Bayesian Mass Estimates of the Milky Way: incorporating incomplete data

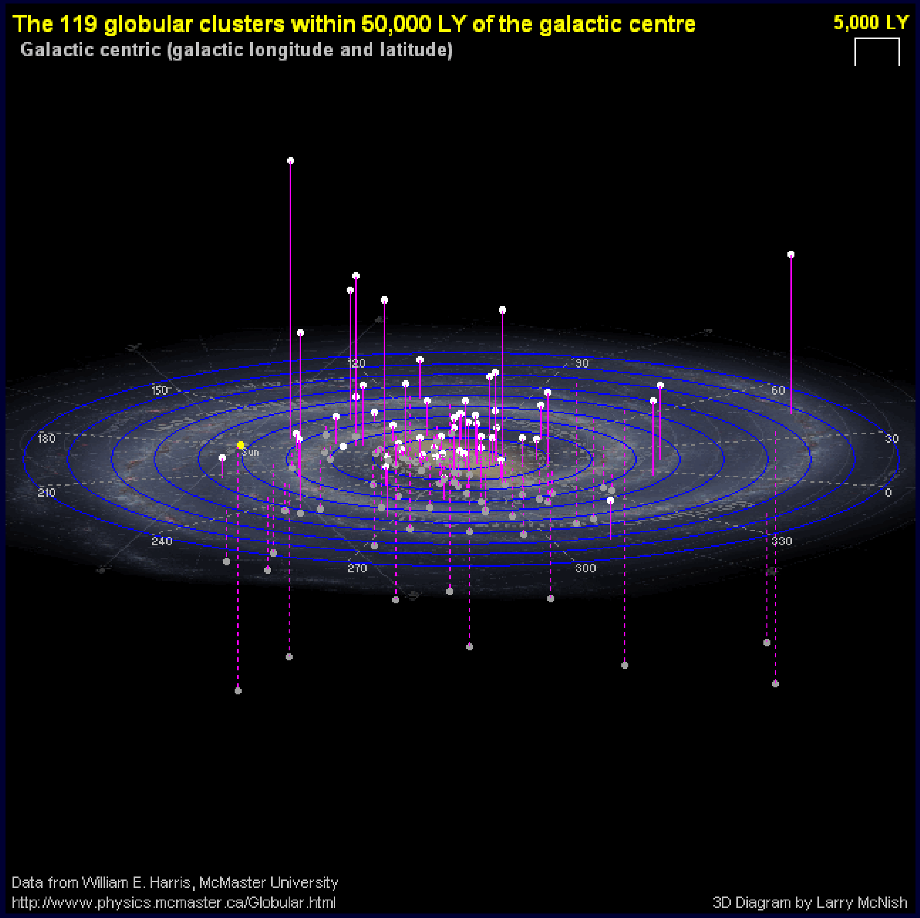
McMaster University

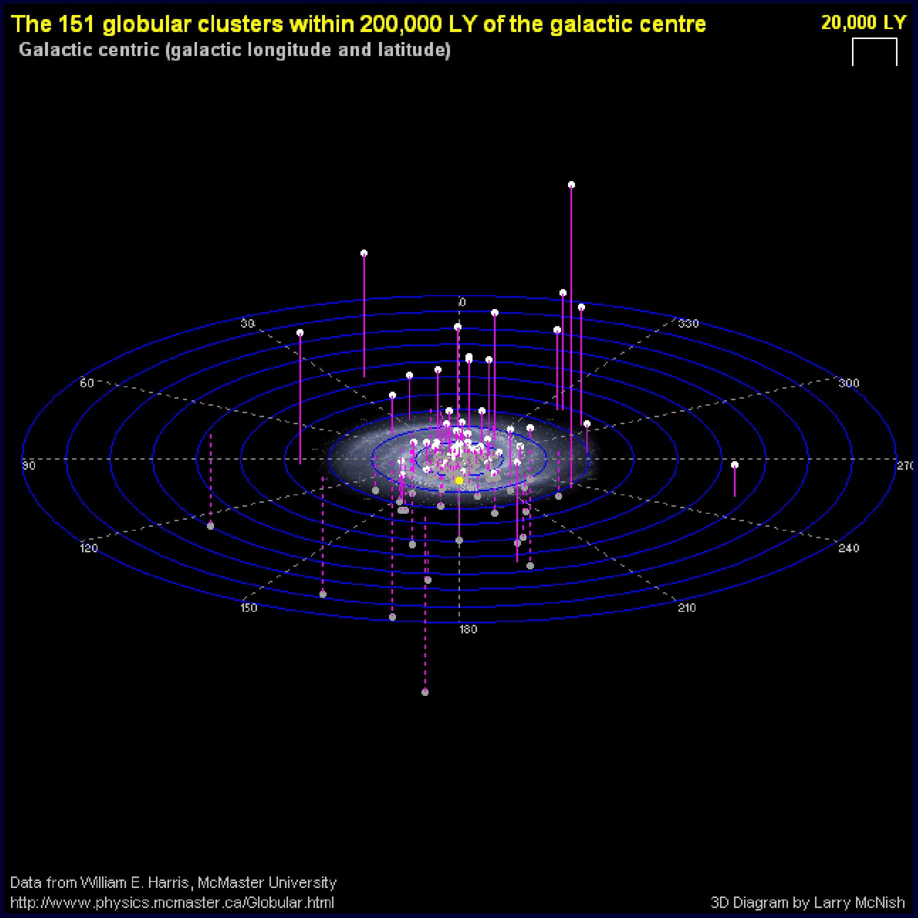
Gwendolyn Eadie, PhD Candidate
Dr. William Harris (McMaster), Dr. Lawrence Widrow (Queen's)
AstroStat Seminar - March 23, 2015

