Stellar-dynamical modelling

Eugene Vasiliev

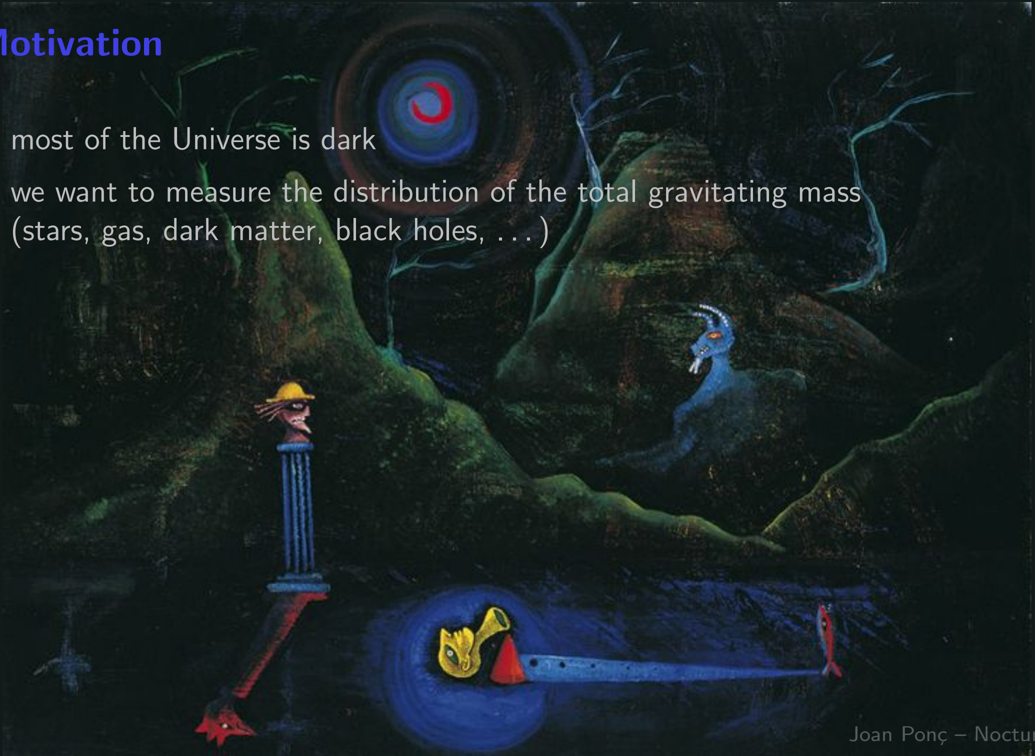
University of Surrey, May 2026

Motivation

most of the Universe is dark

Joan Ponç – Nocturno

Motivation

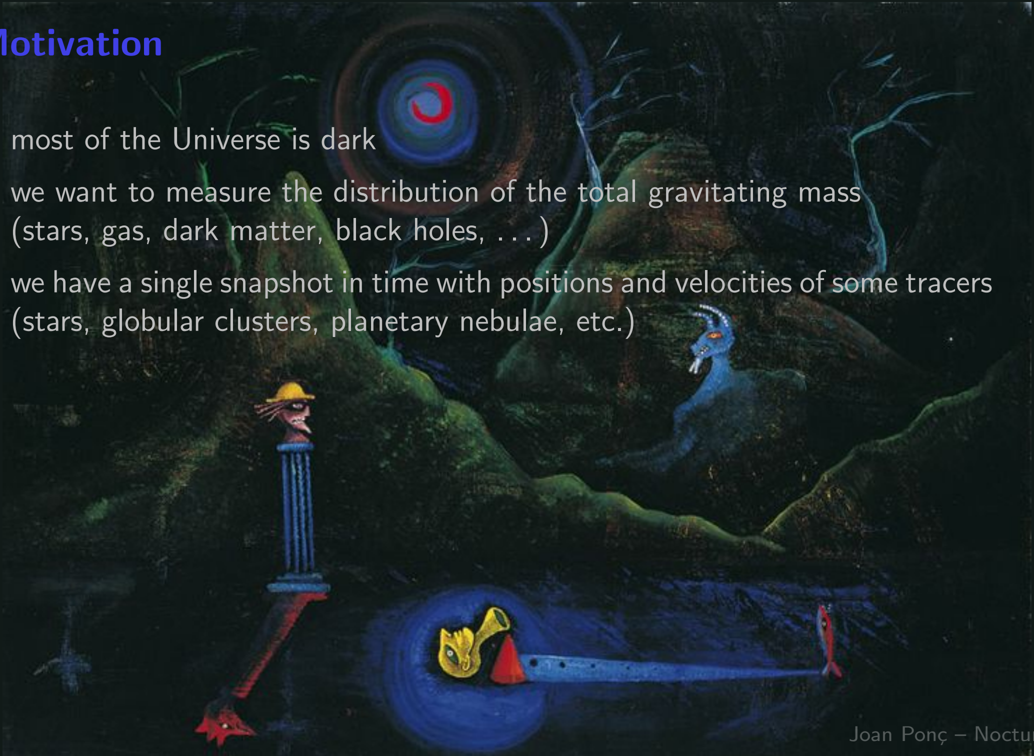


most of the Universe is dark

we want to measure the distribution of the total gravitating mass (stars, gas, dark matter, black holes, . . .)

Joan Ponç – Nocturno

Motivation



most of the Universe is dark

we want to measure the distribution of the total gravitating mass (stars, gas, dark matter, black holes, . . .)

we have a single snapshot in time with positions and velocities of some tracers (stars, globular clusters, planetary nebulae, etc.)

Joan Ponç – Nocturno

Distribution function in stellar dynamics

A stellar system composed of a large number $N \gg 1$ of “identical” stars can be described by a distribution function (DF):

$f(\mathbf{x}, \mathbf{v};\ t)$ – probability density in the phase space (at time t),

$\int f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3x\, \mathrm{d}^3v = 1$ (unit-normalised) or M_total (mass-normalised).

Distribution function in stellar dynamics

A stellar system composed of a large number $N \gg 1$ of “identical” stars can be described by a distribution function (DF):

$f(\mathbf{x}, \mathbf{v};\ t)$ – probability density in the phase space (at time t),

$\int f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3x\, \mathrm{d}^3v = 1$ (unit-normalised) or M_total (mass-normalised).

For a multicomponent system, define a separate DF f_k for each species k, or an “extended distribution function” (EDF) with additional arguments:

$f(\mathbf{x}, \mathbf{v};\ \boldsymbol{\eta};\ t)$ ($\boldsymbol{\eta}$ may be age, mass, chemical composition, etc.)

Distribution function in stellar dynamics

A stellar system composed of a large number $N \gg 1$ of "identical" stars can be described by a distribution function (DF):

$f(\mathbf{x}, \mathbf{v};\ t)$ – probability density in the phase space (at time t),

$\int f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3x\, \mathrm{d}^3v = 1$ (unit-normalised) or M_{total} (mass-normalised).

For a multicomponent system, define a separate DF f_k for each species k, or an "extended distribution function" (EDF) with additional arguments:

$f(\mathbf{x}, \mathbf{v};\ \boldsymbol{\eta};\ t)$ ($\boldsymbol{\eta}$ may be age, mass, chemical composition, etc.)

DF offers a complete description of the stellar population:

density $\rho(\mathbf{x}) = \displaystyle\int f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3v$,

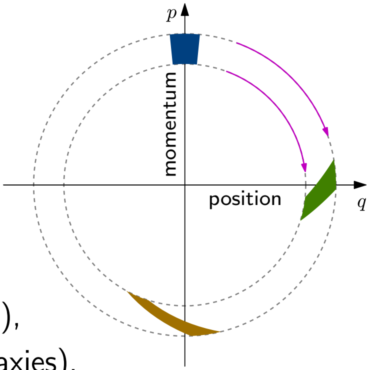
mean velocity $\overline{\mathbf{v}}(\mathbf{x}) = \dfrac{1}{\rho(\mathbf{x})} \displaystyle\int \mathbf{v}\, f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3v$,

second moment of velocity $\overline{v_{ij}^2}(\mathbf{x}) = \dfrac{1}{\rho(\mathbf{x})} \displaystyle\int v_i\, v_j\, f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3v$, etc.

Evolution of the distribution function

As the system evolves, stars move along their trajectories, but the DF at the location of any star is conserved (Liouville's theorem):

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, H] &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} + \frac{\partial f}{\partial \mathbf{v}}\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \\ &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\,\mathbf{v} - \frac{\partial f}{\partial \mathbf{v}}\frac{\partial \Phi}{\partial \mathbf{x}} \boxed{= 0}\end{aligned}$$

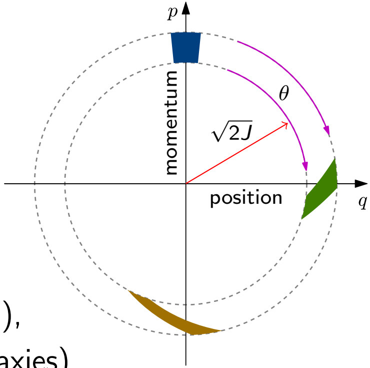


Collisionless Boltzmann (or Vlasov) equation (CBE), generally valid for large-N systems (e.g., entire galaxies), in which the relaxation time is long compared to their age (even if the system is not in a steady state!).

