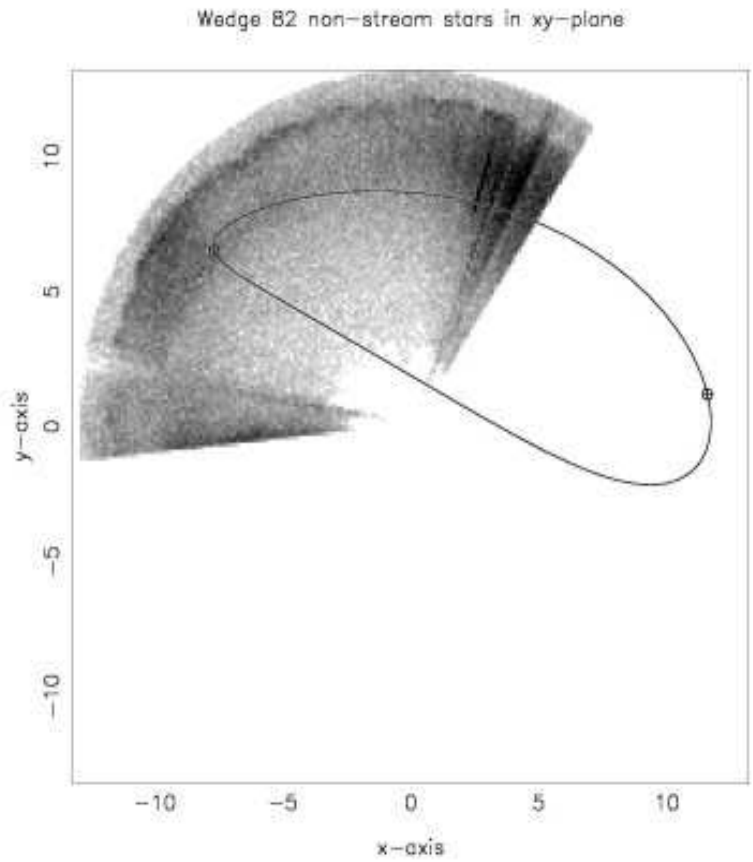
# Real Data

- Applied algorithm to 2.5 degree wide wedge along the celestial equator from SDSS dataset.



# Real Data Results



# Future Plans

- Apply different distributions for halo stars
- Search for multiple streams
- Search for other structures
- Search over multiple ‘wedges’ simultaneously

# Questions

?