Measuring Mass

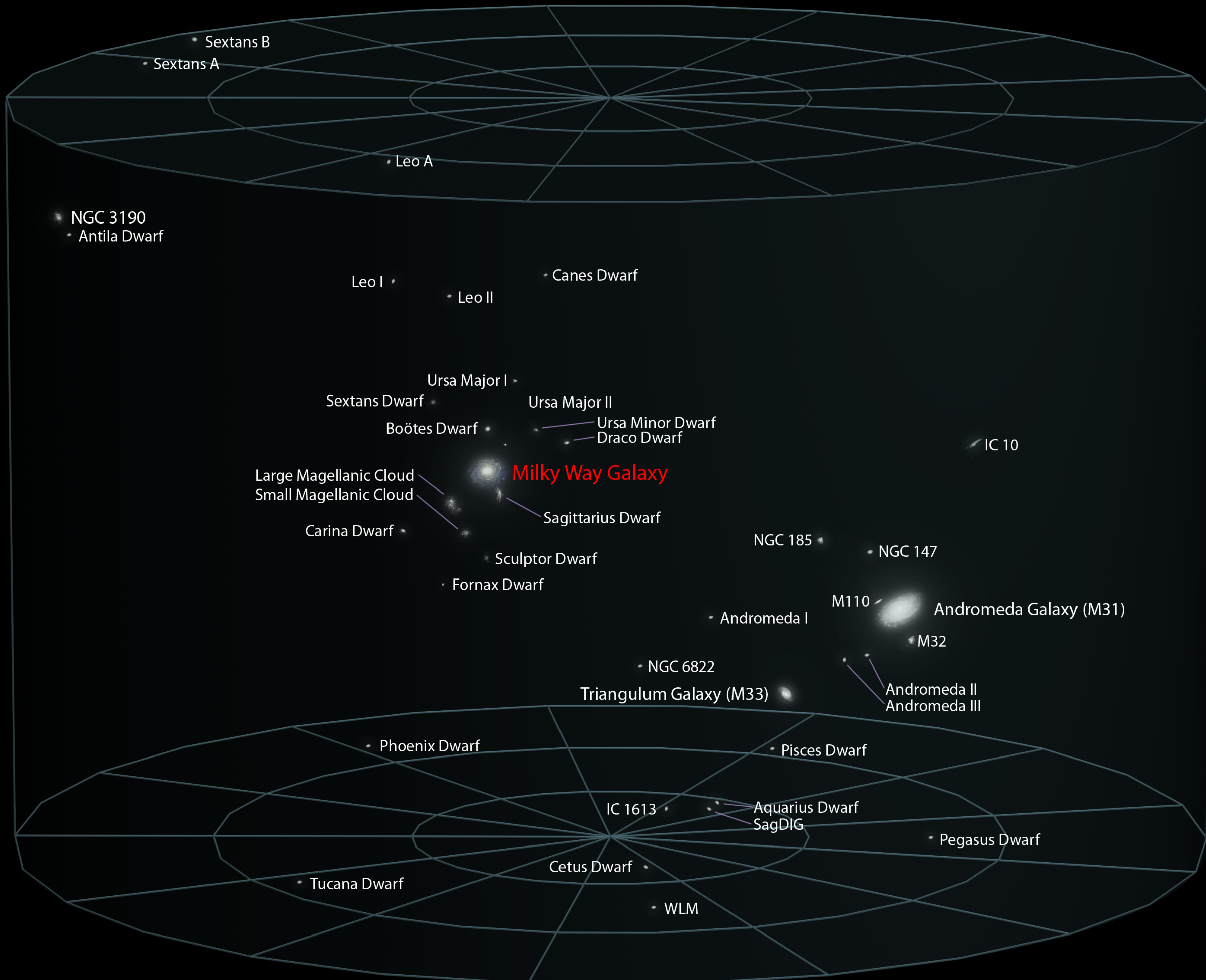
- Motivation
  - Mass-luminosity relationships
  - Globular Cluster (GC) population studies
  - Dark matter halos
  - Compare to cosmological simulations
- Observed Satellites
  - GC
  - dwarf galaxies
  - planetary nebulae
  - halo stars

