Determining the Dark Matter distribution in galaxies with Deep Learning (2111.08725)

As part of the darkmachines projects challenges: https://darkmachines.org/

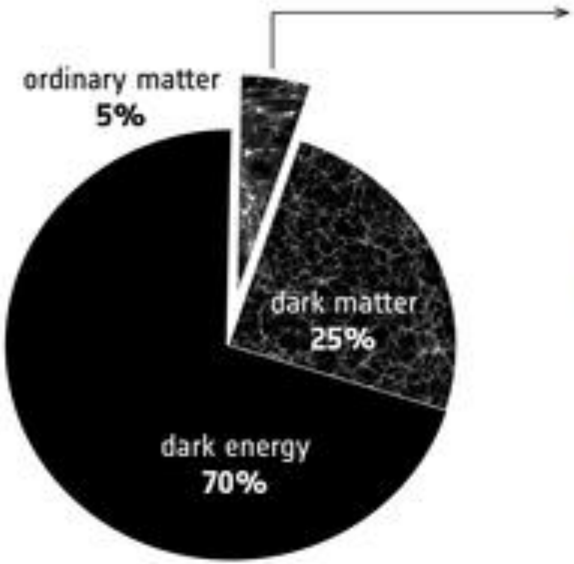
Martín de los Rios *IFT/UAM (Madrid Spain) - ICTP/SAIFR (Sao Paulo Brasil)*
Mihael Petac *CAC (Slovenia) - LUPM (France)*
Bryan Zaldivar *IFIC (Valencia Spain)*
Nina Bonaventura *Cosmic Dawn Center (Copenhagen Denmark)*
Francesca Calore *CNRS LAPTh (Annecy France)*
Fabio Iocco *Dipartimento di Fisica INFN (Napoles Italy)*

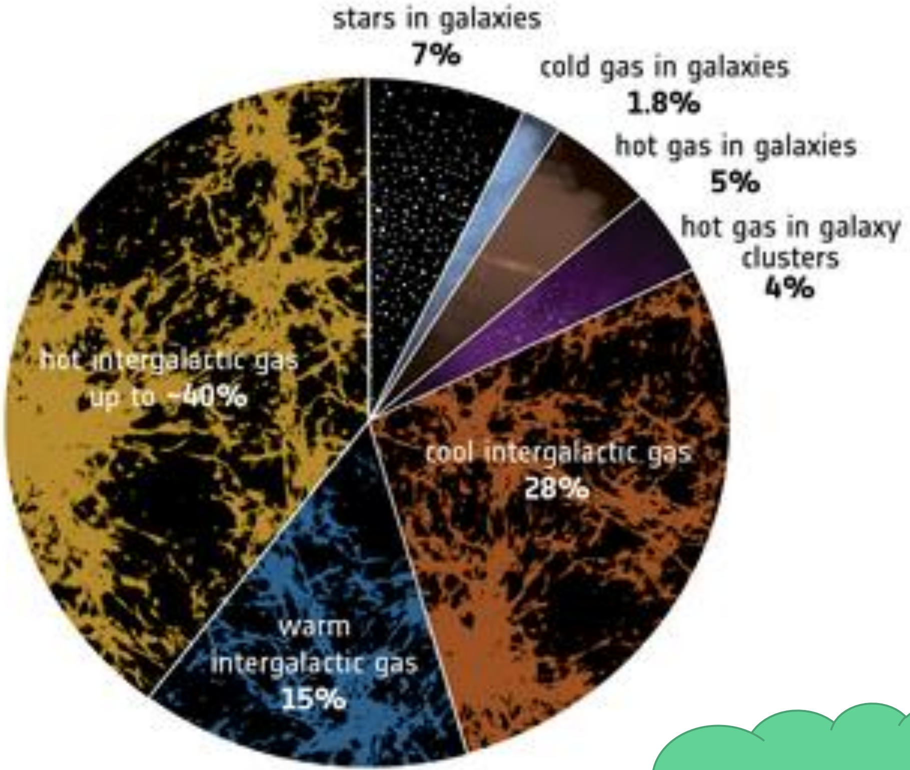
Outline

- Brief Introduction to Dark Matter
- Brief Introduction to Supervised Learning
- Constructions of the dataset
  - TNG100 Simulations (1707.03401, 1707.03395, 1707.03395, ...)
  - SKIRT (2003.00721)
  - MARTINI (https://github.com/kyleaoman/martini)
- Results
  - Prediction of the Dark matter profile
  - Comparison between different architectures
  - Comparison between different inputs
  - Comparison with Rotation Curve method
- Conclusions and Future work

