Generative Adversarial Network for Identifying the Dark Matter Distribution of a Dwarf Spheroidal Galaxy

Sung Hak Lim
Rutgers University

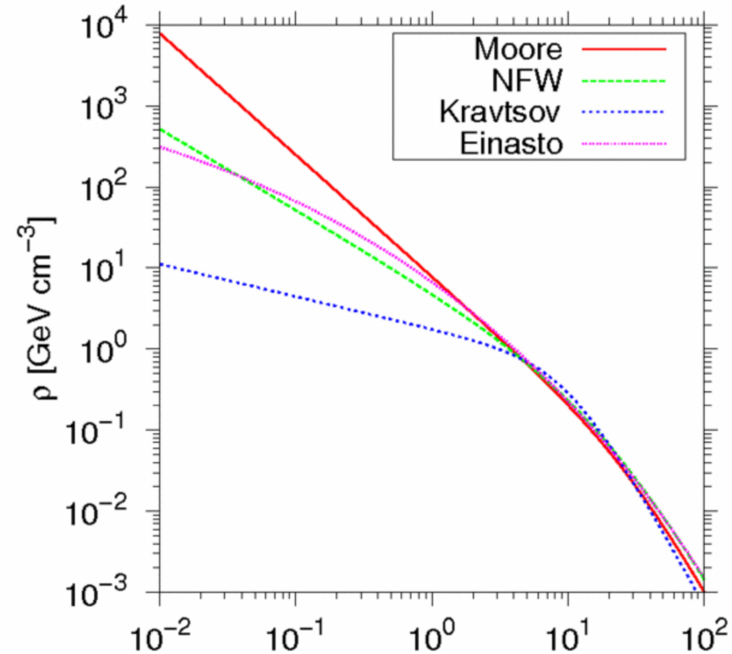
4th Inter-experiment Machine Learning workshop
Oct. 2020

K. Hayashi, S. Horigome, **S. H. Lim**, S. Matsumoto, M. M. Nojiri,
Work in progress...

Dwarf Spheroidal Galaxy and Dark Matter

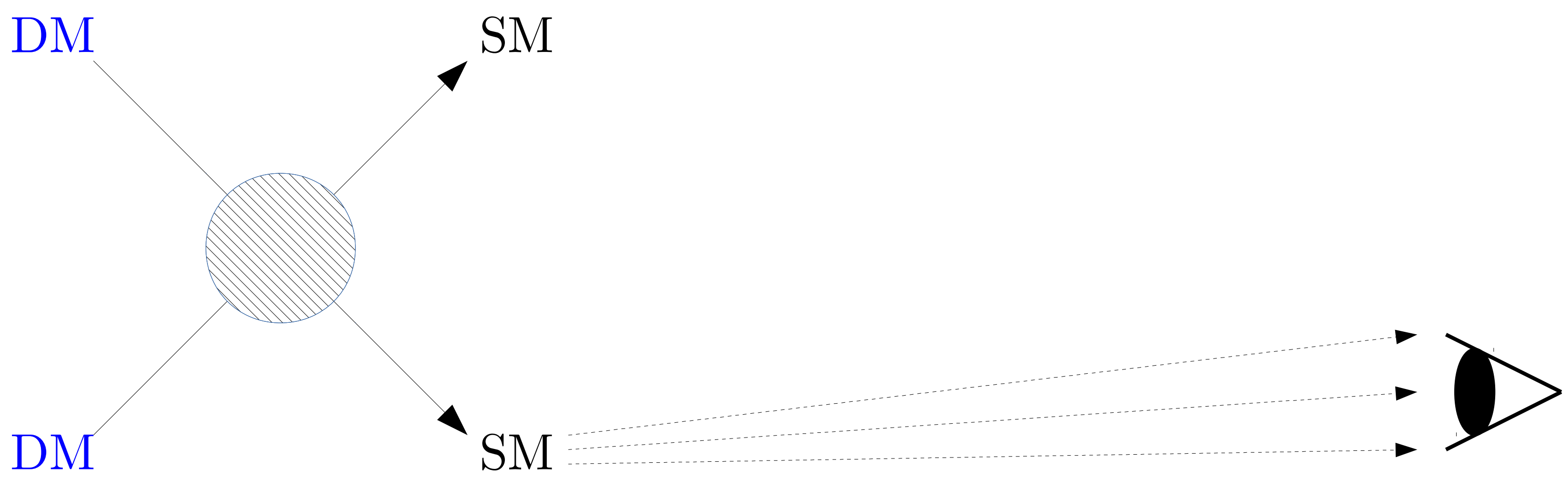
- Dwarf Spheroidal Galaxy (dSph): a small, faint galaxy with little dust, and an old stellar population.
- Why this galaxy is useful?
- dSph may contain lots of dark matter (DM). The visible stars can be considered as a probe of the DM. It is a good source for studying dark matter kinematics.
- For example, understanding the kinematics of DM is essential in the indirect detection of DM.