Evolution of the distribution function

As the system evolves, stars move along their trajectories, but the DF at the location of any star is conserved (Liouville's theorem):

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, H] &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} + \frac{\partial f}{\partial \mathbf{v}}\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \\ &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\,\mathbf{v} - \frac{\partial f}{\partial \mathbf{v}}\frac{\partial \Phi}{\partial \mathbf{x}} = 0\end{aligned}$$

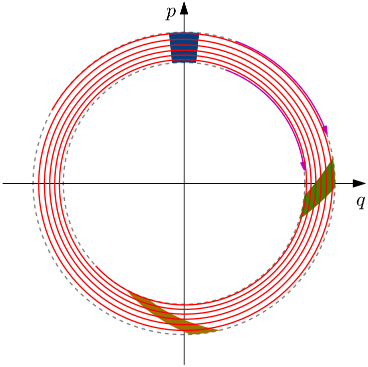


Collisionless Boltzmann (or Vlasov) equation (CBE), generally valid for large-N systems (e.g., entire galaxies), in which the relaxation time is long compared to their age (even if the system is not in a steady state!).

Note that CBE looks the same in any canonical coordinates $\mathbf{p}, \mathbf{q}$, including the action–angle variables $\mathbf{J}, \boldsymbol{\theta}$, where $J \equiv \frac{1}{2\pi}\oint p\,\mathrm{d}q$.

Phase mixing and the Jeans theorem

In realistic galactic potentials, orbital frequencies $\boldsymbol{\Omega}$ depend on the orbit, hence an initially localised ensemble of points eventually attains a uniform distribution in phase angles $\boldsymbol{\theta}$ and fills the entire accessible region for its integrals of motion.



Jeans theorem

In a steady state, the DF may depend only on the integrals of motion:

$$f(\mathbf{x},\mathbf{v}) = f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\,\Phi)\big)$$

Here $\mathcal{I}$ may be energy E, or actions $\mathbf{J}$ – anything that is conserved by the given potential.

As the phase mixing timescale is usually $\gg \Omega^{-1}$, the apparently well-mixed state of galaxies is likely caused by other processes, such as *violent relaxation* in a rapidly varying potential.

Equilibrium models in stellar dynamics

Distribution function of stars $f(\mathbf{x}, \mathbf{v}, t)$ satisfies [sometimes] the collisionless Boltzmann equation:

$$\frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial t} + \mathbf{v}\,\frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial \mathbf{x}} - \frac{\partial \Phi(\mathbf{x}, t)}{\partial \mathbf{x}}\,\frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial \mathbf{v}} = 0.$$

Potential ⇔ mass distribution

not measured directly on human timescales

In order to infer anything about the potential from a time-dependent DF, need to make further assumptions about the initial state of the system, e.g., that the stars belong to a single stream or were perturbed from an equilibrium configuration in a specific way, etc.

Equilibrium models in stellar dynamics

Distribution function of stars $f(\mathbf{x}, \mathbf{v}, t)$ satisfies [sometimes] the collisionless Boltzmann equation:

$$\mathbf{v}\,\frac{\partial f(\mathbf{x}, \mathbf{v}\quad)}{\partial \mathbf{x}} - \frac{\partial \Phi(\mathbf{x}\quad)}{\partial \mathbf{x}}\,\frac{\partial f(\mathbf{x}, \mathbf{v}\quad)}{\partial \mathbf{v}} = 0.$$

3D
(want to infer)

Steady-state assumption $\Longrightarrow$ Jeans theorem:

$$f(\mathbf{x}, \mathbf{v}) = f\left(\mathcal{I}(\mathbf{x}, \mathbf{v};\ \Phi)\right)$$

3D – 6D
(observed)

integrals of motion ($\leq$ 3D?), e.g., $\mathcal{I} = \{E, L, \ldots\}$

Equilibrium models in stellar dynamics

Distribution function of stars $f(\mathbf{x}, \mathbf{v}, t)$ satisfies [sometimes] the collisionless Boltzmann equation:

$$\mathbf{v}\,\frac{\partial f(\mathbf{x}, \mathbf{v} \quad)}{\partial \mathbf{x}} - \frac{\partial \Phi(\mathbf{x} \quad)}{\partial \mathbf{x}}\,\frac{\partial f(\mathbf{x}, \mathbf{v} \quad)}{\partial \mathbf{v}} = 0.$$

3D
(want to infer)

Steady-state assumption $\Longrightarrow$ Jeans theorem:

$$f(\mathbf{x}, \mathbf{v}) = f\left(\mathcal{I}(\mathbf{x}, \mathbf{v};\ \Phi)\right)$$

3D – 6D
(observed)

integrals of motion ($\leq$ 3D?), e.g., $\mathcal{I} = \{E, L, \dots\}$

When observations provide more than 3 phase-space coordinates,
the potential is overconstrained! (can detect non-stationarity, but not measure it)

Purpose of dynamical modelling

Determine the mass distribution of a stellar system from the kinematics of some tracer population(s) under the assumption of dynamical equilibrium.

Methods

increase in complexity and cost

0. Virial theorem: $2\,K + W = 0$

kinetic energy potential energy

virial mass estimators: $G\,M \propto r\,\sigma^2$

1. Jeans equations

2. Distribution functions

3. Orbit superposition

4. Made-to-measure

1. Jeans equations

Simple 1d example:

start from the steady-state CBE

$$v\,\frac{\partial f(x,v)}{\partial x} - \frac{\mathrm{d}\Phi(x)}{\mathrm{d}x}\,\frac{\partial f(x,v)}{\partial v} = 0.$$

1. Jeans equations

Simple 1d example:

start from the steady-state CBE

$$v \, \frac{\partial f(x,v)}{\partial x} - \frac{\mathrm{d}\Phi(x)}{\mathrm{d}x} \, \frac{\partial f(x,v)}{\partial v} = 0.$$

Multiply by velocity and integrate over v:

$$\int v^2 \frac{\partial f}{\partial x} \, \mathrm{d}v - \frac{\mathrm{d}\Phi}{\mathrm{d}x} \int \frac{\partial f}{\partial v} \, v \, \mathrm{d}v = 0.$$

1. Jeans equations

Simple 1d example:

start from the steady-state CBE

$$v\,\frac{\partial f(x,v)}{\partial x} - \frac{\mathrm{d}\Phi(x)}{\mathrm{d}x}\,\frac{\partial f(x,v)}{\partial v} = 0.$$

Multiply by velocity and integrate over v:

$$\int v^2\,\frac{\partial f}{\partial x}\,\mathrm{d}v - \frac{\mathrm{d}\Phi}{\mathrm{d}x}\int \frac{\partial f}{\partial v}\,v\,\mathrm{d}v = 0.$$

The second term can be integrated by parts:

$$\frac{\partial \int f\,v^2\,\mathrm{d}v}{\partial x} - \frac{\mathrm{d}\Phi}{\mathrm{d}x}\left[\underbrace{\cancel{(f\,v)\Big|_{-\infty}^{+\infty}}}_{\substack{\text{vanishes because}\\ f\,v \to 0 \text{ as } |v| \to \infty}} - \int f\,\underbrace{\frac{\partial v}{\partial v}}_{=1}\,\mathrm{d}v\right] = 0.$$

1. Jeans equations

Simple 1d example:

start from the steady-state CBE

$$v\,\frac{\partial f(x,v)}{\partial x} - \frac{\mathrm{d}\Phi(x)}{\mathrm{d}x}\,\frac{\partial f(x,v)}{\partial v} = 0.$$

Multiply by velocity and integrate over v:

$$\int v^2\,\frac{\partial f}{\partial x}\,\mathrm{d}v - \frac{\mathrm{d}\Phi}{\mathrm{d}x}\int \frac{\partial f}{\partial v}\,v\,\mathrm{d}v = 0.$$

The second term can be integrated by parts:

$$\frac{\partial \int f\,v^2\,\mathrm{d}v}{\partial x} - \frac{\mathrm{d}\Phi}{\mathrm{d}x}\left[\cancel{(f\,v)\Big|_{-\infty}^{+\infty}} - \int f\,\frac{\partial v}{\partial v}\,\mathrm{d}v\right] = 0.$$

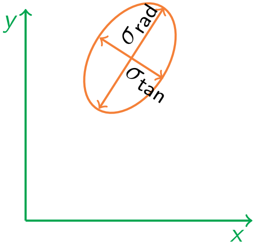
Finally we get the equation linking 0th and 2nd moments of f over velocity:

$$\frac{\mathrm{d}\left(\rho\,\overline{v^2}\right)}{\mathrm{d}x} + \frac{\mathrm{d}\Phi}{\mathrm{d}x}\,\rho = 0.$$

$$\left.\begin{array}{l}\text{gravitational force}\\ \text{pressure gradient}\end{array}\right\}\ \text{hydrostatic equilibrium}$$

Jeans equation in spherical systems

Unlike fluids, pressure in stellar systems can be anisotropic, i.e. $\overline{v_i v_j}$ is a tensor. In the spherical non-rotating case, it has two independent components: radial $\sigma^2_{\rm rad}$ and tangential $\sigma^2_{\rm tan}$ velocity dispersions, and the Jeans equation is



$$\frac{\mathrm{d}\left(\rho\,\sigma^2_{\rm rad}\right)}{\mathrm{d}r} + \frac{\mathrm{d}\Phi}{\mathrm{d}r} + \frac{2\beta}{r}\,\rho\,\sigma^2_{\rm rad} = 0,$$

where $\beta(r) \equiv 1 - \dfrac{\sigma^2_{\rm tan}(r)}{2\sigma^2_{\rm rad}(r)}$ is the anisotropy coefficient:

$\beta = 1$ – purely radial orbits,
$\beta > 0$ – radially anisotropic case,
$\beta = 0$ – isotropic case,
$\beta < 0$ – tangentially anisotropic case,
$\beta = -\infty$ – purely circular orbits.