Globular Cluster distribution



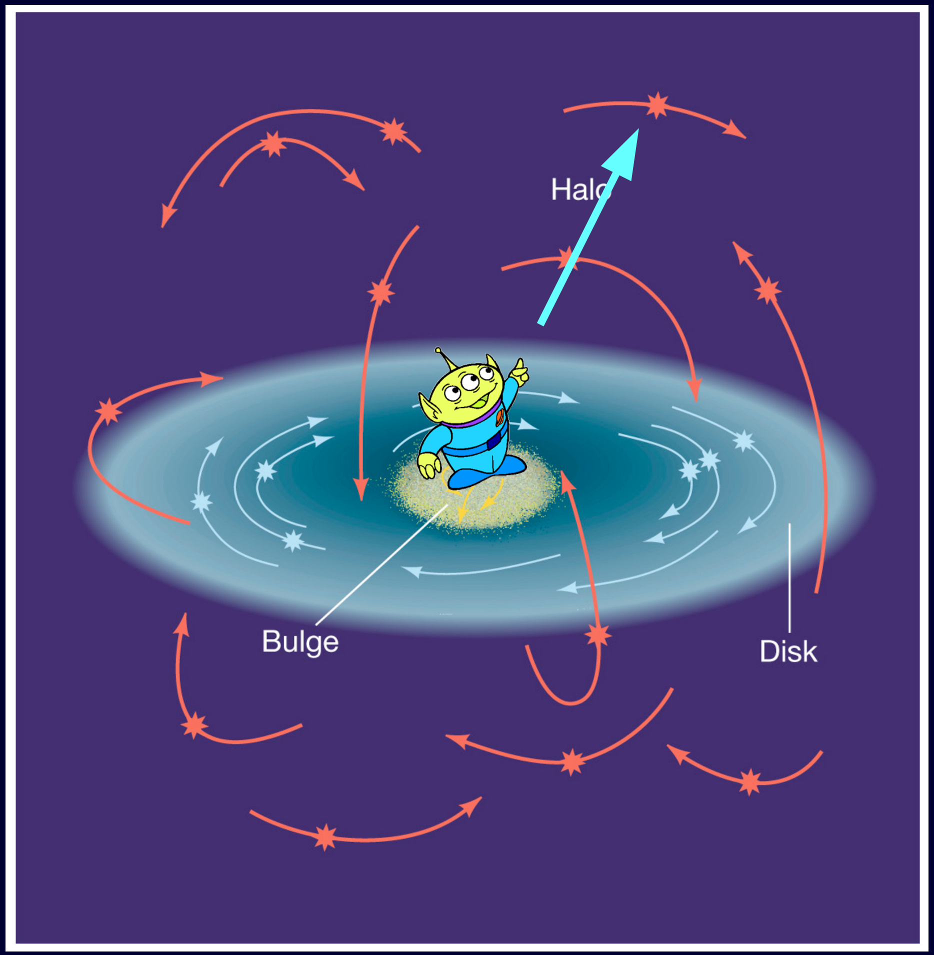


http://calgary.rasc.ca/globulars.htm



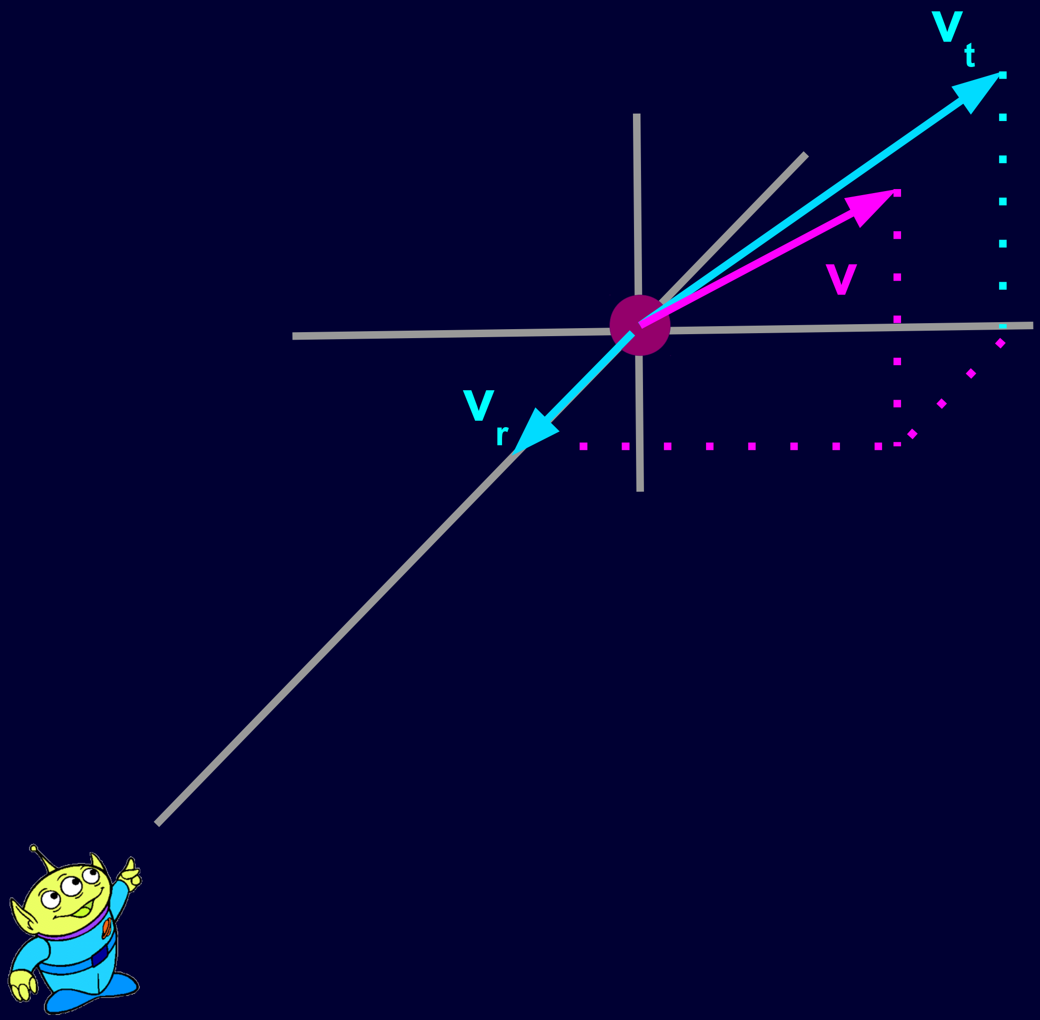
Wiki Commons, author: Andrew Z. Colvin

Galactocentric Measurements



Chaisson & McMillan, *Astronomy*, 2004

Galactocentric Velocities



- Kinematic data
  - v_r radial velocity
  - v_t tangential velocity
  - r distance

Galactocentric *vs* Heliocentric Reference Frames

- Milky Way (MW) mass models are simplest to implement from Galactocentric point of view
- We have a combination of *heliocentric* data that is
  - **Complete** (known velocity vector)
  - **Incomplete** (missing proper motion component)

Galactocentric *vs* Heliocentric Reference Frames

- Milky Way (MW) mass models are simplest to implement from Galactocentric point of view
- We have a combination of *heliocentric* data that is
  - **Complete** (known velocity vector)
  - **Incomplete** (missing proper motion component)
- In the past, incorporating incomplete data into analyses meant using galaxy mass estimators that relied only on line of sight velocities.
- *Our method: use both complete and incomplete data simultaneously in the Galactocentric frame*

Bayesian method:
incorporate both complete and incomplete data

- How this works
- Simulations and testing
- Preliminary application of method to the Milky Way

Eadie, Harris, & Widrow (2015), Astrophysical Journal
(in press, posted to astro-ph in the next couple days)

Using Bayes' Theorem

- Little & Tremaine (1987)
- Bayes' Theorem

$$p\left(\boldsymbol{\theta}|y\right) = \frac{p\left(y|\boldsymbol{\theta}\right)p\left(\boldsymbol{\theta}\right)}{p\left(y\right)}$$

Using Bayes' Theorem

- Little & Tremaine (1987)
- Bayes' Theorem

Likelihood prior

Probability of **model parameters** →

$$p\left(\boldsymbol{\theta}|y\right) \propto p\left(y|\boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right)$$

Using Bayes' Theorem

- Little & Tremaine (1987)
- Bayes' Theorem

Likelihood prior

Probability of **model parameters** →

$$p(\boldsymbol{\theta}|y) \propto p(y|\boldsymbol{\theta})\, p(\boldsymbol{\theta})$$

- Distribution Function: *probability* of finding a satellite with (**r**,**v**)

⇒ $f(r, v|\boldsymbol{\theta})$

Using Bayes' Theorem

- Little & Tremaine (1987)
- Bayes' Theorem

Likelihood prior

Probability of **model parameters** →

$$p\left(\boldsymbol{\theta}|y\right) \propto p\left(y|\boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right)$$

- Distribution Function: *probability* of finding a satellite with (**r**,**v**)

$$\Rightarrow f\left(\boldsymbol{r},\boldsymbol{v}|\boldsymbol{\theta}\right)$$

Likelihood:

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_i f\left(\boldsymbol{r_i},\boldsymbol{v_i}|\boldsymbol{\theta}\right)$$

Deriving the Distribution Function (DF)

- Relative energy:

$$\mathcal{E} = -\frac{v^2}{2} + \Psi(r)$$

- Model: potential, mass density, and mass profile

$\Phi(r)$ $\rho(r)$ $M(r)$

- write density as a function of relative potential

$\rho(\Psi)$

- solve an Abel transform (Binney & Tremaine)

Deriving the Distribution Function (DF)

- For isotropic cases:

$$f(\mathcal{E}) = \frac{\sqrt{2}}{4\pi^2}\frac{d}{d\mathcal{E}}\int_0^{\mathcal{E}} \frac{d\Psi}{\sqrt{\Psi - \mathcal{E}}}\frac{d\rho}{d\Psi}$$

- DF goes into the likelihood

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_i f(\boldsymbol{r_i}, \boldsymbol{v_i}|\boldsymbol{\theta})$$

Example: Hernquist Model

- potential, mass density, and mass profile

$$\Phi(r) = -\frac{M_{tot}}{r+a}$$

$$\rho(r) = \frac{aM_{tot}}{2\pi r\,(r+a)^3}$$

$$M(r) = M_{tot}\frac{r^2}{(r+a)^2}$$

parameters: M_{tot} , a

In the case of an isotropic velocity distribution:

$$f(q) = \frac{M_{tot}}{8\sqrt{2}\pi^3 a^3 v_g^3\,(1-q^2)^{5/2}}\left[3\arcsin(q) + q\sqrt{(1-q^2)}\,(1-2q^2)\,(8q^4-8q^2-3)\right]$$

Hernquist (1990), ApJ 356: 359-364.

$$q = \sqrt{\frac{a\mathcal{E}}{M_{tot}}}$$

$$v_g = \sqrt{\frac{M_{tot}}{a}}$$

Example Posterior Distribution

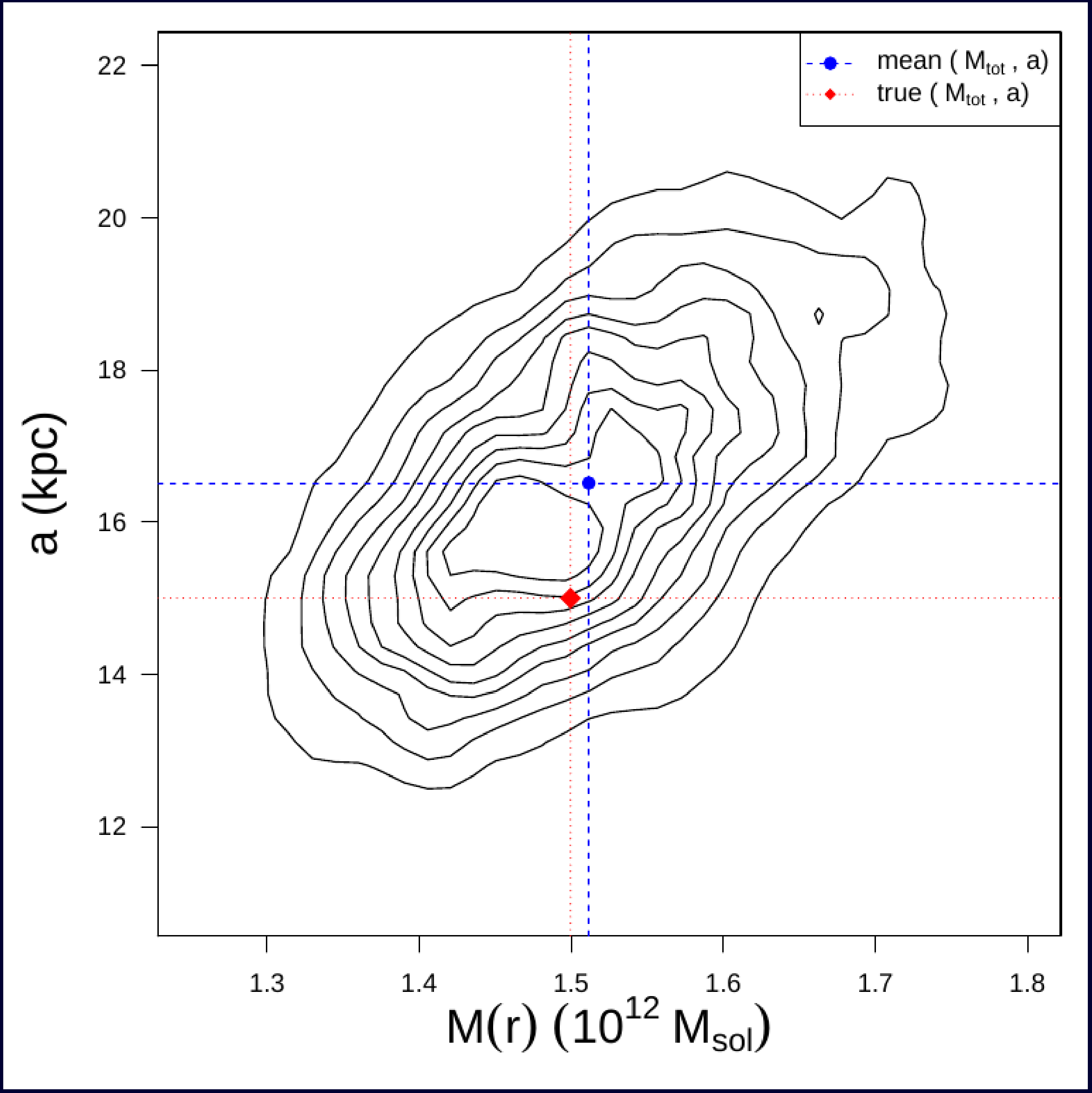
(Isotropic Hernquist model, simulated data)