The Fornax dwarf galaxy
https://www.eso.org/public/images/eso1007a/



Example dark matter profiles
C. Rott, arXiv:0912.5183

Indirect Detection and Dark Matter Distribution



The signal flux is proportional to the squared dark matter density (J-factor)

$$\Phi(\Delta\Omega) \propto \rho_{\rm DM}^2$$

and it is important to know the dark matter density precisely for the high-quality analysis.

However, accessible information from dSph is limited since we are watching far away stars from one direction.
ML may help systematically analyzing dSph with limited information.

As the first part of this project,
we introduce an inference model for **DM distribution of dSph**, using **GAN**.

Kinematics of dSph

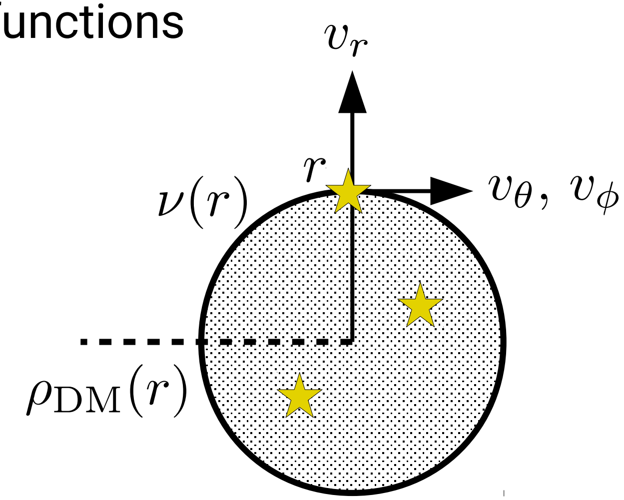
Spherically symmetric case

Kinematics of visible stars of dSph is parameterized by unary functions so that NN analysis can be useful.

Number density of the stars: $\nu(r)$

Variance of radial velocity $\overline{v_r^2}(r)$ $\quad v_i \sim N(0, \overline{v_i^2}(r))$

Variance of tangential velocity $\overline{v_t^2}(r) = \overline{v_\theta^2} + \overline{v_\phi^2}$



For the tangential velocity, we often use the following anisotropy function instead of the velocity itself.

$$\beta_{\mathrm{ani}}(r) = 1 - \frac{\overline{v_\theta^2} + \overline{v_\phi^2}}{2\overline{v_r^2}} = \begin{cases} 1 & \text{radially biased} \\ -\infty & \text{tangentially biased} \end{cases}$$

These parameter functions can be identified by analyzing the kinematics of the visible stars.

$$\{(x, y, z, v_r, v_\theta, v_\phi)_i | i = 1, \cdots, N\}$$

Also, one good thing is that the statistical modeling of the stellar distribution is factorizable.

$$p(\vec{r}, \vec{v}) = p(\vec{v}|\vec{r})p(\vec{r})$$

Position: density estimation

$$p(\vec{r}) = \nu(r)$$

Velocity: parameter inference

$$p(v_i|\vec{r}) \sim N(0, \overline{v_i^2}(r))$$

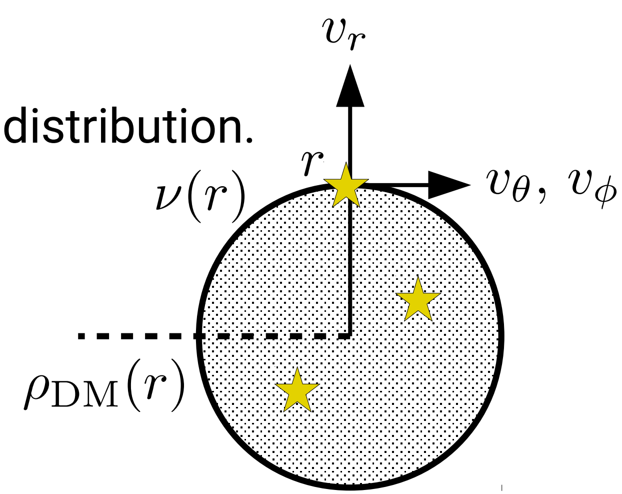
We could fit those functions one-by-one.