Same Φ and ρ can produce different $\sigma_{\rm r,t}$ profiles, depending on β (more on that later).

Jeans equations in axisymmetric systems

Four components of the velocity ellipsoid tensor $(\sigma_R^2, \sigma_z^2, \overline{v_R v_z}, \overline{v_\phi^2})$, two eqns:

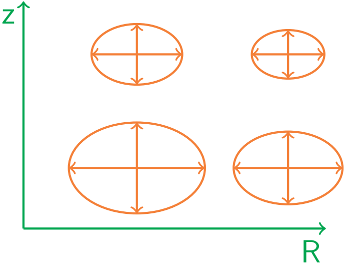
$$0 = \rho \frac{\partial \Phi}{\partial z} + \frac{\partial(\rho\, \sigma_z^2)}{\partial z} + \frac{\partial(\rho\, \overline{v_R v_z})}{\partial R} + \frac{\rho\, \overline{v_R v_z}}{R},$$

$$0 = \rho \frac{\partial \Phi}{\partial R} + \frac{\partial(\rho\, \sigma_R^2)}{\partial R} + \frac{\partial(\rho\, \overline{v_R v_z})}{\partial z} + \frac{\rho\left(\sigma_R^2 - \overline{v_\phi^2}\right)}{R}.$$

Further assumptions about the shape and orientation of the velocity ellipsoid:

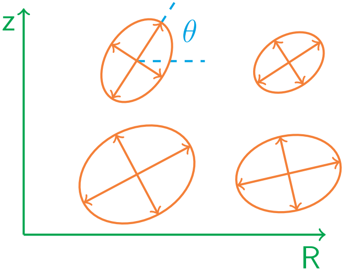
cylindrical alignment $(\overline{v_R v_z} = 0)$

Jeans Anisotropic method (JAM) [Cappellari 2008]



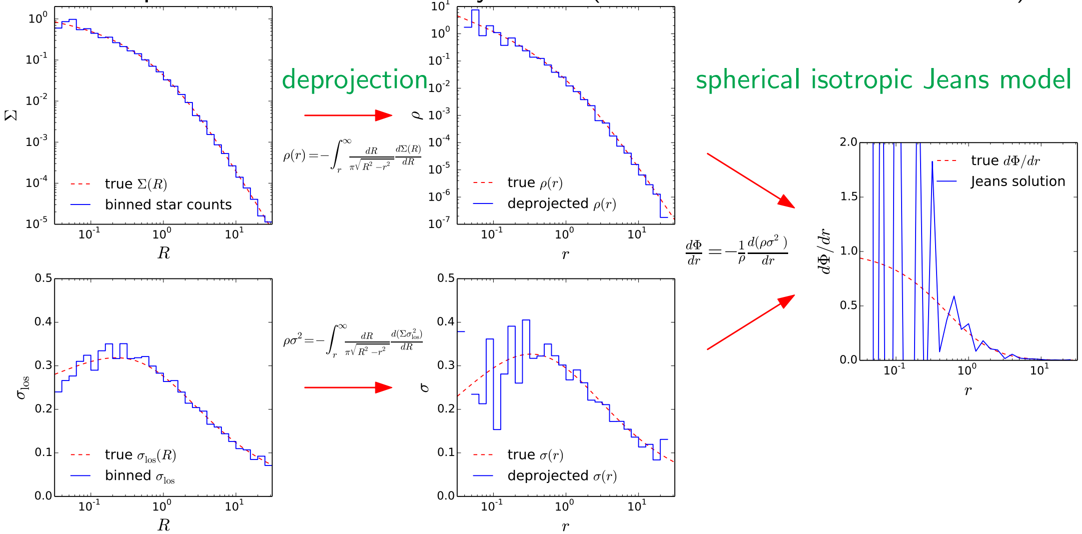
spherical alignment

JAM-sph [Cappellari 2019]



A note on forward vs. inverse modelling

It is formally possible to take the observed quantities (density, velocity dispersions) and "invert" them to obtain the gravitational potential directly; however, this inversion process involves derivatives of noisy measurements, and in practice the result is barely usable (but see Rehemtulla+ 2022, Medina+ 2025).



deprojection

spherical isotropic Jeans model

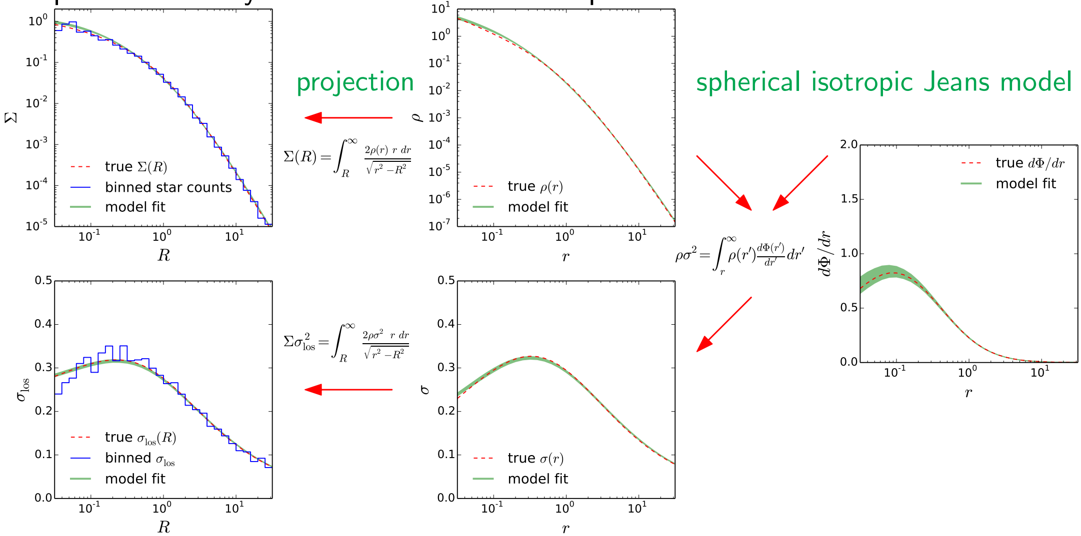
$$\rho(r) = -\int_r^\infty \frac{dR}{\pi\sqrt{R^2 - r^2}} \frac{d\Sigma(R)}{dR}$$

$$\rho\sigma^2 = -\int_r^\infty \frac{dR}{\pi\sqrt{R^2 - r^2}} \frac{d(\Sigma\sigma_{\rm los}^2)}{dR}$$

$$\frac{d\Phi}{dr} = -\frac{1}{\rho}\frac{d(\rho\sigma^2\ )}{dr}$$

A note on forward vs. inverse modelling

A more reliable way is to start with a parametric model for $\Phi(r)$ and $\rho(r)$, determine the kinematic properties and project the model into data space, then compute the model likelihood against observational data, and repeat the procedure many times to find the best-fit parameters and their uncertainties.



projection

$$\Sigma(R) = \int_R^\infty \frac{2\rho(r)\ r\ dr}{\sqrt{r^2 - R^2}}$$

spherical isotropic Jeans model

$$\rho\sigma^2 = \int_r^\infty \rho(r') \frac{d\Phi(r')}{dr'} dr'$$

$$\Sigma\sigma_{\rm los}^2 = \int_R^\infty \frac{2\rho\sigma^2\ \ r\ dr}{\sqrt{r^2 - R^2}}$$

Jeans equations in dynamical modelling

Typically one solves the Jeans equation(s) for $\sigma(\mathbf{x})$ given [measured] $\rho(\mathbf{x})$ and assumed $\Phi(\mathbf{x})$, then compares the velocity dispersions with observations.

