Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 8 March 2013 (MN LATEX style file v2.2)

A non-parametric method for mass modelling spherical systems

P. Steger[1⋆], D. von Rickenbach[1], J. I. Read[1,2]
[1] *Institute for Astronomy, Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland*
[2] *Department of Physics, University of Surrey, Guildford, GU2 7XH, UK*

8 March 2013

ABSTRACT
Abstract here.

Key words:

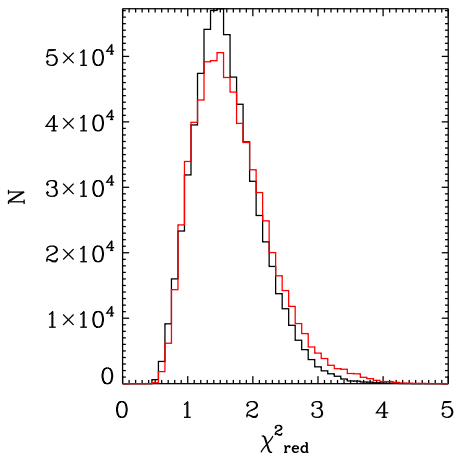

Figure 1. Example figure.

1 INTRODUCTION

Intro here. Example reference: Carlin et al. (2012).

2 METHOD

The collisionless Boltzmann equation for a spherical system with gravitational potential Φ,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \nabla_{\vec{x}} f \cdot \vec{v} - \nabla_{\vec{v}} f \cdot \nabla_{\vec{x}} \Phi = 0, \quad (1)$$

describes the motion of tracer stars with distribution function $f(\vec{x}, \vec{v})$.

In spherical coordinates (r, θ, ϕ), the collisionless Boltzmann equation then reads as

$$\frac{\partial f}{\partial t} + \dot{r}\frac{\partial f}{\partial r} + \dot{\theta}\frac{\partial f}{\partial \theta} + \dot{\phi}\frac{\partial f}{\partial \phi} + \dot{v}_r\frac{\partial f}{\partial v_r} + \dot{v}_\theta\frac{\partial f}{\partial v_\theta} + \dot{v}_\phi\frac{\partial f}{\partial v_\phi} = 0 \quad (2)$$

with velocities

$$\dot{r} = v_r, \quad (3)$$
$$\dot{\theta} = v_\theta / r \quad (4)$$
$$\dot{\phi} = v_\phi / r \sin\theta. \quad (5)$$

The assumption of steady state hydrodynamic equilibrium gives $\partial f/\partial t = 0$ and $\bar{v}_r = 0$, and using spherical symmetry $\bar{v}_\theta = 0$, $\bar{v}_\phi = 0$, with a unique tangential velocity dispersion $\sigma_\phi^2 = \sigma_\theta^2 = \sigma_t^2$ yields

$$\frac{1}{\nu}\frac{\partial}{\partial r}(\nu\sigma_r^2) + 2\frac{\sigma_r^2 - \sigma_t^2}{r} = -\frac{\partial \Phi}{\partial r} = -\frac{GM(<r)}{r^2} \quad (6)$$

with enclosed mass $M(<r)$, gravitational constant $G = 6.67398 \cdot 10^{-11}\mathrm{m}^3/\mathrm{kg\,s}^2$. The departure from spherical hydrostatic equilibrium $\sigma_r^2 = \sigma_t^2$ is measured by the anisotropy parameter

$$\beta \equiv 1 - \frac{\sigma_t^2}{\sigma_r^2} \quad (7)$$

with values in the range from $-\infty$ (purely circular orbits) through 0 (hydrostatic equilibrium) to 1 (purely radial orbits).