How can we infer the DM distribution from these functions?

Interaction between DM and stars: Jeans Equation

Spherically symmetric case

These visible stars are assumed to be moving under the gravitational potential of DM distribution.
We may use the equation of motion in order to identify the DM distribution.



Jeans equation

$$\frac{d\overline{v_r^2}}{dr} + \left[\frac{1}{\nu}\frac{d\nu}{dr} + \frac{2\beta_{\rm ani}}{r}\right]\overline{v_r^2} = -\frac{d\Phi}{dr}$$

Poisson equation for gravity

$$\frac{1}{r^2}\frac{d}{dr}r^2\frac{d\Phi}{dr} = 4\pi G\,\rho_{\rm DM},\quad \frac{d\Phi}{dr} = \frac{G}{r^2}\int_0^r dr'\,\rho_{\rm DM}(r')\,4\pi r'^2$$

For spherically symmetric case, a closed form solution of the Jeans equation exists.

$$\overline{v_r^2}(r) = \frac{1}{z(r)\nu(r)}\int_r^\infty dr'\,z(r')\nu(r')\frac{d\Phi}{dr'},\quad \log z(r) = \int_{r_0}^r dr'\,\frac{2\beta_{\rm ani}(r')}{r'}$$

After discretizing all the operations above,
the relation between the variance of radial velocity and the DM density becomes linear.

$$\overline{v_r^2}(r) = \sum_{r'} A(r, r')\rho_{\rm DM}(r')$$

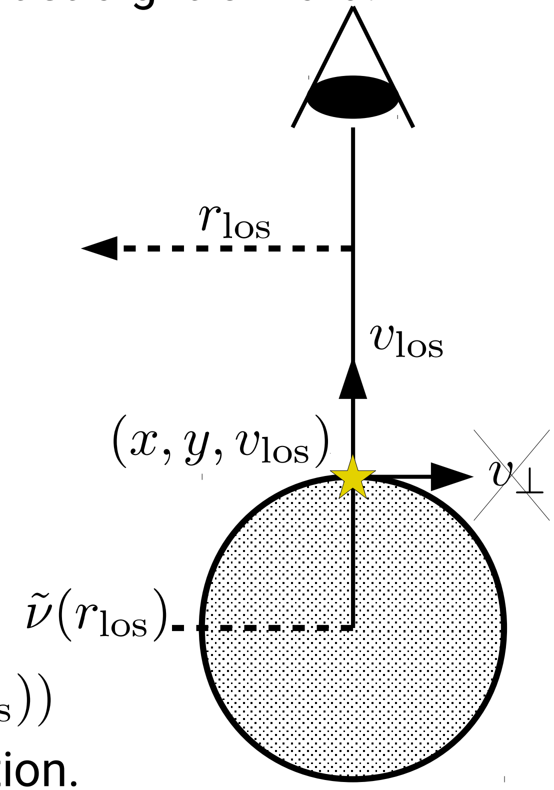
This linear expression is convenient for embedding this to an NN architecture.

Line-of-sight observables

Spherically symmetric case

However, we cannot access the full data so that the analysis is not straightforward.
We are observing far away stars from a single direction.

$$(x, y, z, v_r, v_\theta, v_\phi) \to (x, y, v_{\rm los})$$



The full parameters can not be inferred directly from the data,

Number density of stars: $\nu(r)$

Variance of radial velocity $\overline{v_r^2}(r)$

Velocity anisotropy $\beta_{\rm ani}(r)$

But we have to infer them from projected profiles

Number density of stars: $\tilde{\nu}(r_{\rm los})$

Variance of line-of-sight velocity $\overline{v_{\rm los}^2}(r_{\rm los})$ $\quad v_{\rm los} \sim N(0, \overline{v_{\rm los}^2}(r_{\rm los}))$

Two degrees of freedom of the velocity are lost during the projection.
The velocity anisotropy is assumed to compensate the lost information.

cf. Analytical solution for the inversion exist (Abel transformation)

$$\nu(r) = -\frac{1}{\pi}\int_r^\infty \frac{d\tilde{\nu}}{dr_{\rm los}} \frac{dr_{\rm los}}{\sqrt{r_{\rm los}^2 - r^2}}$$

See also Ullio and Valli, JCAP 07 (2016) 025, arXiv:1603.07721 for full closed form solution of the dark matter density for the symmetric case.