For instance, in spherical systems the solution for the radial dispersion is

$$\sigma_r^2(r) = \frac{1}{\rho(r)\,g(r)} \int_r^\infty \mathrm{d}s\;\rho(s)\,g(s)\,\frac{\mathrm{d}\Phi(s)}{\mathrm{d}s}, \quad g(r) \equiv \exp\left[2\int_0^r \frac{\beta(s)}{s}\,\mathrm{d}s\right],$$

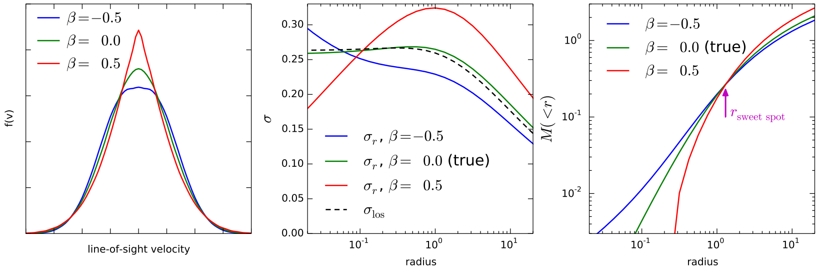
and then the projected (line-of-sight) velocity dispersion is

$$\sigma_{\mathrm{los}}^2(R) = \frac{2}{\Sigma(R)} \int_R^\infty \left(1-\beta(r)\,\frac{R^2}{r^2}\right)\,\frac{\sigma_r^2(r)\,\rho(r)\,r}{\sqrt{r^2-R^2}}\,\mathrm{d}r.$$

But the anisotropy profile $\beta(r)$ cannot be determined from the line-of-sight velocity component alone, so different assumptions about $\beta(r)$ produce different deprojected $\sigma_r(r)$ and hence different inferred mass profiles, leading to a “mass–anisotropy degeneracy” (MAD) [Binney & Mamon 1982].

Jeans equations: ways to cope with MADness

- The mass enclosed within a particular “sweet-spot radius” ($\simeq$ half-light radius of tracers) is insensitive to β [Walker+ 2009, Wolf+ 2010, Churazov+ 2010].



- Using two different tracer populations, one can determine the mass at two different radii $\Rightarrow$ constrain the slope of the mass profile [Walker & Peñarrubia 2011].
- Adding proper motions makes it possible to measure or constrain β directly.
- One can derive higher-order Jeans equations and use the information about the shape of the velocity distribution, not just its dispersion [Merrifield & Kent 1990, Richardson & Fairbairn 2014, Read & Steger 2017, Bañares-Hernández+ 2026].
- DF-based methods also utilise the full velocity distribution.

2. DF-based models: two approaches

a. From the given *total* potential $\Phi(\mathbf{x})$ and density of *stellar tracers* $\rho(\mathbf{x})$ (which need not be related), obtain $f\big(\mathcal{I}(\mathbf{x}, \mathbf{v};\ \Phi)\big)$.

2. DF-based models: two approaches

a. From the given *total* potential $\Phi(\mathbf{x})$ and density of *stellar tracers* $\rho(\mathbf{x})$ (which need not be related), obtain $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$.

Spherical isotropic systems: $f(E)$ is given by the Eddington inversion formula

$$f(E) = \frac{1}{\sqrt{8}\,\pi^2}\frac{\mathrm{d}}{\mathrm{d}E}\int_E^{\Phi_\infty}\frac{\mathrm{d}\Phi}{\sqrt{\Phi-E}}\,\frac{\mathrm{d}\rho\big(r(\Phi)\big)}{\mathrm{d}\Phi}.$$

But spherical systems can be *an*isotropic, so $f(E,L)$ is not uniquely specified by $\rho(r)$ and $\Phi(r)$ – various generalisations of the above formula are available for some specific forms of the DF, e.g. $f(E,L)=\hat{f}\big(E+\frac{L^2}{2r_\mathrm{a}^2}\big)\,L^{-2\beta}$ [Ossipkov 1979, Merritt 1985, Cuddeford 1991].

2. DF-based models: two approaches

a. From the given *total* potential $\Phi(\mathbf{x})$ and density of *stellar tracers* $\rho(\mathbf{x})$ (which need not be related), obtain $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$.

Spherical isotropic systems: $f(E)$ is given by the Eddington inversion formula

$$f(E) = \frac{1}{\sqrt{8}\,\pi^2}\frac{\mathrm{d}}{\mathrm{d}E}\int_E^{\Phi_\infty}\frac{\mathrm{d}\Phi}{\sqrt{\Phi-E}}\,\frac{\mathrm{d}\rho\big(r(\Phi)\big)}{\mathrm{d}\Phi}.$$

But spherical systems can be *an*isotropic, so $f(E,L)$ is not uniquely specified by $\rho(r)$ and $\Phi(r)$ – various generalisations of the above formula are available for some specific forms of the DF, e.g. $f(E,L) = \hat{f}\big(E+\frac{L^2}{2r_\mathrm{a}^2}\big)\,L^{-2\beta}$ [Ossipkov 1979, Merritt 1985, Cuddeford 1991].

Axisymmetric systems: both ρ and Φ are 2d functions of R, z;
$f(E, L_z)$ is determined uniquely from a contour integral in the complex plane [Hunter & Qian 1993].

But more generally, f may also depend on the third integral, so is not unique.

2. DF-based models: two approaches

a. From the given *total* potential $\Phi(\mathbf{x})$ and density of *stellar tracers* $\rho(\mathbf{x})$ (which need not be related), obtain $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$.

Spherical isotropic systems: $f(E)$ is given by the Eddington inversion formula

$$f(E) = \frac{1}{\sqrt{8}\,\pi^2}\frac{\mathrm{d}}{\mathrm{d}E}\int_E^{\Phi_\infty}\frac{\mathrm{d}\Phi}{\sqrt{\Phi-E}}\,\frac{\mathrm{d}\rho\big(r(\Phi)\big)}{\mathrm{d}\Phi}.$$

But spherical systems can be *an*isotropic, so $f(E,L)$ is not uniquely specified by $\rho(r)$ and $\Phi(r)$ – various generalisations of the above formula are available for some specific forms of the DF, e.g. $f(E,L) = \hat{f}\big(E+\frac{L^2}{2r_\mathrm{a}^2}\big)\,L^{-2\beta}$ [Ossipkov 1979, Merritt 1985, Cuddeford 1991].

Axisymmetric systems: both ρ and Φ are 2d functions of R, z;
$f(E,L_z)$ is determined uniquely from a contour integral in the complex plane [Hunter & Qian 1993].
But more generally, f may also depend on the third integral, so is not unique.

Most general (triaxial) case: all functions are 3d, but no general approach for determining $f(\mathcal{I})$ from $\rho(\mathbf{x})$ and $\Phi(\mathbf{x})$ is known.

2. DF-based models: two approaches

a. From the given *total* potential $\Phi(\mathbf{x})$ and density of *stellar tracers* $\rho(\mathbf{x})$ (which need not be related), obtain $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$.

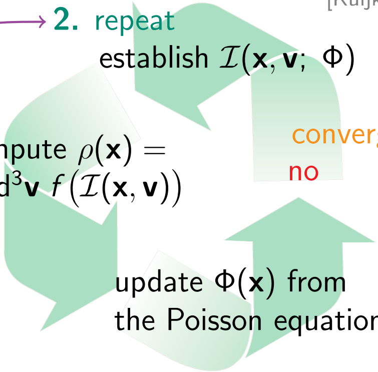
b. From an assumed $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$,
compute the density $\rho(\mathbf{x}) = \iiint f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)\,\mathrm{d}^3\mathbf{v}$
and then the corresponding potential $\Phi(\mathbf{x})$ from the Poisson equation.

2. DF-based models: two approaches

a. From the given *total* potential $\Phi(\mathbf{x})$ and density of *stellar tracers* $\rho(\mathbf{x})$ (which need not be related), obtain $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$.

b. From an assumed $f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)$,
compute the density $\rho(\mathbf{x}) = \iiint f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big)\,\mathrm{d}^3\mathbf{v}$
and then the corresponding potential $\Phi(\mathbf{x})$ from the Poisson equation.

Some spherical isotropic models exist (polytropes, King models, etc.),
but in general this circular dependency is resolved using an iterative approach
[Kuijken & Dubinski 1995, Binney 2014].



1. assume $f(\mathcal{I})$ and an initial guess for Φ →

2. repeat

establish $\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)$

compute $\rho(\mathbf{x}) = \iiint \mathrm{d}^3\mathbf{v}\ f\big(\mathcal{I}(\mathbf{x},\mathbf{v})\big)$

update $\Phi(\mathbf{x})$ from the Poisson equation

converged?

no

yes →

3. enjoy!