Probability of **parameters** (posterior distribution)

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \prod_i f(\boldsymbol{r_i}, \boldsymbol{v_i}|\boldsymbol{\theta}) p(\boldsymbol{\theta})$$

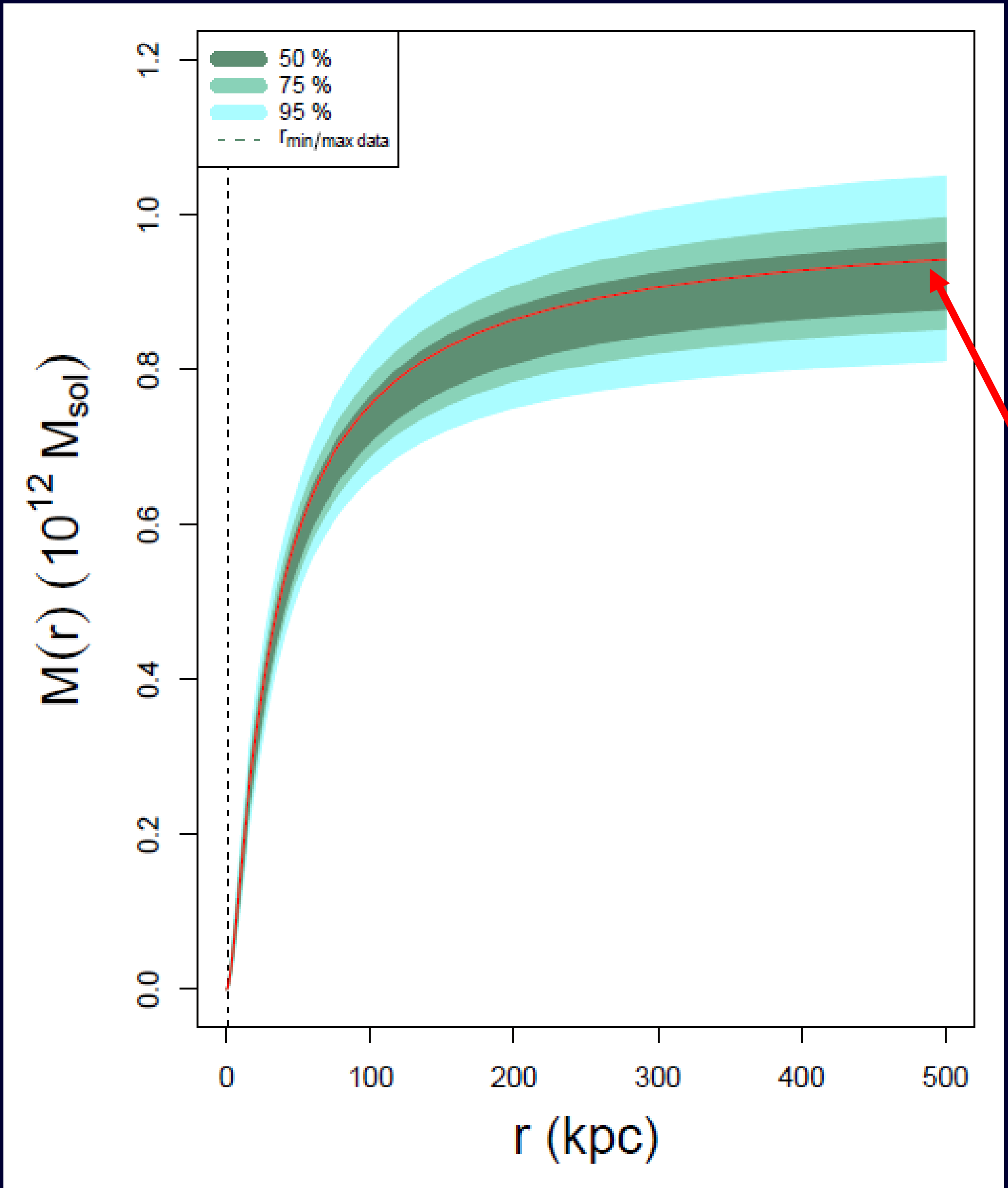
Sample using Metropolis algorithm ---> Markov Chain:

M_{tot} and a pairs



Example Cumulative Mass Profile

(Isotropic Hernquist model)



Credible Regions

true M(r)

Advantage of Bayesian Approach

- If $\mathbf{v}_t$ of the satellites are unknown

?

$$v^2 = v_r^2 + v_t^2$$

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \prod_i f(\boldsymbol{r_i}, \boldsymbol{v_i}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

Advantage of Bayesian Approach

- If $\mathbf{v}_t$ of the satellites are unknown

?

$$v^2 = v_r^2 + v_t^2$$

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \prod_i f(\boldsymbol{r_i}, \boldsymbol{v_i}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

SOLUTION:
Treat the tangential velocities as *parameters*

$$p(\boldsymbol{\theta}, v_t|\boldsymbol{y}) \propto \prod_i f(\boldsymbol{r_i}, v_r|\boldsymbol{\theta}, v_t)p(\boldsymbol{\theta})p(v_t)$$

Method

- Gather kinematic data
- Choose a model (likelihood) and priors
- Sample the Posterior Distribution

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \prod_i f(\boldsymbol{r_i}, \boldsymbol{v_i}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

  - (Metropolis step, hybrid-Gibbs for v_t)
- Result: Markov Chain proportional to $p(\boldsymbol{\theta}|\boldsymbol{y})$