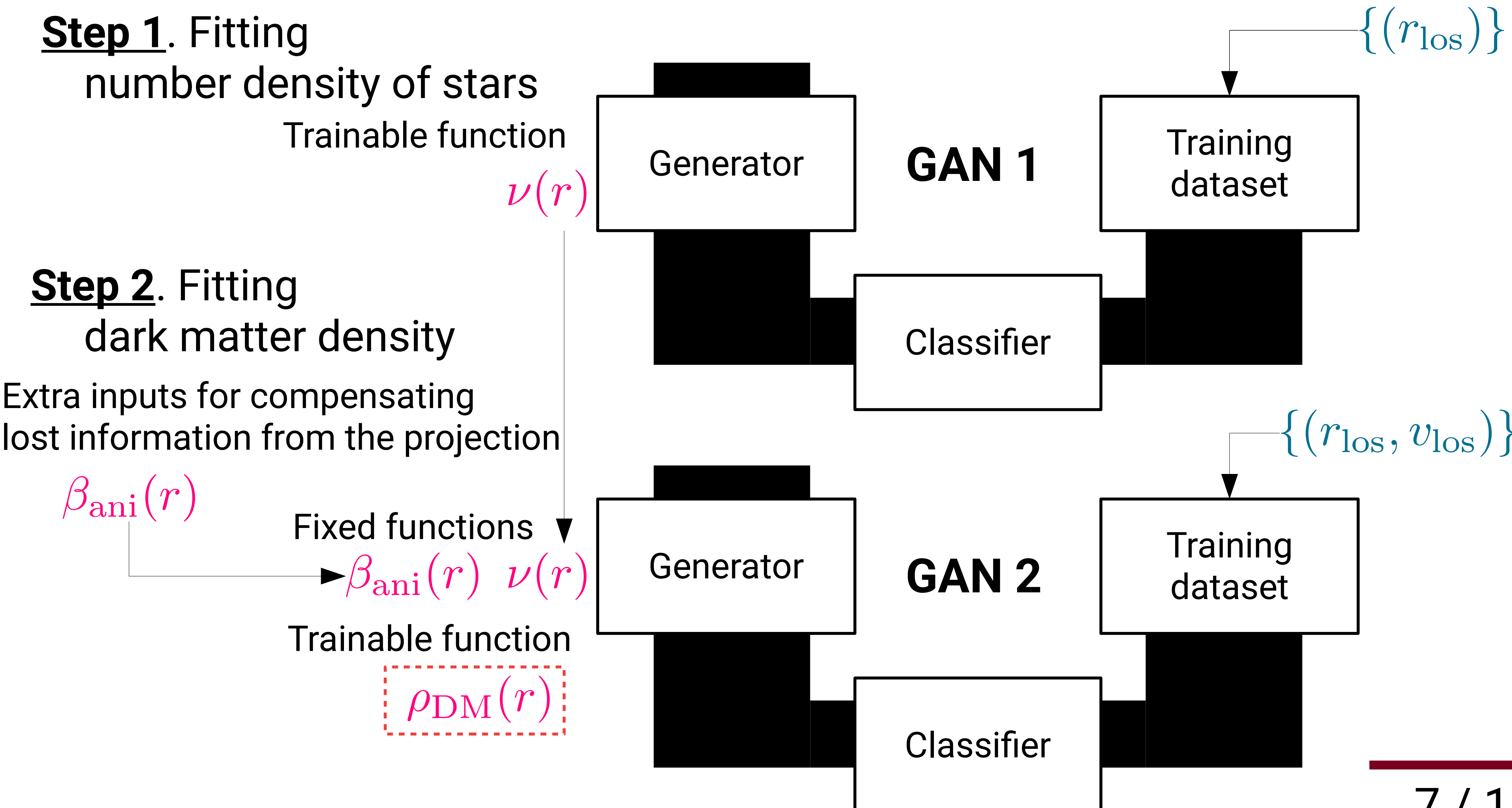
But if we use **GAN**, we do not need this inversion formula, and this may help for dealing with more general cases.

Fitting strategy using GAN

By using GAN, we can directly fit the dark matter density from the sample dataset.

Step 1. Fitting number density of stars

Step 2. Fitting dark matter density



Benchmark point

We consider the following benchmark point for the training sample.