3. Schwarzschild's orbit-superposition method

Introduced by Schwarzschild (1979) as a practical approach
for constructing self-consistent triaxial models with prescribed $\rho(\mathbf{x}) \Leftrightarrow \Phi(\mathbf{x})$.

integrals of motion

To invert the equation $\rho(\mathbf{x}) = \iiint f\big(\mathcal{I}\,(\mathbf{x},\mathbf{v}\,;\ \Phi)\big)\, d^3\mathbf{v}$,

discretise both the density profile and the distribution function:

$\rho(\mathbf{x}) \implies$ cells of a spatial grid;
mass of each cell is $M_c = \iiint\limits_{\mathbf{x}\in V_c} \rho(\mathbf{x})\, d^3x$;

$f(\mathcal{I}) \implies$ collection of orbits with unknown weights:

$$f(\mathcal{I}) = \sum_{k=1}^{N_\mathrm{orb}} w_k\ \delta(\mathcal{I} - \mathcal{I}_k)$$

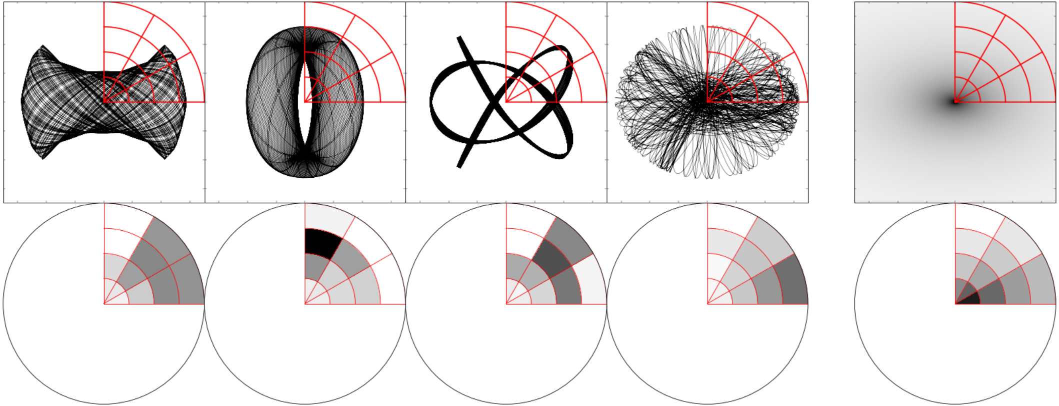
each orbit is a delta-function in the space of integrals of motion

adjustable weight of each orbit [to be determined]

Schwarzschild's orbit-superposition method: self-consistency

orbits in the model

target density



discretised orbit density

(fraction of time t_{kc} that k-th orbit spends in c-th cell)

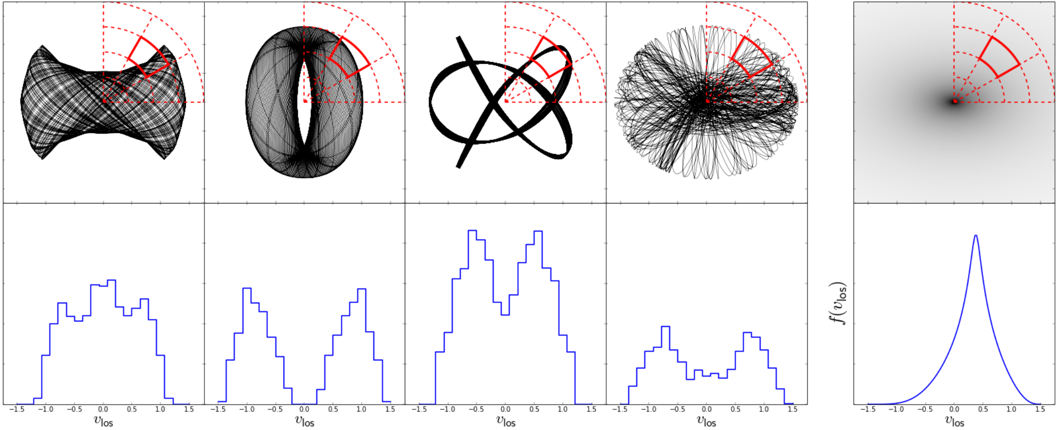
discretized density

(mass M_c in grid cells)

For each c-th cell we require $\sum_k w_k \, t_{kc} = M_c$, where $w_k \geq 0$ is orbit weight

Schwarzschild's orbit-superposition method: kinematics

orbits in the model



orbit LOSVDs

target LOSVD

Schwarzschild's orbit-superposition method: fitting procedure

- Assume some potential $\Phi(\mathbf{x})$
  (e.g., from the deprojected luminosity profile plus parametric DM halo or SMBH)
- Construct the orbit library in this potential:
  for each k-th orbit, store its contribution to the discretised density profile t_{kc}, $c = 1..N_{\rm cell}$ and to the kinematic observables u_{kn}, $n = 1..N_{\rm obs}$
- Solve the constrained optimisation problem to find orbit weights w_k:

$$\text{minimize } \chi^2 + \mathcal{S} \equiv \sum_{n=1}^{N_{\rm obs}} \left(\frac{\sum_{k=1}^{N_{\rm orb}} w_k\, u_{kn} - U_n}{\delta U_n} \right)^2 + \ \mathcal{S}\big(\{w_k\}\big)$$

regularisation term

observational constraints

their uncertainties

$$\text{subject to } w_k \ge 0, \quad k = 1..N_{\rm orb},$$

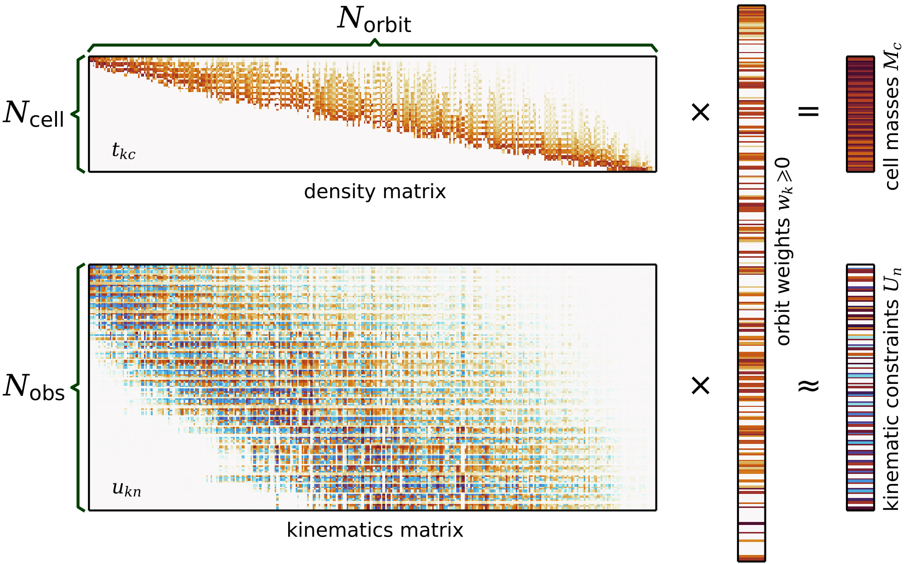
$$\sum_{k=1}^{N_{\rm orb}} w_k\, t_{kc} = M_c, \quad c = 1..N_{\rm cell}$$

density constraints (cell masses)

- Repeat for different choices of potential and find the one that has lowest χ^2

Schwarzschild's orbit-superposition method: fitting procedure

Solve the linear system with non-negativity constraints on the solution vector $w_k \geq 0$ (linear or non-linear optimisation problem)



4. Made-to-measure (M2M) *N*-body models

Introduced by Syer & Tremaine 1996 as a way of constructing "tailored" *N*-body models satisfying some observational constraints.

Ingredients:

- *N*-particle system with time-dependent phase-space coordinates and weights $\{\mathbf{x}_k, \mathbf{v}_k, w_k\}\Big|_{k=1}^{N_\mathrm{body}}$ moving in a potential $\Phi(\mathbf{x})$
- Observational constraints U_n and their uncertainties δU_n, $n = 1..N_\mathrm{obs}$
- Model predictions for these observations: $V_n = \sum_{k=1}^{N_\mathrm{body}} w_k \; \underbrace{K_n(\mathbf{x}_k, \mathbf{v}_k)}_{\text{some predefined kernels}}$

Objective:

- minimise $\Omega \equiv \frac{1}{2} \sum_{n=1}^{N_\mathrm{obs}} \Delta_n^2 + \mathcal{S}\big(\{w_k\}\big)$,

  where $\Delta_n \equiv (V_n - U_n)/\delta U_n$ is the deviation in *n*-th constraint,

  $\mathcal{S}\big(\{w_k\}\big)$ is some measure of smoothness (regularisation term),

  by varying the particle weights w_k.

Made-to-measure models

Objective is satisfied when $\dfrac{\partial \Omega}{\partial w_k} \equiv \sum_{n=1}^{N_{\rm obs}} \dfrac{\Delta_n\, K_n(\mathbf{x}_k, \mathbf{v}_k)}{\delta U_n} + \dfrac{\partial \mathcal{S}}{\partial w_k}$ is 0 for all k

Procedure:

repeat until $\Delta_n \approx 0$

- Evolve the N-body system in time: $\dot{\mathbf{x}}_k = \mathbf{v}_k, \; \dot{\mathbf{v}}_k = -\left.\dfrac{\partial \Phi}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_k}$
- Adjust the particle weights: $\dot{w}_k = -\dfrac{w_k}{\tau_{\rm ch}} \dfrac{\partial \Omega}{\partial w_k}$ (force-of-change)
- To reduce fluctuations, replace $\Delta_n(t)$ by a time-smoothed $\tilde{\Delta}_n(t) \equiv \dfrac{1}{\tau_{\rm sm}} \displaystyle\int_0^\infty \Delta_n(t-\tau)\, \exp\left(-\dfrac{\tau}{\tau_{\rm sm}}\right) d\tau$ in the above expression
- remove particles with too small w_k, split particles with too large w_k
- recompute the potential $\Phi(\mathbf{x})$ from particle positions and weights

Made-to-measure vs. Schwarzschild method

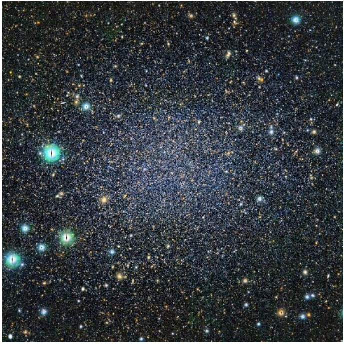
Both represent the DF as a large ensemble of δ-functions with weights as free parameters in the model:

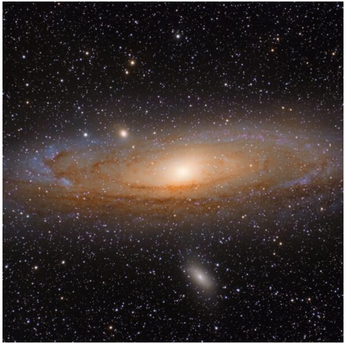
- N-body particles ($\sim 10^5 - 10^6$)
- time-average during evolution
- iteratively adjust weights (handmade gradient descent method)
- may adjust the potential during the fitting procedure
- live N-body system – easy to test the stability
- more expensive in CPU time

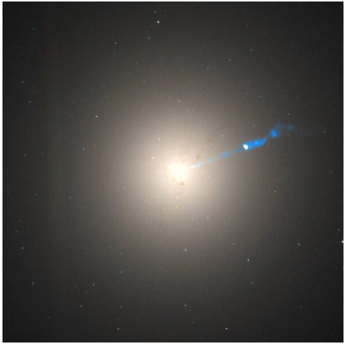
- orbits ($\sim 10^3 - 10^5$)
- compute entire orbits beforehand
- solve a large-scale constrained optimisation problem by black-box routines
- potential fixed in advance (need to construct a new orbit library each time a new potential is chosen)
- need to convert orbit library into an N-body model first

Observational data

	resolved	unresolved	(into individual stars)
X, Y sky-plane position	$+$	$+$	(surface brightness profile)
Z distance	$\mp$	$-$	
$\mu_{X,Y}$ proper motions	$\mp$	$-$	
V_{los} line-of-sight velocity	$\pm$	$+$	(LoS velocity distribution)
	Milky Way and LG galaxies	Galaxies outside the Local Group	


Sculptor dSph [credit: Subaru]


Andromeda [Wikipedia / B.Wesner]


M 87 [HST]

Dynamical modelling with resolved data

Bayes' rule: $$\mathcal{P}\big(\text{model} \mid \text{data}\big) = \frac{\mathcal{P}(\text{data} \mid \text{model})\ \ \mathcal{P}(\text{model})}{\mathcal{P}(\text{data})}$$

any priors
on model
parameters

fixed, thus irrelevant
for model comparison

Dynamical modelling with resolved data

Bayes' rule: $$\mathcal{P}\big(\text{model} \mid \text{data}\big) = \frac{\mathcal{P}(\text{data} \mid \text{model})\;\mathcal{P}(\text{model})}{\mathcal{P}(\text{data})}$$

any priors
on model
parameters

DF-based methods:

$$\mathcal{P}(\text{model} \mid \text{data}) = \prod_{i=1}^{N_\text{stars}} f\big(\mathbf{x}_i, \mathbf{v}_i;\ \boldsymbol{\alpha}\big) = \prod_{i=1}^{N_\text{stars}} f\Big(\mathcal{I}\big(\mathbf{x}_i, \mathbf{v}_i;\ \Phi(\mathbf{x};\ \boldsymbol{\beta})\big);\ \boldsymbol{\alpha}\Big),$$

where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ are the parameters of the DF and the potential, respectively,
which are optimised during model fitting;
in the self-consistent case, Φ is determined by the DF, so $\boldsymbol{\beta}$ are unused.

Dynamical modelling with resolved data

Bayes' rule: $$\mathcal{P}\big(\text{model} \mid \text{data}\big) = \frac{\mathcal{P}(\text{data} \mid \text{model})\ \mathcal{P}(\text{model})}{\mathcal{P}(\text{data})}$$

any priors on model parameters

DF-based methods:

$$\mathcal{P}(\text{model} \mid \text{data}) = \prod_{i=1}^{N_\text{stars}} f\big(\mathbf{x}_i, \mathbf{v}_i;\ \boldsymbol{\alpha}\big) = \prod_{i=1}^{N_\text{stars}} f\Big(\mathcal{I}\big(\mathbf{x}_i, \mathbf{v}_i;\ \Phi(\mathbf{x};\ \boldsymbol{\beta})\big);\ \boldsymbol{\alpha}\Big),$$

where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ are the parameters of the DF and the potential, respectively, which are optimised during model fitting;
in the self-consistent case, Φ is determined by the DF, so $\boldsymbol{\beta}$ are unused.

Jeans models:

Jeans eqs predict the second velocity moment $\overline{v^2}(\mathbf{x})$ or σ^2 given $\rho_\star$ and Φ.

or

assume some functional form for $f(\mathbf{x}, \mathbf{v})$, e.g. a Gaussian $f = \rho(\mathbf{x})\,\mathcal{N}\big(\mathbf{v}\,|\overline{v}, \sigma^2(\mathbf{x})\big)$, or a more compicated form when using higher-order Jeans eqns that also predict kurtosis, then proceed as in DF-based methods.

compute $\sigma^2_\text{obs}(\mathbf{x})$ from binned data, then compare it to the predicted values (loss of information due to binning!)

Dynamical modelling with resolved data: complications

- Missing phase-space coordinates

  need to marginalise the DF over missing dimensions:

  e.g. for 3d data, $f_{\mathrm{projected}}(X, Y, V_{\mathrm{los}}) = \iiint f(\mathbf{x}, \mathbf{v})\, \mathrm{d}Z\, \mathrm{d}V_X\, \mathrm{d}V_Y.$

  These integrals sometimes can be taken analytically (e.g. for second-order Jeans models with particular families of the density and anisotropy profiles), but usually have to be computed numerically – for every star in the catalogue! *[ouch!]*

Dynamical modelling with resolved data: complications

- Missing phase-space coordinates
- Observational uncertainties

  same recipe, marginalise over observational uncertainties:

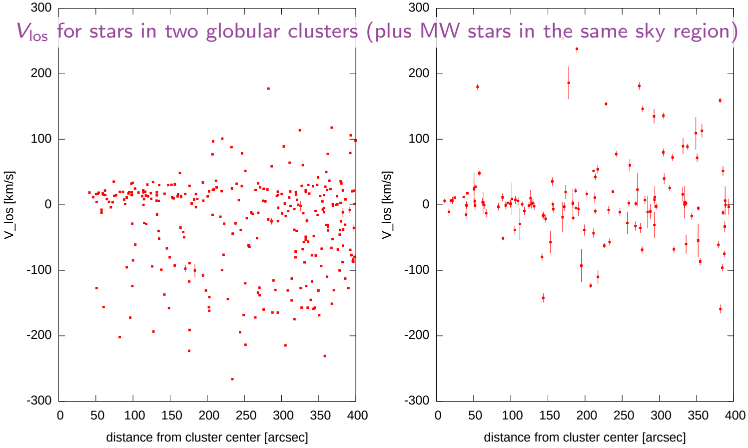
  $$f_{\mathrm{marg}}(\ldots, V_{\mathrm{los},i}) = \int f(\ldots, V'_{\mathrm{los}})\, \mathcal{N}\left(V'_{\mathrm{los}} \mid V_{\mathrm{los},i}, \epsilon_{V_{\mathrm{los}},i}\right) \mathrm{d}V'_{\mathrm{los}},$$

  where $V_{\mathrm{los},i}$ is the measured value for i-th star and $\epsilon_{V_{\mathrm{los}},i}$ is its error estimate.

Dynamical modelling with resolved data: complications

- Missing phase-space coordinates
- Observational uncertainties
- Foreground contamination

some stars in the catalogue do not belong to the stellar system of interest.



V_{los} for stars in two globular clusters (plus MW stars in the same sky region)

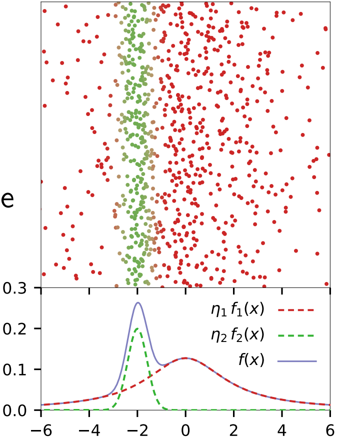
Dynamical modelling with resolved data: complications

- Missing phase-space coordinates
- Observational uncertainties
- Foreground contamination

  some stars in the catalogue do not belong to the stellar system of interest.
  $\Rightarrow$ introduce a mixture model
  $f_{\text{mix}}(V_{\text{los}}) = (1-\eta)\, f_{\text{model}}(\dots, V_{\text{los}}) + \eta\, f_{\text{fg}}(V_{\text{los}})$,
  where η is the fraction of [foreground] interlopers,
  f_{fg} is their assumed velocity distribution
  (e.g. a broad Gaussian with mean $\overline{V}_{\text{fg}}$ and width σ_{fg}),
  and the parameters $\eta, \overline{V}_{\text{fg}}, \sigma_{\text{fg}}$ are chosen to maximise the overall likelihood of the observed catalogue.