Integrating both sides of equation 6 gives the main equation of this paper,

$$\sigma_r^2(R) = \frac{1}{\nu(R)} \exp\left(2\int_{r_{min}}^{r} \frac{\beta(s)}{s}\mathrm{d}s\right) \cdot \quad (8)$$
$$\int_R^\infty \frac{GM(r)\nu(r)}{r^2} \exp\left(2\int_{r_{min}}^{r} \frac{\beta(s)}{s}\mathrm{d}s\right)\mathrm{d}r.$$

For distant spherical systems, only the projected velocity dispersion σ_{LOS} can be measured, which in our case is given by

⋆ E-mail: psteger@phys.ethz.ch

$$\sigma^2_{\rm LOS}(R) = \frac{2}{\Sigma(R)} \int_R^\infty \left(1 - \beta \frac{R^2}{r^2}\right) \frac{\nu(r)\sigma_r^2(r) r}{\sqrt{r^2 - R^2}} {\rm d}r, \qquad (9)$$

where $\Sigma(R)$ denotes the surface mass density at radius R.

In the following, we present a fully non-parametric method for the solution of equation 9 for the total gravitating mass density $\rho(r)$, given observed $\nu(r)$ and $\sigma_{\rm LOS}(r)$, where r denotes the projected two-dimensional radius from the center of mass of the spherical system. We get the enclosed mass $M(<r)$ from the density via

$$M(<r) = \int_0^r \rho(r) r^2 dr, \qquad (10)$$

which shows up in eq. 8. In principle, the above-mentioned method can be generalized to investigate alternative gravity models, if the acceleration $GM(r)/r^2$ is replaced with the respective form of $-\partial\Phi/\partial r$.

The degeneracy between mass M and velocity anisotropy β is accounted for: For any non-isothermal system, we let vary the anisotropy $\beta(r)$ as well. We checked that in the case of a simple Hernquist profile, $\beta(r) \approx 0$ is retrieved correctly.

The main idea is to let an Monte Carlo Markov Chain sample the parameter space $[\nu_i, \sigma_{{\rm LOS},i}, \rho]$ for distinct populations $i = 1...N$ of stellar or gaseous tracers.

Two different approaches are taken into account for the sampling of densities:

Tracer densities ν_i are expected to fall with increasing radii. To ensure this, one can explicitly build a monotonic function

$$\nu_i(r) = \int_0^r \tilde{\nu}_i(r') dr' \qquad (11)$$

with parameters $\tilde{\nu}_i(r') > 0$ discretized in bins. The other possibility is to let $\tilde{\nu}_i(r')$ vary freely, or just sample $\nu_i(r') = \tilde{\nu}_i|_{r'}$ directly. As it turns out, the MCMC prefers decreasing densities anyhow, so the ν_i parameters were allowed to vary freely throughout the rest of this paper.

Furthermore, the sampling might be done in a linear or logarithmic fashion. Wherever negative components are required, $\beta_i(r)$, linear sampling was chosen, s.t.

$$\beta_i^{(n+1)}(r) = \beta_i^{(n)}(r) + \delta\beta_i(r) \qquad (12)$$

with new parameter $\beta_i^{(n+1)}(r)$ in iteration $n+1$ determined from its old value $\beta_i^{(n)}(r)$ at iteration n and the parameter stepsize $\delta\beta_i(r)$ drawn from a random uniform distribution. On the other hand, for any positive definite parameter that needs to span a range in logarithmic space, $\nu_i(r)$ and $\rho(r)$, we sample logarithmically,

$$\nu_i^{(n+1)}(r) = 10^{\tilde{\nu}_i^{(n+1)}}, \qquad \tilde{\nu}_i^{(n+1)}(r) = \tilde{\nu}_i^{(n)}(r) + \delta\tilde{\nu}_i(r) \quad (13)$$

In a next step, $\sigma_{{\rm LOS},i}(r)$ is calculated ν_i, $\rho(r)$, and $\beta_i(r)$ according eq. **??**. This is done numerically, involving three integrations, which are performed with extrapolations up to $2 \cdot r_{max,{\rm data}}$, s.t. contributions from $\rho(r > r_{max}$ hinders an artificial falloff of $\sigma_{\rm LOS}$. $\rho(r > r_{max})$ is modeled as a powerlaw with variable slope, adding one more parameter for the MCMC to sample.