Number density of stars: Plummer model

$$\nu(r) = \frac{3}{4\pi} \frac{1}{(1 + r^2/r_{\rm half}^2))^{5/2}}$$

$$r_{\rm half} = 1\,{\rm kpc}$$

Dark matter density: NFW profile

$$\rho_{\rm DM}(r) = \rho_0 \left(\frac{r}{r_s}\right)^{-\gamma} \left[1 + \left(\frac{r}{r_s}\right)^{\alpha}\right]^{-(\beta-\gamma)/\alpha}$$

$$\rho_0 = 4 \times 10^7\,{\rm M}_\odot/{\rm pc}^3$$
$$r_s = 1\,{\rm kpc}$$
$$(\alpha, \beta, \gamma) = (1, 3, 1)$$

Velocity anisotropy: Osipkov-Meritt model (Remind that this function is an <u>input</u> for the fitting.)

$$\beta_{\rm ani}(r) = \frac{r^2}{r^2 + r_a^2}$$

$$r_a = 1\,{\rm kpc}$$

Number of stars in the training sample: 10,000

We <u>**do not assume**</u> any particular set of number density profile and dark matter density during the GAN training.

We will show the results only using <u>the same</u> anisotropy function for the training sample and GAN training.
But <u>**they can be different in general**</u>, and studying an impact on that will be a future study.

Fit: number density of visible stars

Generator model: **Normalizing flow** from Gaussian distribution N(0,1) to the number density

$$\nu(r) = \mathrm{NF}(r)$$

(Sylvester flow is used.
Inverse flow is calculated by Newton's method)

Projection: the sampled dataset from NF is projected to line-of-sight observables

$$(r, \theta, \phi) \to (r_{\mathrm{los}})$$

Classifier model: closed form solution of the CE minimization.
(Kernel density estimation is used.)

$$p(y|r_{\mathrm{los}}) = \frac{p(r_{\mathrm{los}}|y)}{\sum_y p(r_{\mathrm{los}}|y)} \qquad p(r_{\mathrm{los}}|y) = \tilde{\nu}(r_{\mathrm{los}}) = \frac{1}{N_y}\sum_{i=1}^{N_y} K_h(r_{\mathrm{los}} - r_{\mathrm{los}}^{(i)})$$

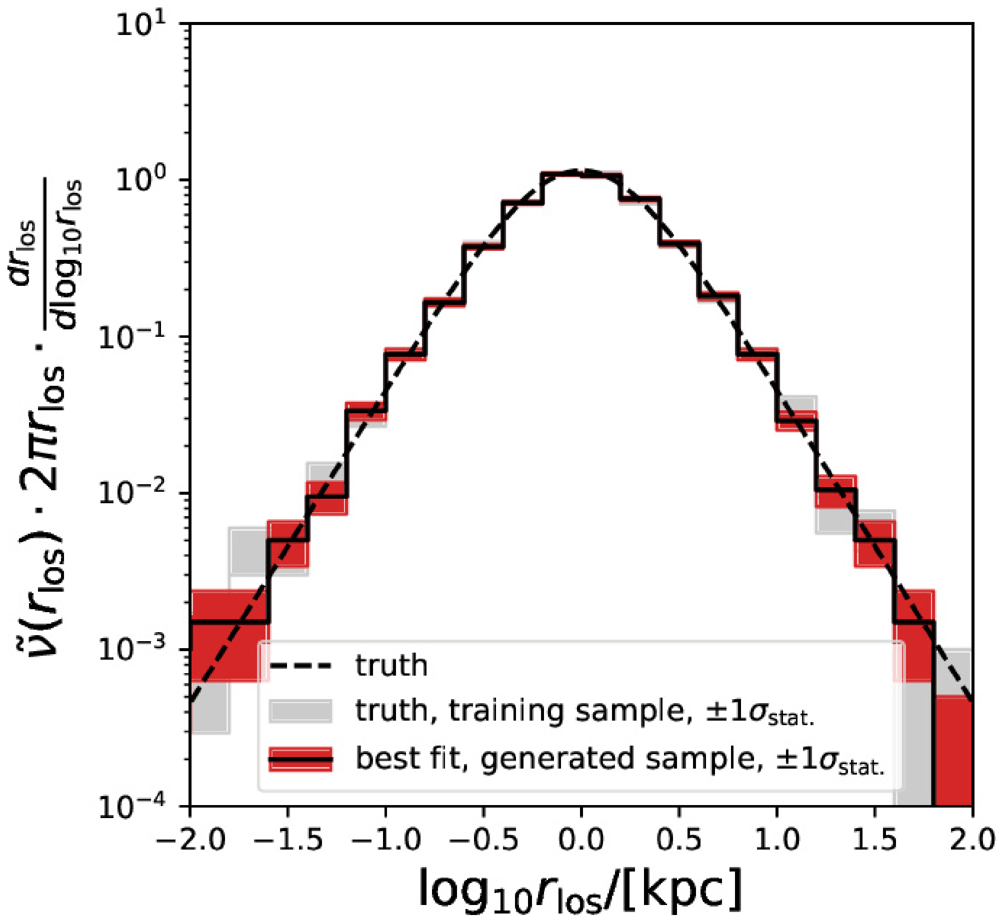
Additional regularizer: monotone condition on the number density.
The solution with increasing density is penalized.