  One can then determine the posterior interloper probability for each star i: $p_{\text{fg},i} = \dfrac{\eta\, f_{\text{fg}}(V_{\text{los},i})}{f_{\text{mix}}(V_{\text{los},i})}$.

  One can extend the mixture model to handle multiple populations in the stellar system.



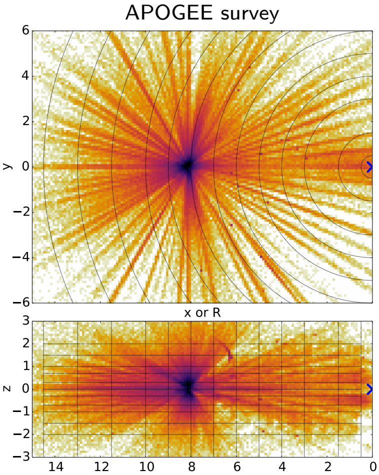
Dynamical modelling with resolved data: complications

- Missing phase-space coordinates
- Observational uncertainties
- Foreground contamination
- Survey footprint (a.k.a. selection function S)

  need to account for the observable volume in the normalisation of the DF in the model:

  use $$f_{\text{survey}} \equiv \frac{f(\mathbf{x},\mathbf{v})}{\iiint \mathrm{d}^3\mathbf{x} \iiint \mathrm{d}^3\mathbf{v}\; f(\mathbf{x},\mathbf{v})\; S(\mathbf{x},\mathbf{v})}.$$

  This integral can be very expensive!



Dynamical modelling with resolved data: complications

- Missing phase-space coordinates
- Observational uncertainties
- Foreground contamination
- Survey footprint (a.k.a. selection function S)

  need to account for the observable volume in the normalisation of the DF in the model:

  use $$f_\mathrm{survey} \equiv \frac{f(\mathbf{x},\mathbf{v})}{\iiint \mathrm{d}^3\mathbf{x} \iiint \mathrm{d}^3\mathbf{v}\; f(\mathbf{x},\mathbf{v})\; S(\mathbf{x},\mathbf{v})}.$$

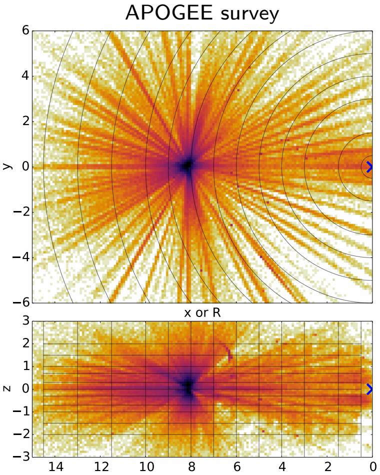
  This integral can be very expensive!

  Alternative approach:

  use conditional velocity distribution:

  $$\mathfrak{f}(\mathbf{v}\mid\mathbf{x}) \equiv \frac{f(\mathbf{x},\mathbf{v})}{\iiint \mathrm{d}^3\mathbf{v}\; f(\mathbf{x},\mathbf{v})} = \frac{f(\mathbf{x},\mathbf{v})}{\rho(\mathbf{x})}.$$

  (MW disc: Binney & Vasiliev 2023, 2024; nuclear stellar disc: Sormani+ 2022; nuclear star cluster: Vasiliev+ 2026)

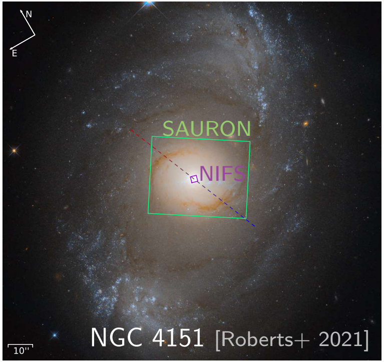


Dynamical modelling with integrated-light data

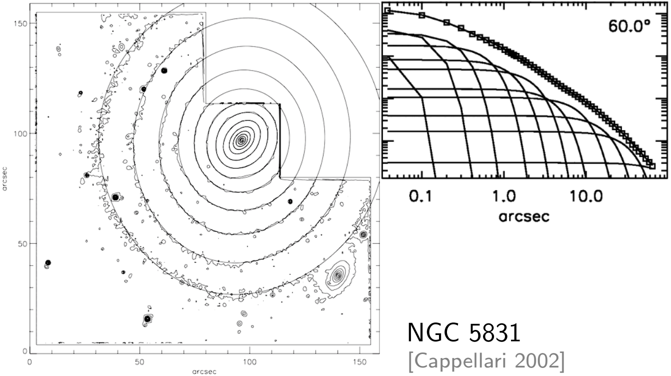
(spatially resolved, but each spaxel contains light of many stars)

Photometry usually covers a much larger spatial region than kinematic data; surface brightness profile is typically fitted with some parametric models (Sérsic, Nuker, disks, etc.) or multi-Gaussian profiles (MGE).

Major problem: deprojection (determination of 3d shape from the 2d image) is quietly sidestepped in most cases by using ellipsoidal profiles.

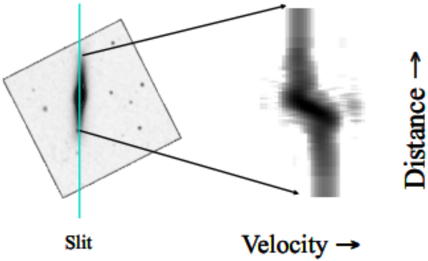


NGC 4151 [Roberts+ 2021]

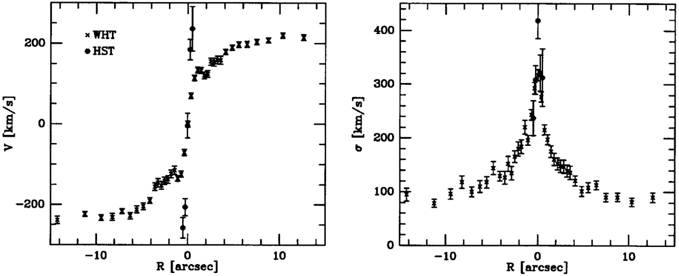


NGC 5831 [Cappellari 2002]

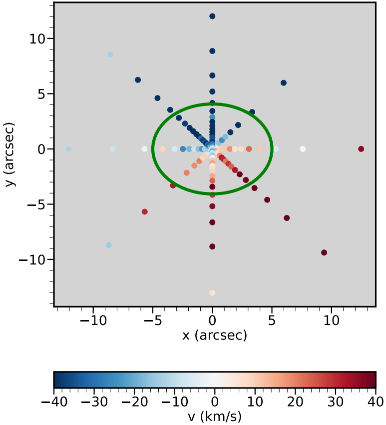
Kinematics: long-slit spectroscopy



[credit: Wikipedia]

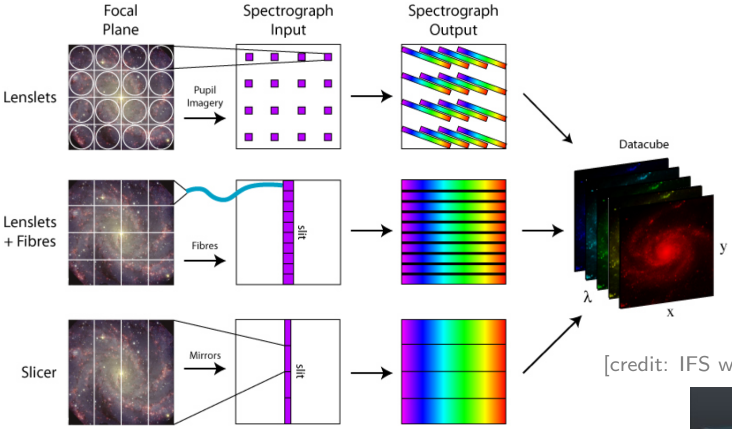


NGC 4342 [Cretton & van den Bosch 1999]

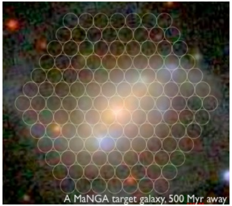


Abell 1314 [de Nicola+ 2026]

Kinematics: integral-field spectroscopic units



[credit: IFS wiki]



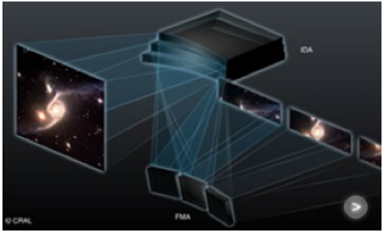
MaNGA IFU

Typical field of view:

$\sim 30-60''$ with natural seeing ($\sim 1''$ PSF) – SAURON, MaNGA, SAMI, MUSE;

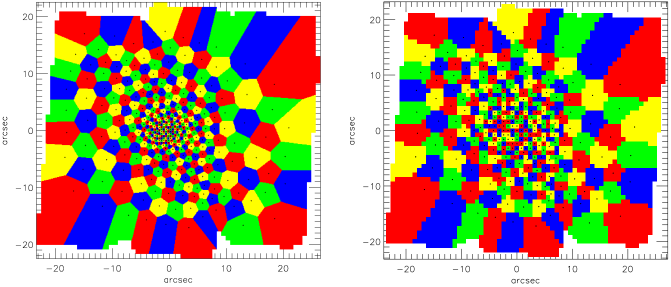
$\sim 3-8''$ with adaptive optics ($\sim 0.1''$ PSF) – NIFS, SINFONI, MUSE-AO, JWST NIRSpec.