Brief Introduction to Dark Matter

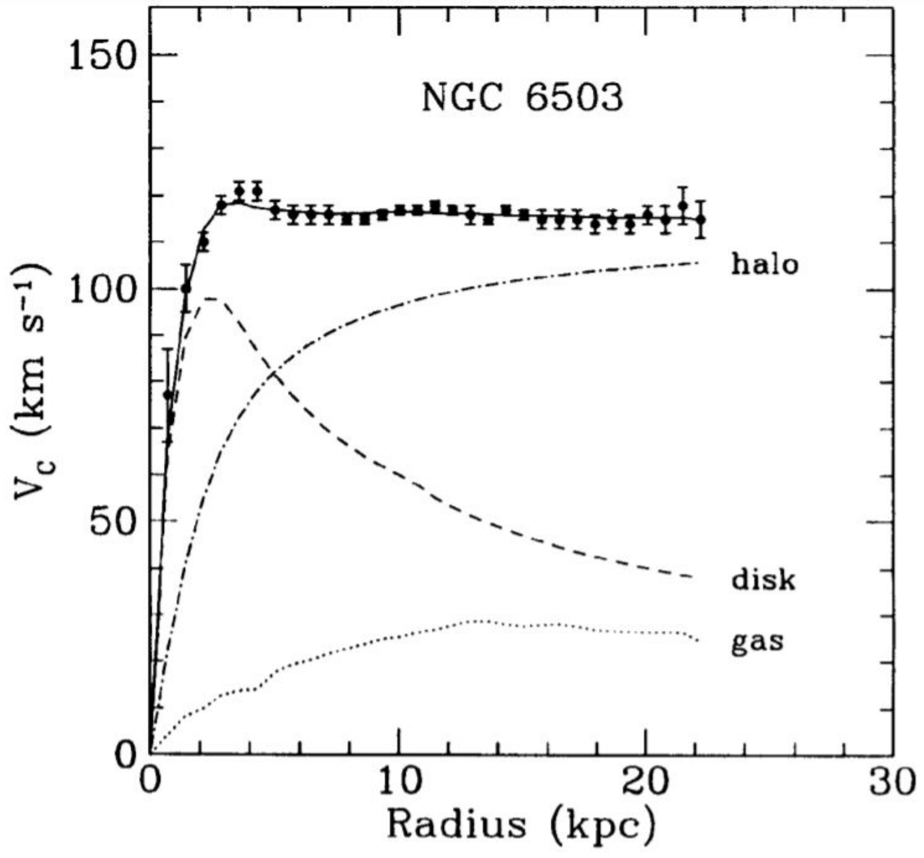
Cosmic Budget