$$\mathcal{L}_{\mathrm{reg,phy}} = \int \log_{10} r \left[\mathrm{ReLU}\left(\frac{d\,\log\nu}{d\,\log_{10} r}\right)\right]^2$$

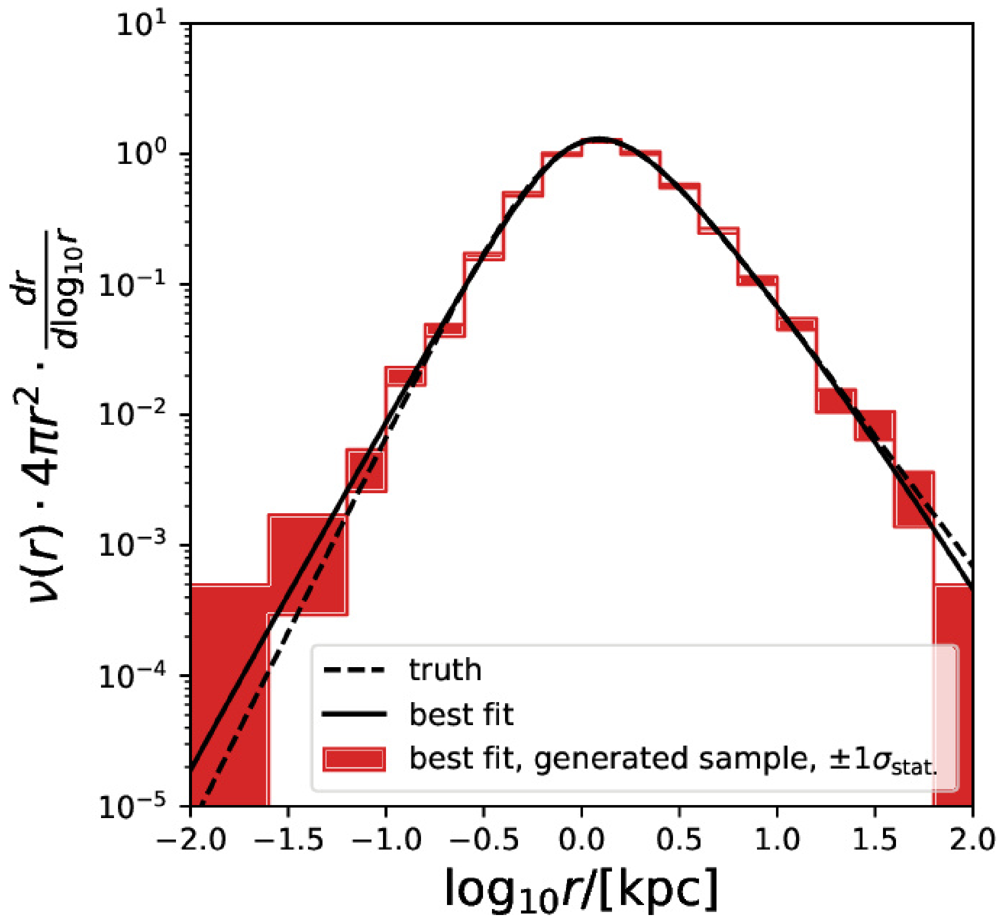
(this regularizer is identical to rectified Gaussian prior on the derivative)

Goodness of fit: likelihood ratio, KS test, …
(We are comparing two 1D probability densities of r_{los})

Results



Projected number density
for evaluating the goodness of fit.