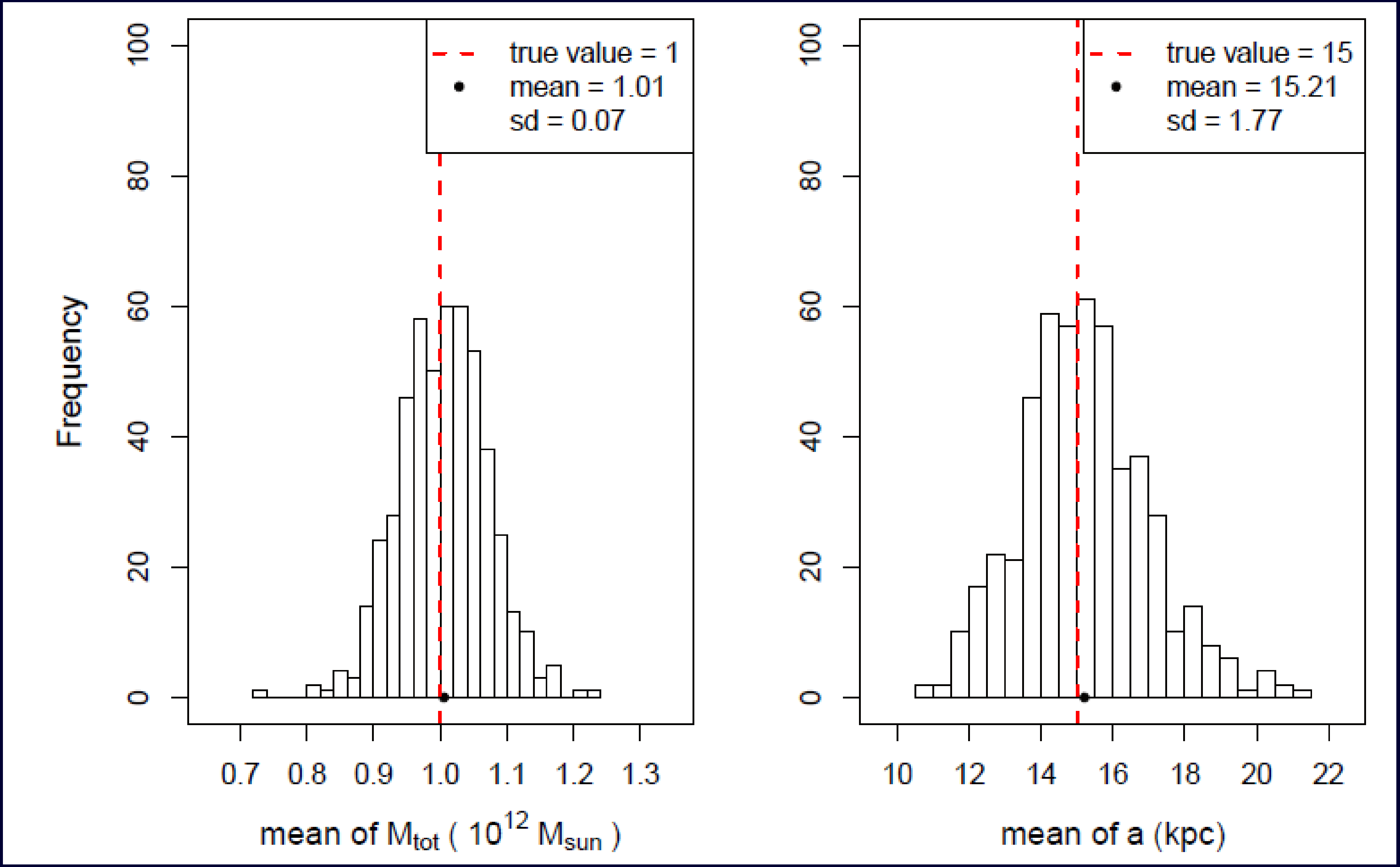
$(M_{tot}, a, v_{t1}, v_{t2}, \ldots, v_{tn})$

Simulations & Testing

Scenario	Simulated Data	Data Availability
1	Isotropic	complete
2	Isotropic	50% incomplete
3	Anisotropic	50% incomplete

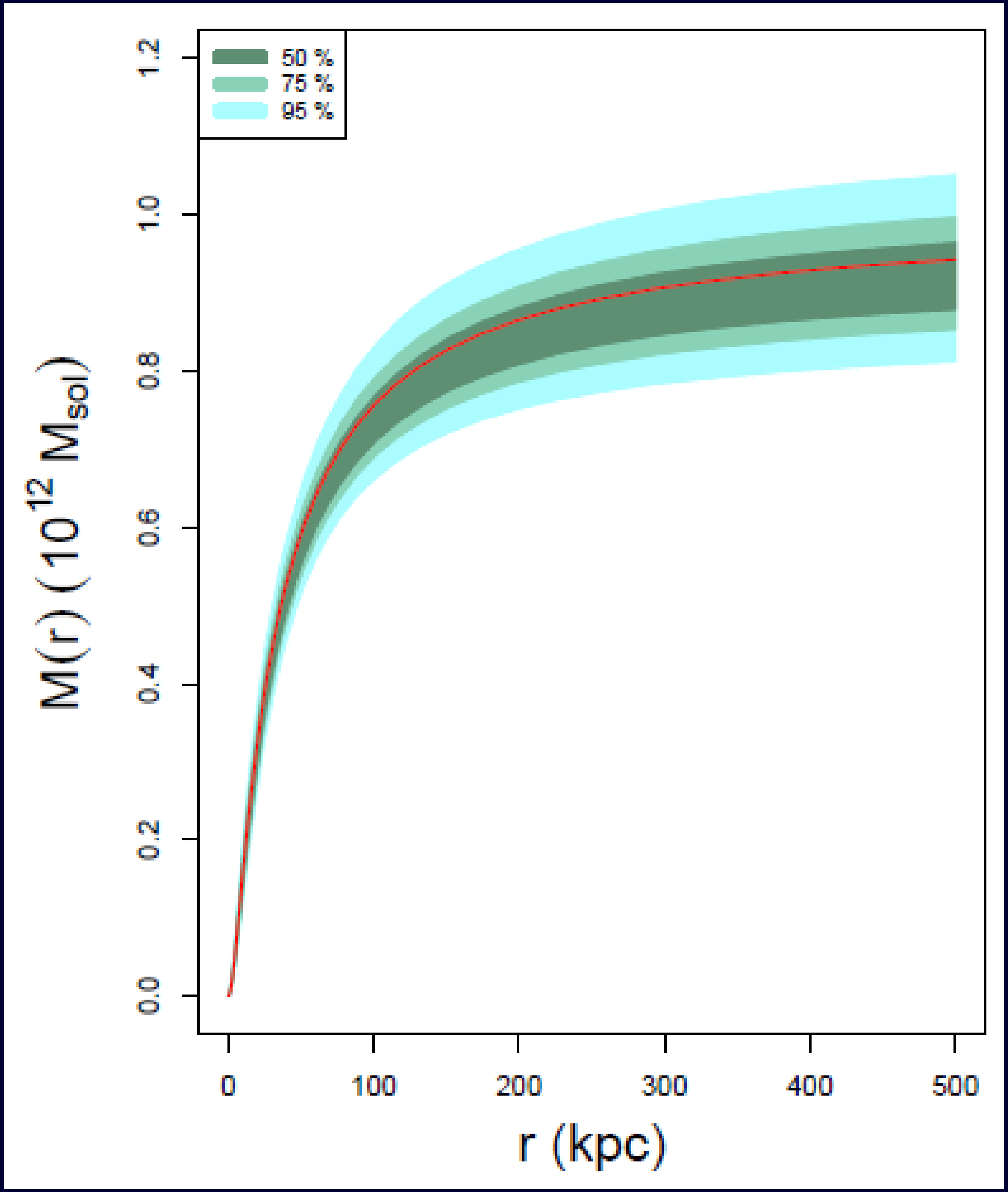
Analyze each scenario assuming isotropic Hernquist model