The last step involves comparison of $\nu_i(r)$, $\sigma_i(r)$ and $beta_i(r)$, if available, to the data for the tracer populations to get an error function

$$\chi^2 = \sum_{i=1}^{N} \chi^2_{\nu,i} + \chi^2_{\sigma,i} + \chi^2_{\beta,i} \qquad (14)$$

$$\chi^2_{\nu,i} = \sum_{j=1}^{N_{bin}} \left(\frac{\nu_{i,{\rm data}}(r_j) - \nu_{i,{\rm model}}(r_j)}{\epsilon_\nu(r_j)} \right)^2 \qquad (15)$$

and accordingly for $\chi^2_{\sigma,i}$ and $\chi^2_{\beta,i}$. In absence of a measured $\beta_i(r)$, we set $\chi^2_{\beta,i} = 0$.

The model for iteration $n+1$ is accepted if

$$\exp(\chi^2_n - \chi^2_{n+1}) < \varepsilon, \qquad \varepsilon \in [0, 1) \qquad (16)$$

for a uniform random ε. If not, the model is rejected.

The stepsize can vary for each bin, and is changed during an initialization phase. If the acceptance rate of models lies between 0.24 and 0.26, it is decreased by factors of 1.01, else, it is increased by the same amount. After a burn-in phase of several 100 accepted models with $\chi^2 < \chi^2_{end} = 70$, the stepsize is frozen and the MCMC starts storing the accepted models for further statistical analysis. A default 10^5 models are taken, where nothing else is indicated.

2.1 Binning characteristics

The number of bins for ν, σ, β, ρ are free parameters. They are set to fulfill

$$n_\nu = n_\sigma = n_\beta = n_\rho \leqslant n_{\rm data} \qquad (17)$$

with number of datapoints $n_{\rm data}$. This choice simplifies integration greatly, and prevents invention of information on scales smaller than the frequency of datapoints.

TODO: check that nbin is set s.t.

$$\chi^2_{red} = \frac{\chi^2}{n_{\rm data} - n_\nu - n_\sigma - n_\beta - n_\rho - 1} \qquad (18)$$

is minimized, and still the whole parameter space is tracked.

The whole parameter space for β_i is sampled with $n_\beta = 12$ if $n_{\rm data} = 30$ in the case of 10000 tracers.

The dark matter density is calculated by subtracting the measured baryon density from the dynamical mass density.

2.2 Priors

Following priors are included in the MCMC, and can help to reject unphysical models from the start:

1) cprior: $M(r = 0) = 0$, the central mass is set to 0;
2) bprior: $\rho(r) \geqslant \rho_b(r) - \epsilon_{\rho,b}(r) \forall r \geqslant 0$, ensures that no models with overall densities below the measured baryon density (reduced by the measurement error) are considered any further;

3) lbprior: $M(r > r_{max}) \leqslant M(< r_{max})/3.$, rejects any model which has more than 33% of the overall mass up to the outermost radius in the extrapolated bins;

4) rising ρ prior: $(\rho(r + \Delta r) - \rho(r))/\rho(r) \leqslant 0.5$, prevents ρ rising more than 50% for the next bin. There is no reason for the overall mass density to rise outwards in dynamically old systems. It might be favourable for convergence, though, if a dip in $\rho(r)$ does not lead to immediate rejection of all models with correct $\rho(r + \Delta r)$;

5) $\beta_i(r + \Delta r) - \beta_i(r) < 0.5$: prevent any sudden jumps in β_i;

We show in the appendix what effects the disabling of any of these priors have.