The inferred number density.

Fit: dark matter density and variance of velocity

Generator model: Two-level model with reparameterization layer and a generic function model for ρ_{DM}.

$$p(v_r, v_\theta, v_\phi|\vec{r}) \sim (N(0, \overline{v_r^2}(r)), N(0, \overline{v_t^2}(r)), N(0, \overline{v_t^2}(r)))$$

$$v_i = \sqrt{\overline{v_i^2}(r)} \times N(0,1)$$

$$\overline{v_r^2}(r) = \sum_{r'} A(r, r')\rho_{\mathrm{DM}}(r')$$ discretized solution of the Jeans equation

$$\overline{v_t^2}(r) = 2(1-\beta_{\mathrm{ani}}(r))\overline{v_r^2}(r)$$ Use anisotropy function.

Projection:

the sampled dataset from normal distributions is projected to line-of-sight observables

$$(r, \theta, \phi, v_r, v_\theta, v_\phi) \to (r_{\mathrm{los}}, v_{\mathrm{los}})$$

Classifier model: Two-level model with logistic regression on v_{los}^2

$$\mathrm{logit} \circ p(y|r_{\mathrm{los}}, v_{\mathrm{los}}) = \mathrm{MLP}_1(r_{\mathrm{los}})v_{\mathrm{los}}^2 + \mathrm{MLP}_2(r_{\mathrm{los}})$$ because $v_{\mathrm{los}} \sim N(0, \overline{v_{\mathrm{los}}^2}(r))$

Additional regularizer: monotone condition on the dark matter density.
The solution with increasing density is penalized.

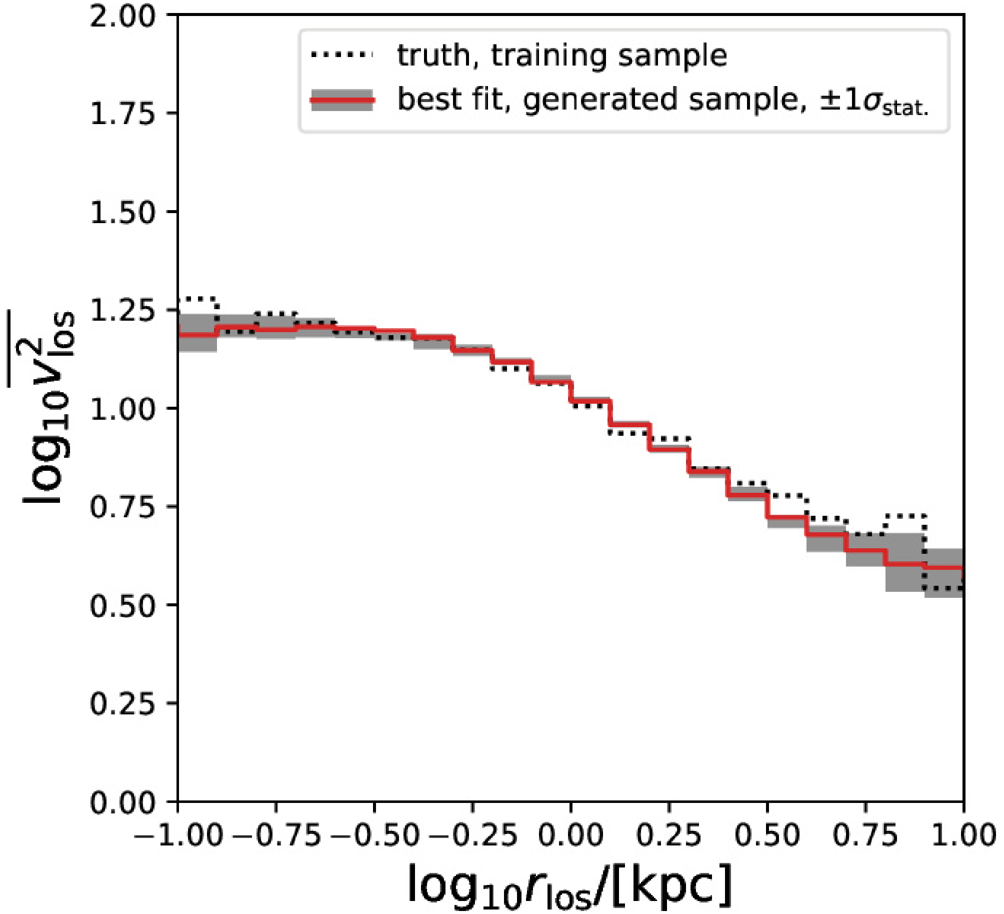
$$\mathcal{L}_{\mathrm{reg,phy}} = \int \log_{10} r \left[\mathrm{ReLU}\left(\frac{d\ \log \rho_{\mathrm{DM}}}{d\ \log_{10} r}\right)\right]^2$$