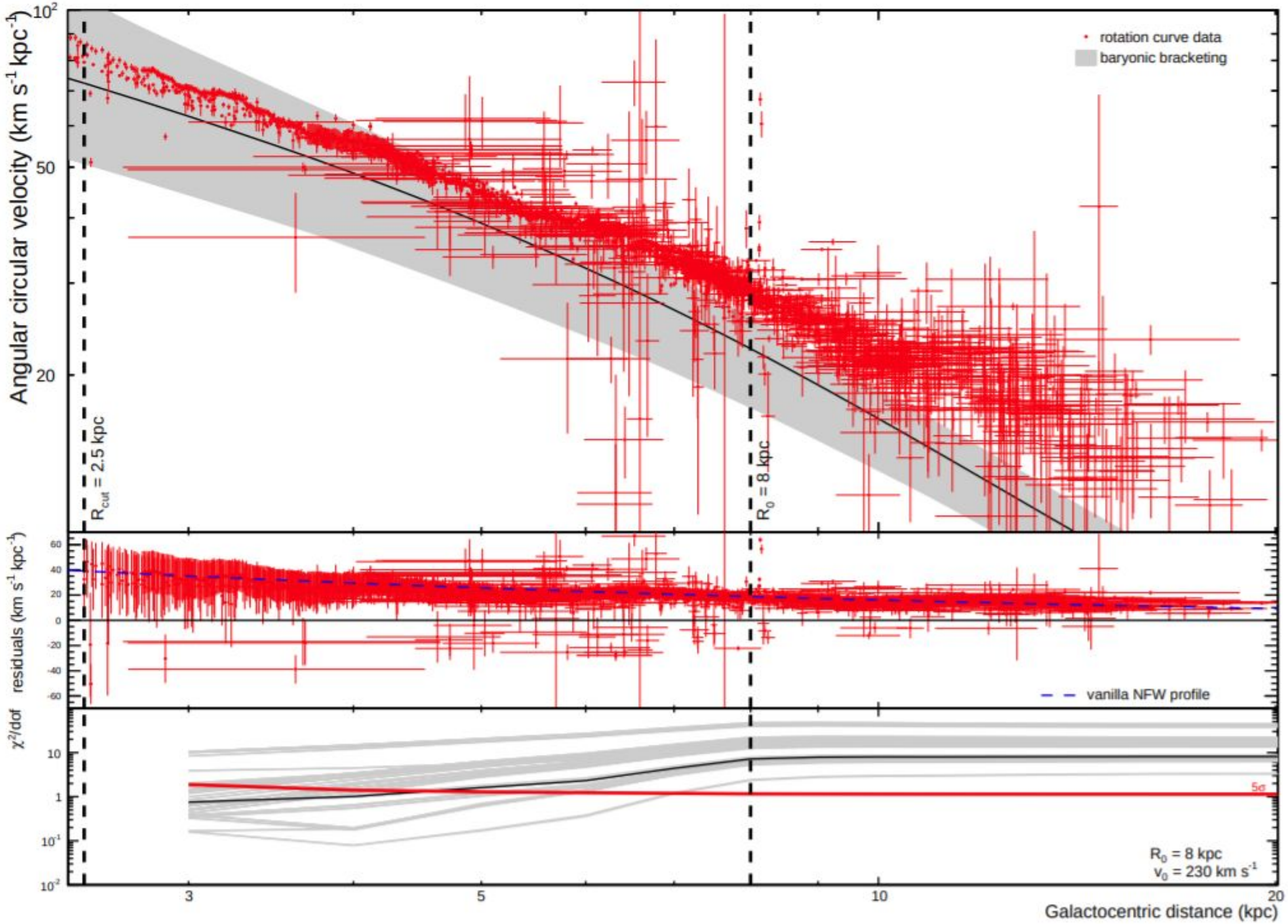
https://sci.esa.int/web/xmm-newton/-/60430-the-cosmic-budget-of-ordinary-matter