3 TESTS

3.1 Hernquist model

To check the correct function of the integration routine, we used the analytic formulas from Hernquist 1990 TODO: cite:

$$M(r) = r \quad (19)$$
$$\nu(r) = r \quad (20)$$
$$\beta(r) = 0 \quad (21)$$

as inputs, and calculated $\sigma_{\rm LOS}(r)$. It turns out to be similar to the analytic value (Baes, Dejonghe 2002) of

$$\sigma_r(r) = r(1+r)^3 \ln\left(\frac{1+r}{r}\right) - \frac{r(25+52r+42r^2+12r^3)}{12(1+r)}$$
$$I(r) = \frac{1}{2\pi}\frac{(2+r^2)X(r)-3}{(1-r^2)^2}$$
$$X(r) = \begin{cases}(1-r^2)^{-1/2}\operatorname{arcsech} r, & \text{for } 0 \leqslant r \leqslant 1,\\ (r^2-1)^{-1/2}\operatorname{arcsecans} r, & \text{for } 1 \leqslant r \leqslant \infty\end{cases},$$
$$\sigma_{\rm LOS}(r) = \frac{1}{I(r)}\frac{1}{24\pi(1-r^2)^3} \times [3r^2(20-35r^2+28r^4-8r^6)X(r) + (6-65r^2+68r^4-24r^6)] - \frac{r}{2},$$

as is seen in figure 2

TODO: description of Hernquist simulation files. We then let ν, ρ, β vary. The MCMC correctly recovered the mass distribution, see fig. 4.

3.2 Data quality

How many tracer stars are needed to determine the overall density profile reliably? To address this question, we performed three runs with a restricted set of tracer particles. In the first, 10^3 particles were chosen out of the 10^6 simulated particles.

With 10^4 particles, the confidence intervals shrink, see fig. 4

If these 10^4 particles are split into two populations of each 5e3 particles, TODO: with scalelengths of about 1/10

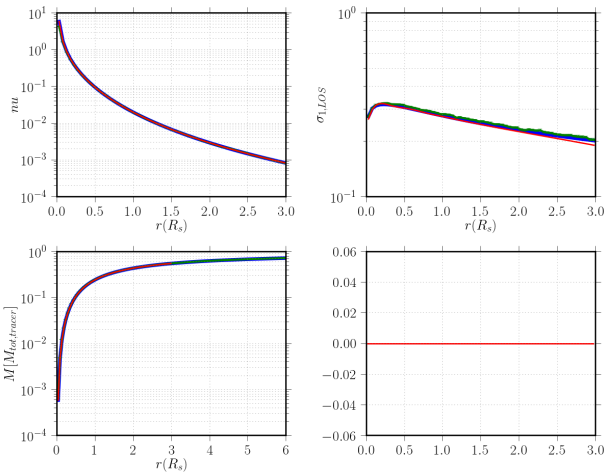

Figure 2. Check of the integration routine.

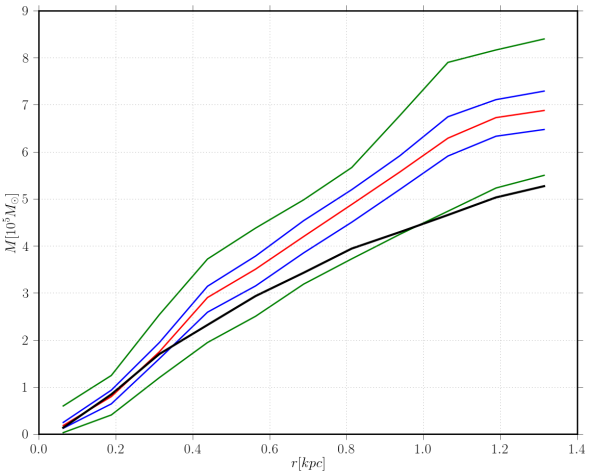

Figure 3. Hernquist profile found by MCMC model (red) for 1000 tracer particles

It has to be highlighted that most of the second population particles are inside the first two bins, so the overall convergence is not visibly affected above the third bin.

4 RESULTS

Results from Walker mock data, and possibly some early dwarf spheroidal observations.