Scenario 1: distribution of estimates



Eadie, Harris, & Widrow (2015) ApJ, in press

Scenario 1: example mass profile



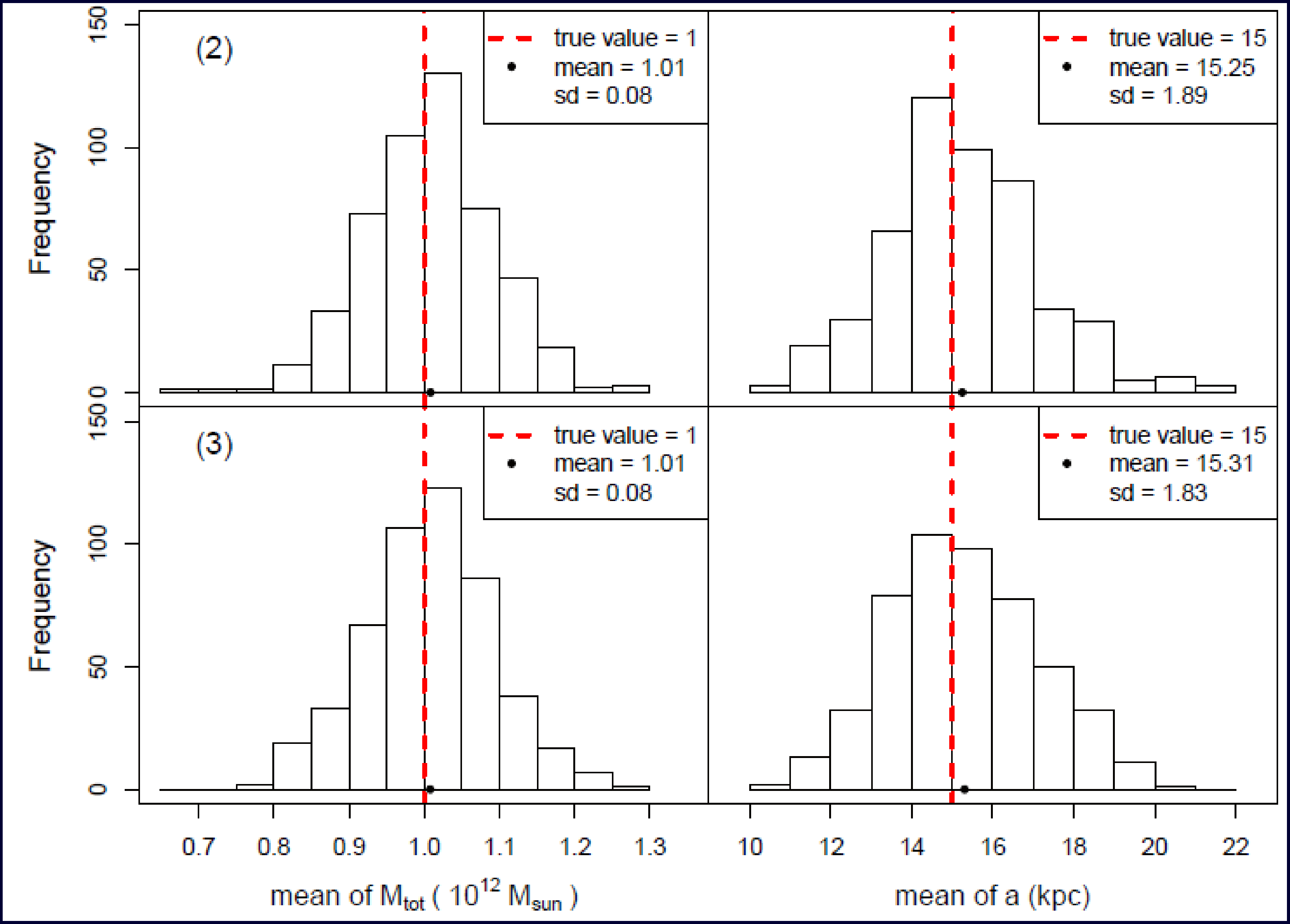
Eadie, Harris, & Widrow (2015), in press

Simulations & Testing

Scenario	Simulated Data	Data Availability
1	Isotropic	complete
2	Isotropic	50% incomplete
3	Anisotropic	50% incomplete

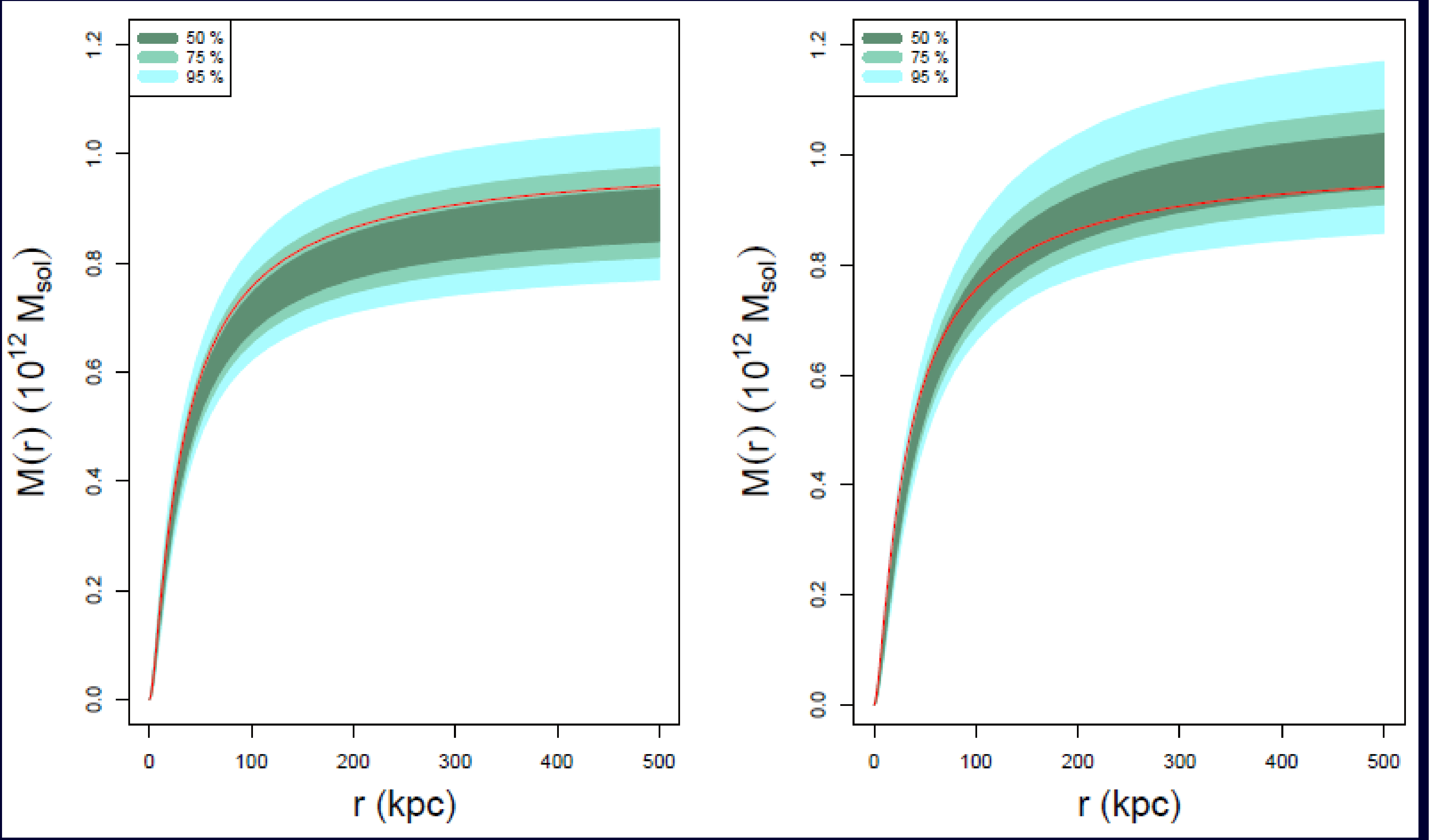
Analyze each scenario assuming isotropic Hernquist model

Scenario 2 & 3: distributions of estimates



Eadie, Harris, & Widrow (2015), in press

Scenario 2 & 3: example mass profiles



Eadie, Harris, & Widrow (2015), in press

On to real data!