(this regularizer is identical to rectified Gaussian prior on the derivative)

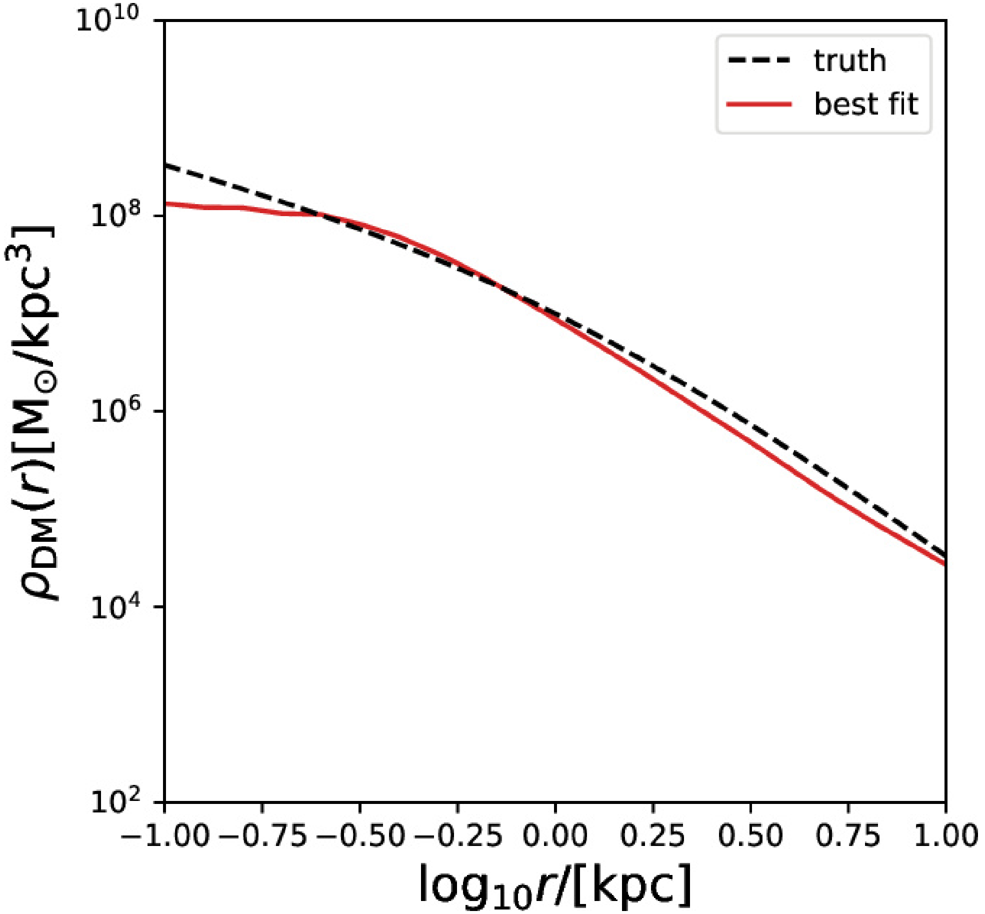
Goodness of fit: F distribution

(We are comparing variances of v_{los} of training samples and generated samples.)

Results



Projected variance of the line-of-sight velocity for evaluating the goodness of fit.



The inferred dark matter density

Conclusion and Discussion

- We introduced an ML-based inference model for dark matter distribution of the dwarf spheroidal galaxy.
- The event generator for dSph can be converted to the inference model by using GAN.
- This method is not limited to analyzing dSph using a particular anisotropy function.
- This method does not require explicit inversion formula from projected distributions.
- The discussion on this talk is limited to a spherically symmetric case, but it is easily extendable to more generic cases thanks to the flexibility of ML architectures.
- If there are foreground stars, we may also use anomaly detection methods such as ANODE
  B. Nachman and D. Shih, Phys. Rev. D 101, 075042 (2020), arXiv:2001.04990
  M. Buckley, L. Necib, D. Shih and J. Tamanas, work in progress, (link)
  and so on.