We apply our method to another set of mock data, the spherical models for the Gaia challenge by Walker and Penarrubia. They consist of dynamical tracer populations with density distribution

$$\nu_*(r) = \nu_0 \left(\frac{r}{r_*}\right)^{-\gamma_*}\left[1+\left(\frac{r}{r_*}\right)^{\alpha_*}\right]^{(\gamma_*-\beta_*)/\alpha_*} \quad (22)$$

inside dark matter halos of the form

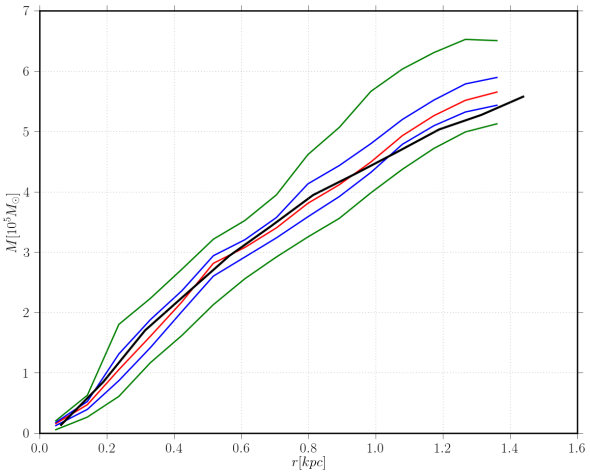

Figure 4. Hernquist profile found by MCMC model (red) for 10^4 tracer particles. Green curve shows the enclosed mass derived from data.

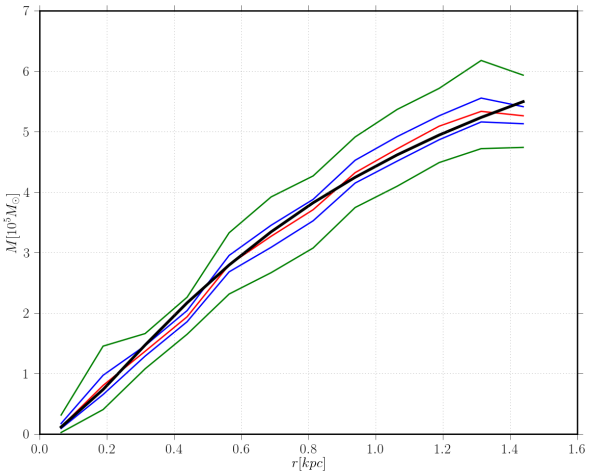

Figure 5. Hernquist profile found by MCMC model (red) for $2 \cdot 10^3$ tracer particles. Green curve shows the enclosed mass derived from data.

$$\rho_{\rm DM} = \rho_0 \left(\frac{r}{r_{\rm DM}}\right)^{-\gamma_{\rm DM}} \left[1 + \left(\frac{r}{r_{\rm DM}}\right)^{\alpha_{\rm DM}}\right]^{(\gamma_{\rm DM}-\beta_{\rm DM})/\alpha_{\rm DM}} \tag{23}$$

with scale radii $r_*, r_{\rm DM}$, inner and outer logarithmic slopes of $\gamma_*, \gamma_{\rm DM}$ and $\beta_*, \beta_{\rm DM}$, with transition parameters $\alpha_*, \alpha_{\rm DM}$.

The anisotropy follows the functional form of Osipkov 1979 and Merritt 1985,

$$\beta_{\rm anisotropy}(r) = 1 - \frac{\sigma_\theta^2}{\sigma_r^2} = \frac{r^2}{r^2 + r_a^2}. \tag{24}$$

with scale radius r_a, turning over from nearly isotropic at $r \to 0$ to radially isotropic at $r_* = r_a$.

Of these distributions, finite samplings are taken and converted to mock observational data including spectral indices, systemic velocities, proper motions, binary motion.

TODO: plot of Walker tests

5 CONCLUSIONS

Conclusions.

6 ACKNOWLEDGEMENTS

JIR would like to acknowledge support from SNF grant PP00P2_128540/1.

REFERENCES

Carlin J. L., Majewski S. R., Casetti-Dinescu D. I., Law D. R., Girard T. M., Patterson R. J., 2012, ApJ, 744, 25

7 APPENDIX

7.1 Effects of turning off priors