Satellite data:

- 88 satellites, covering 3kpc < r < 261kpc
  - 59 GCs
  - 29 Dwarf galaxies

Data compiled from: Dinescu et al. (1999), Casseti-Dinescu et al (2010, 2013), Harris (1996), Boylan-Kochlin (2013), and Watkins et al (2010)

Satellite data:

- 88 satellites, covering 3kpc < r < 261kpc
  - 59 GCs

    (26 are missing tangential velocities)
  - 29 Dwarf galaxies

    (18 are missing tangential velocities)

Data compiled from: Dinescu et al. (1999), Casseti-Dinescu et al (2010, 2013), Harris (1996), Boylan-Kochlin (2013), and Watkins et al (2010)

Satellite data:

- 88 satellites, covering 3kpc < r < 261kpc
  - 59 GCs

    (26 are missing tangential velocities)
  - 29 Dwarf galaxies

    (18 are missing tangential velocities)

*Aside from the incomplete data, all other data have already been converted to the Galactocentric reference frame in previous studies

Data compiled from: Dinescu et al. (1999), Casseti-Dinescu et al (2010, 2013), Harris (1996), Boylan-Kochlin (2013), and Watkins et al (2010)

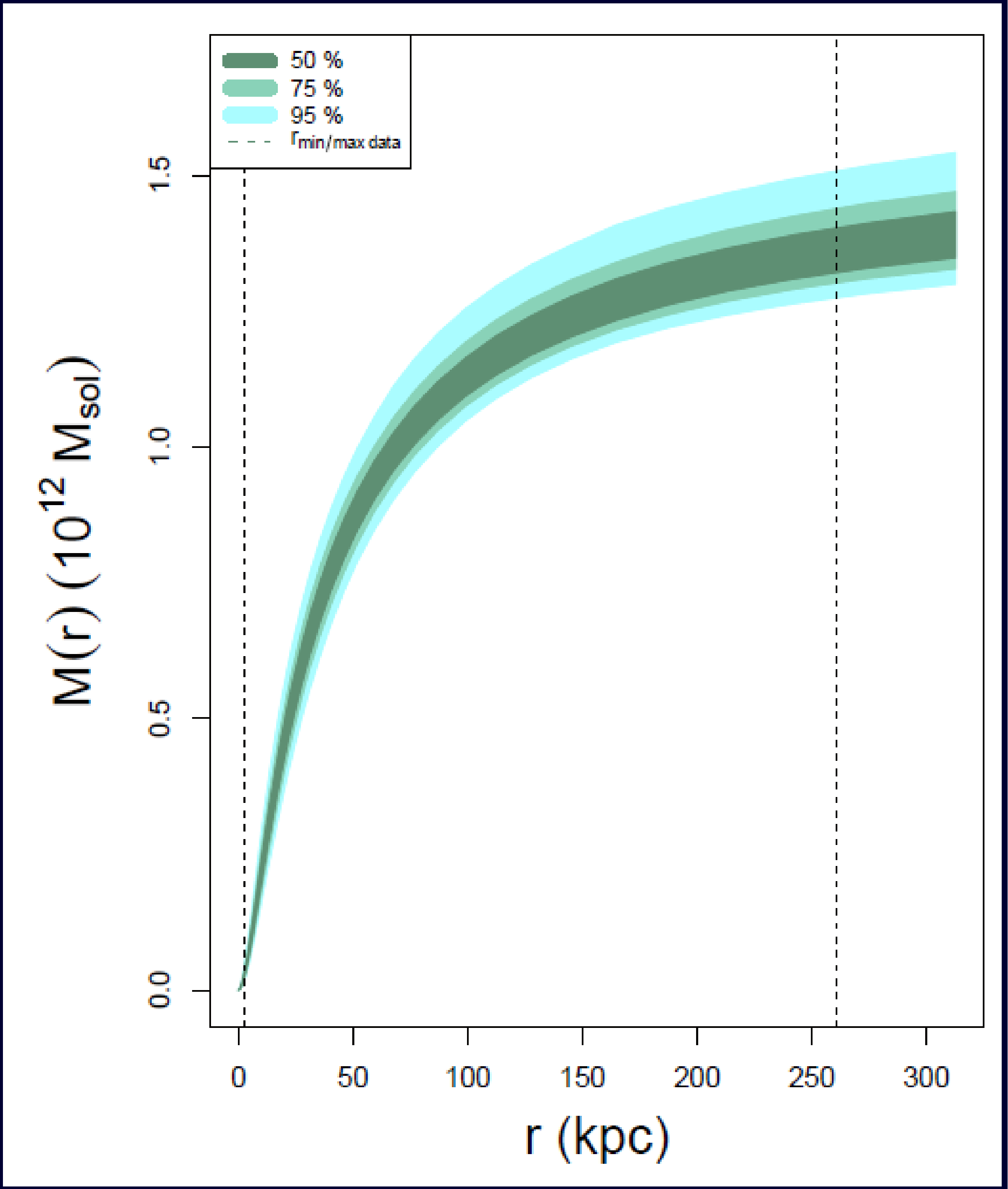
MW Mass profile

- Isotropic Hernquist model assumed
- Total mass estimate:

  $1.55 \times 10^{12}\ M_{sol}$

  (1.42, 1.73)
- Mass within 260 kpc
  - $1.37 \times 10^{12}\ M_{sol}$

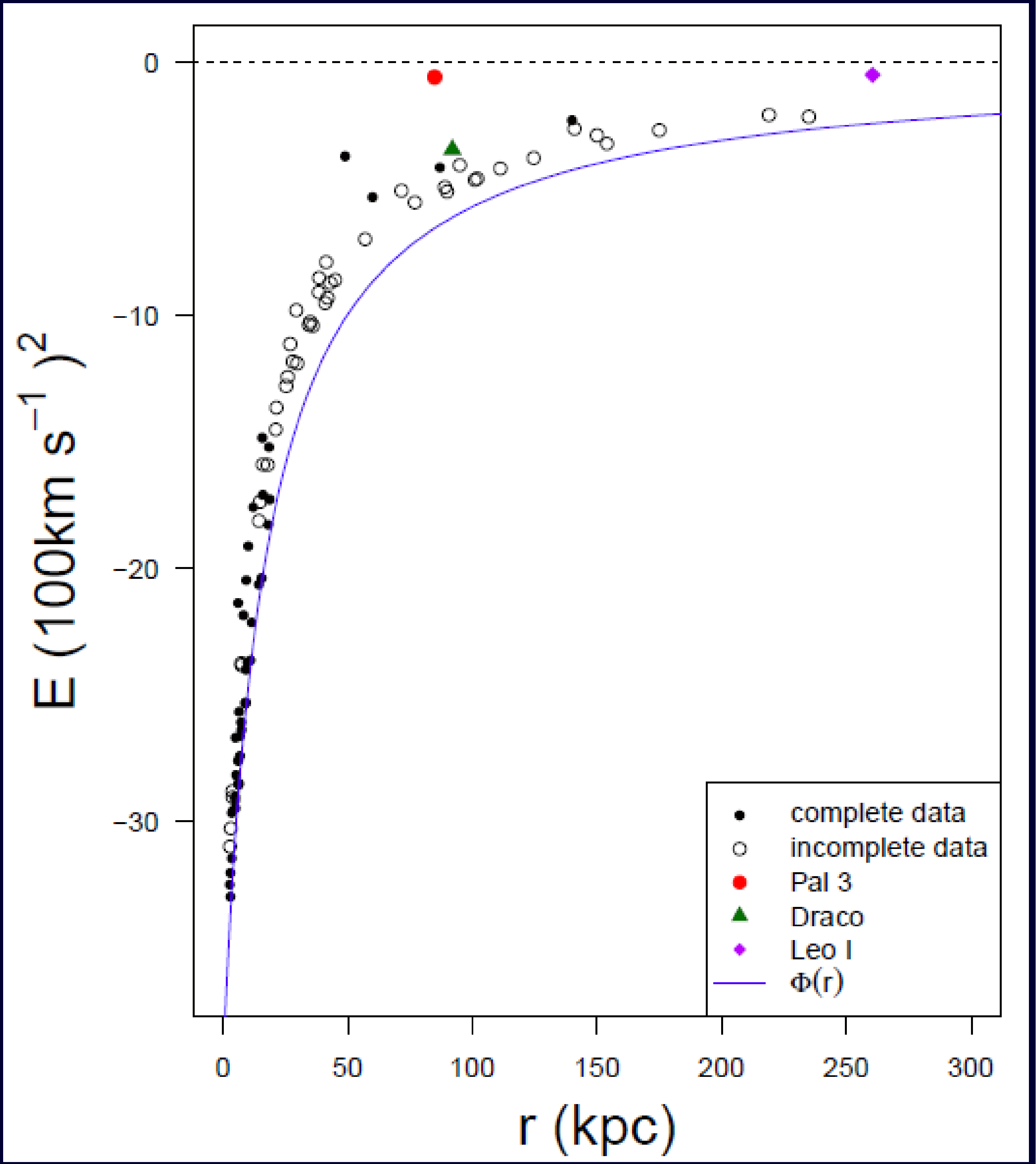
    (1.27, 1.51)



Eadie, Harris, & Widrow (2015), in press

Energy Profile

- Isotropic Hernquist model assumed
- Incomplete data
  - estimate of v_t from posterior distribution
- Gravitational potential
  - parameter estimates from posterior distribution



Preliminary Check: Sensitivity Analysis

Preliminary Check: Sensitivity Analysis

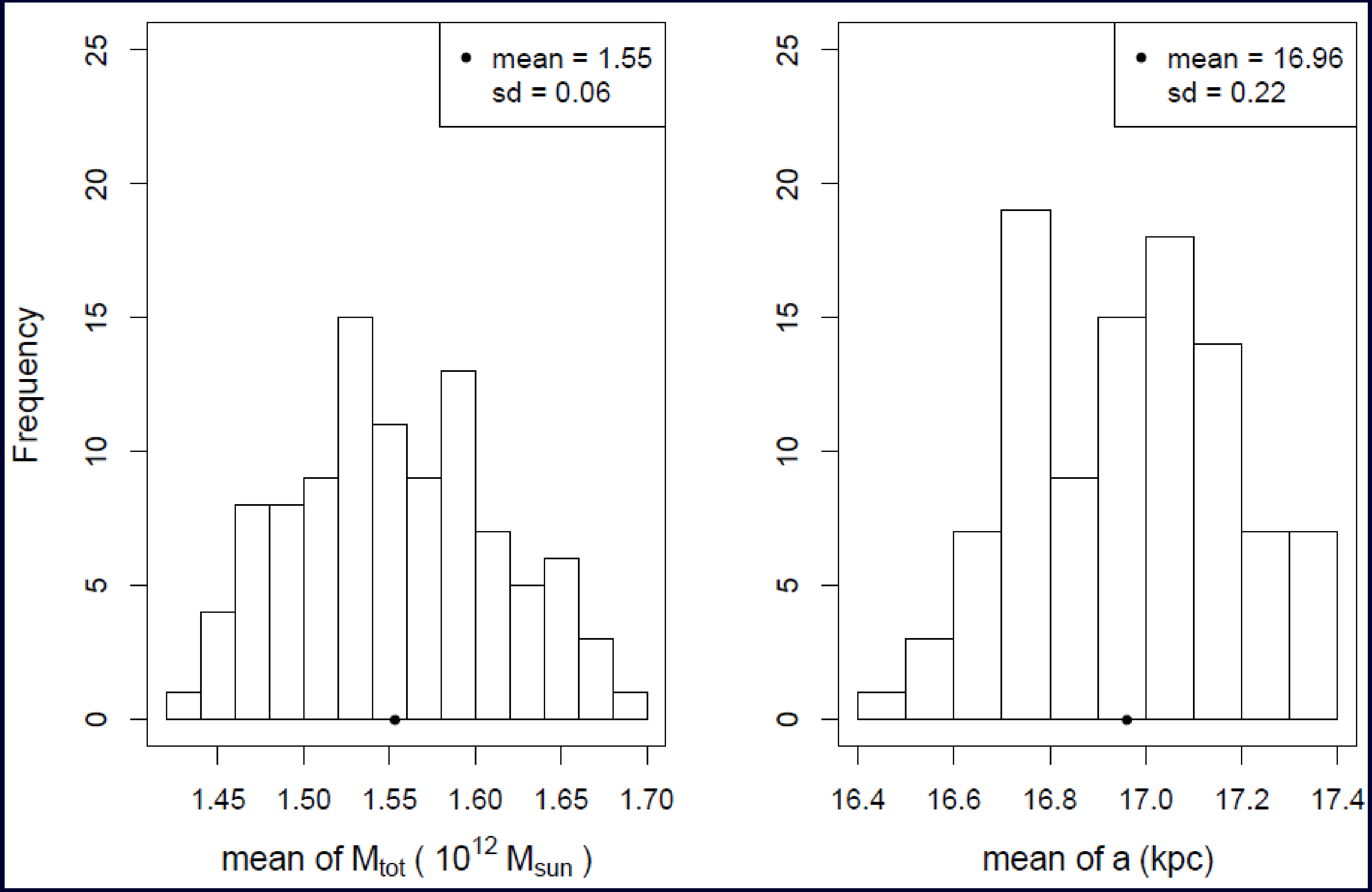
- Create 100 data sets with different tangetial velocities
  - Adjust velocities via random draw from a normal distribution with variance equal to the uncertainty

$$v_{t,new} = v_t + N(0, \Delta v_t)$$

Sensitivity Analysis Results

100 synthetic data sets

$$v_{t,new} = v_t + N(0, \Delta v_t)$$



Next step:
Include uncertainties via a hierarchical model

- incorporate uncertainty in **r** and **v**

- The problem becomes a hierarchical one...

$$f(r*, v*|r_{true}, v_{true}, \sigma_{r,v})p(\sigma_{r,v})p(r_{true}, v_{true}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

observed r and v

= a probability distribution involving the known measurement uncertainty

hyperprior

Conclusion

- Developed a Bayesian method to incorporate complete and incomplete data in the Galactic Mass estimation problem
- Simulations showed method is robust and effective when there is a mix of complete and incomplete data
- Preliminary analysis gives very encouraging results
  - Consistent between models
  - Results consistent with other methods

Eadie, Harris, & Widrow, ApJ 2015 (in press) will be posted to astro-ph in the next couple days

Some Future Work

- Milky Way
  - Proper hierarchical Bayesian analysis incorporating measurement uncertainties
  - Implementing the NFW model into the code
- Models where satellites do not follow the same distribution as dark matter halo particles
- Looking ahead to GAIA data
- R package Galactic Mass Estimator (GME)

Thank you!