Low to medium spectral resolution (1500–10000).



muse
multi unit spectroscopic explorer

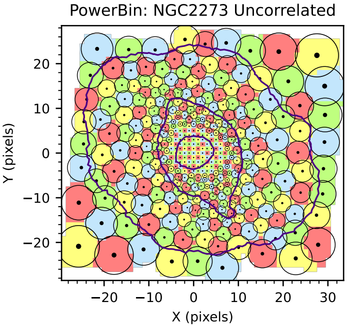
IFU kinematics: spatial binning

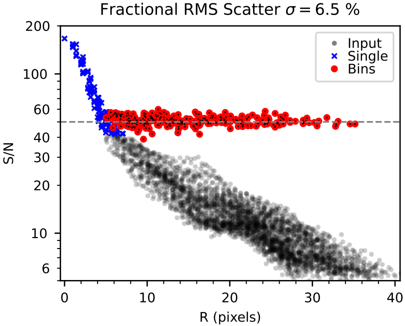


Group nearby spaxels to keep the S/N ratio roughly constant.

Voronoi binning

[Cappellari & Copin 2003]

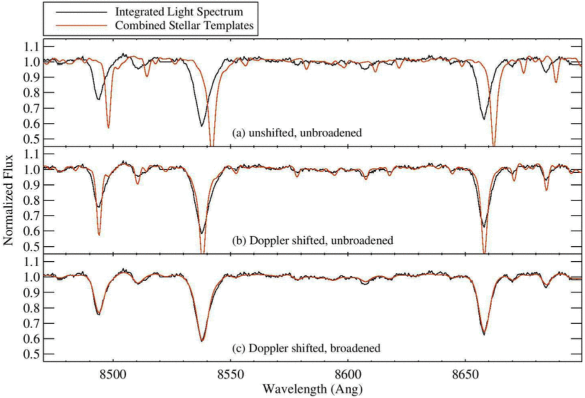

PowerBin: NGC2273 Uncorrelated

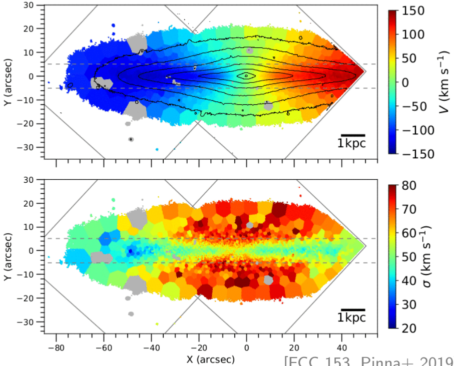

Fractional RMS Scatter $\sigma = 6.5$ %

New approach: PowerBin

[Cappellari 2025]

IFU kinematics: velocity distribution



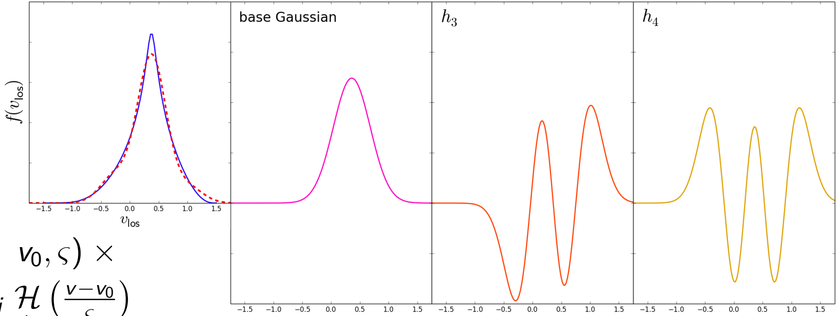


[FCC 153, Pinna+ 2019]

spectra of a few template stars $\circledast$ shift and broadening by adjustable velocity distribution function $\Rightarrow$ spectrum in each spaxel

IFU kinematics: velocity distribution

Gauss–Hermite parameterisation of LOSVDs [van der Marel & Franx 1993; Gerhard 1993]

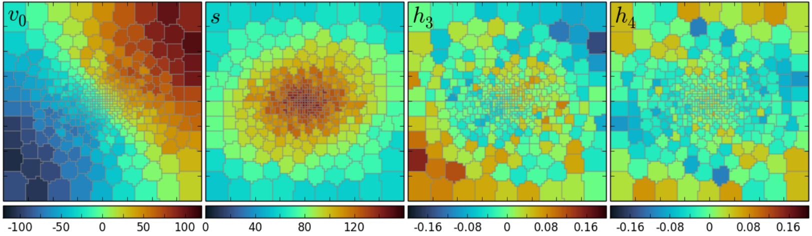


base Gaussian

$$f(v) = \mathcal{N}(v;\ v_0, \varsigma) \times \sum_i h_i\, \mathcal{H}\left(\frac{v-v_0}{\varsigma}\right)$$

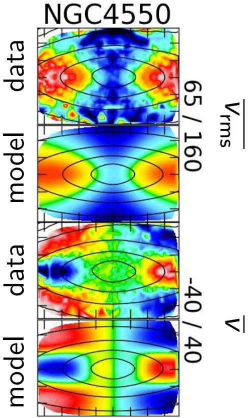
GH coefficients

Hermite polynomials

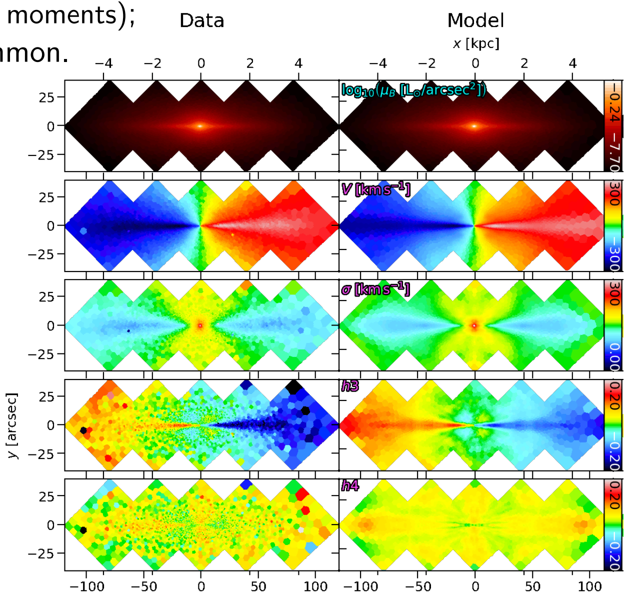


Dynamical modelling with integrated-light data

Most studies use Jeans models (with only $\overline{v}$ and σ)
or Schwarzschild models (with GH moments);
DF and M2M models are less common.

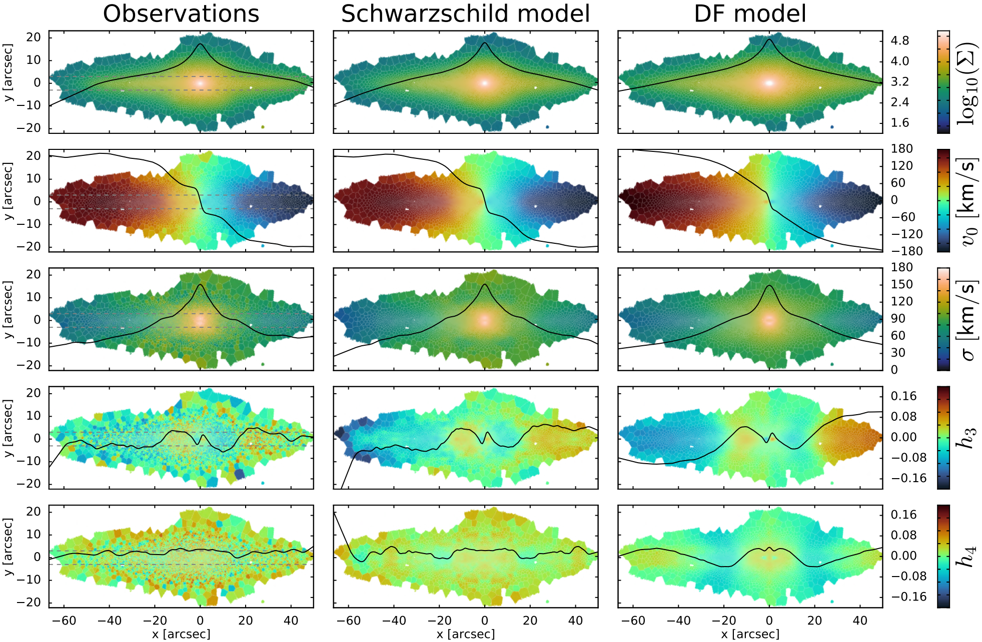


JAM, NGC 4550 [Cappellari+ 2011]



Schwarzschild, NGC 3115 [Poci+ 2019]

Dynamical modelling with integrated-light data



FCC 170 [Galán-de Anta+ 2023]

Summary

- different methods have different tradeoffs (performance, accuracy, interpretability)
- use multiple methods for cross-checking
- ...

Remedios Varo – Star maker