See Matteo Martinelli talk



NGC 6503 Rotation Curve

Katherine Freese
https://arxiv.org/abs/0812.4005



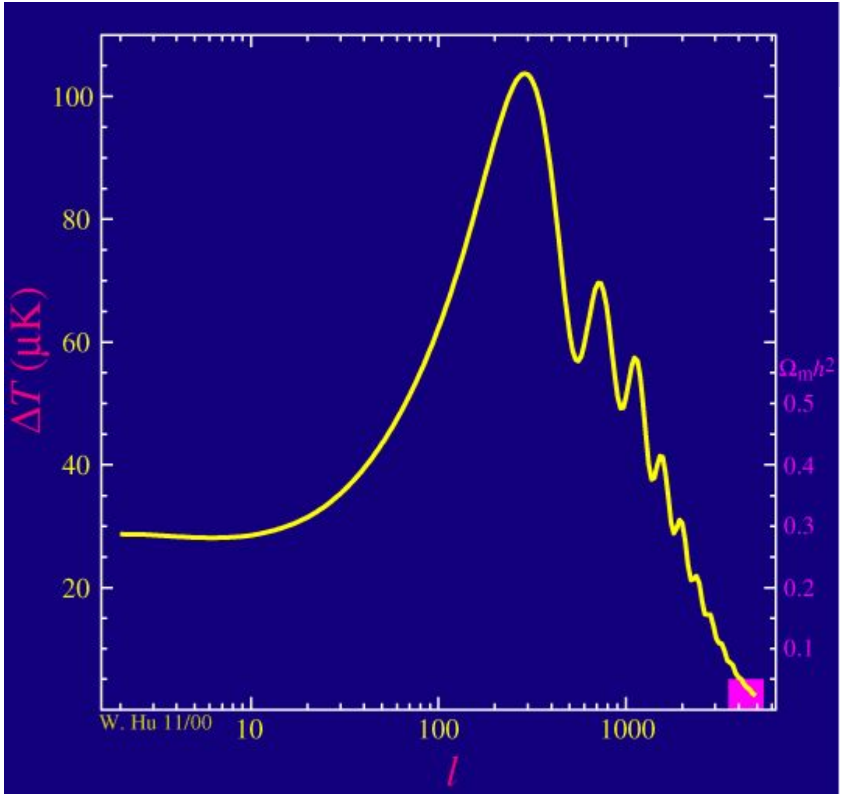
Evidence for dark matter in the inner Milky Way

(https://arxiv.org/abs/1502.03821)

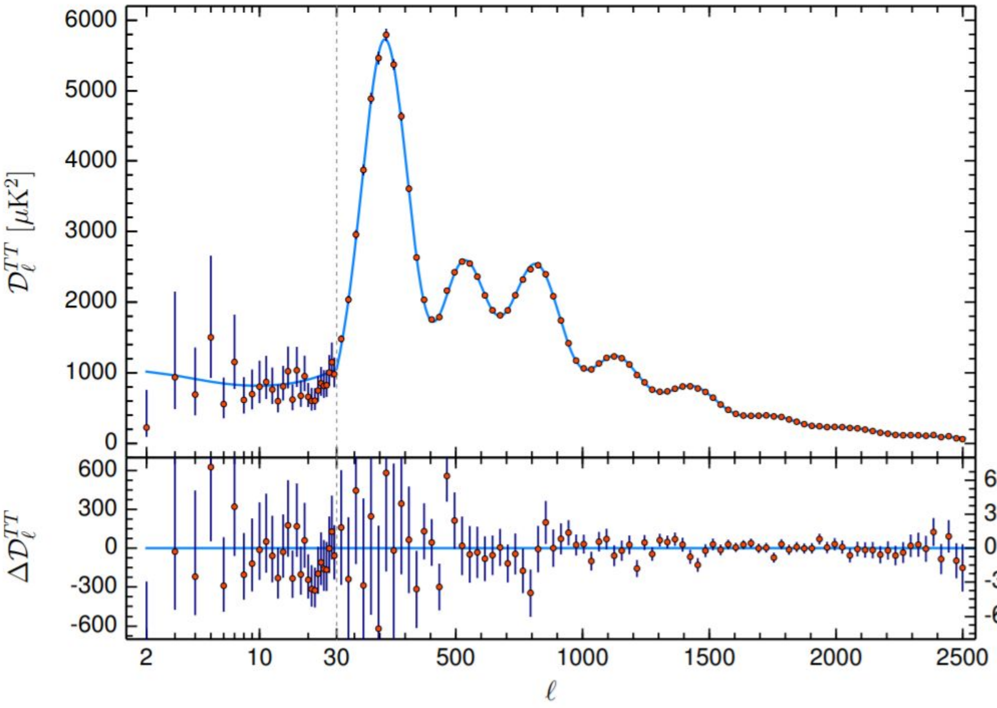
Fabio Iocco, Miguel Pato & Gianfranco Bertone

Bullet Cluster

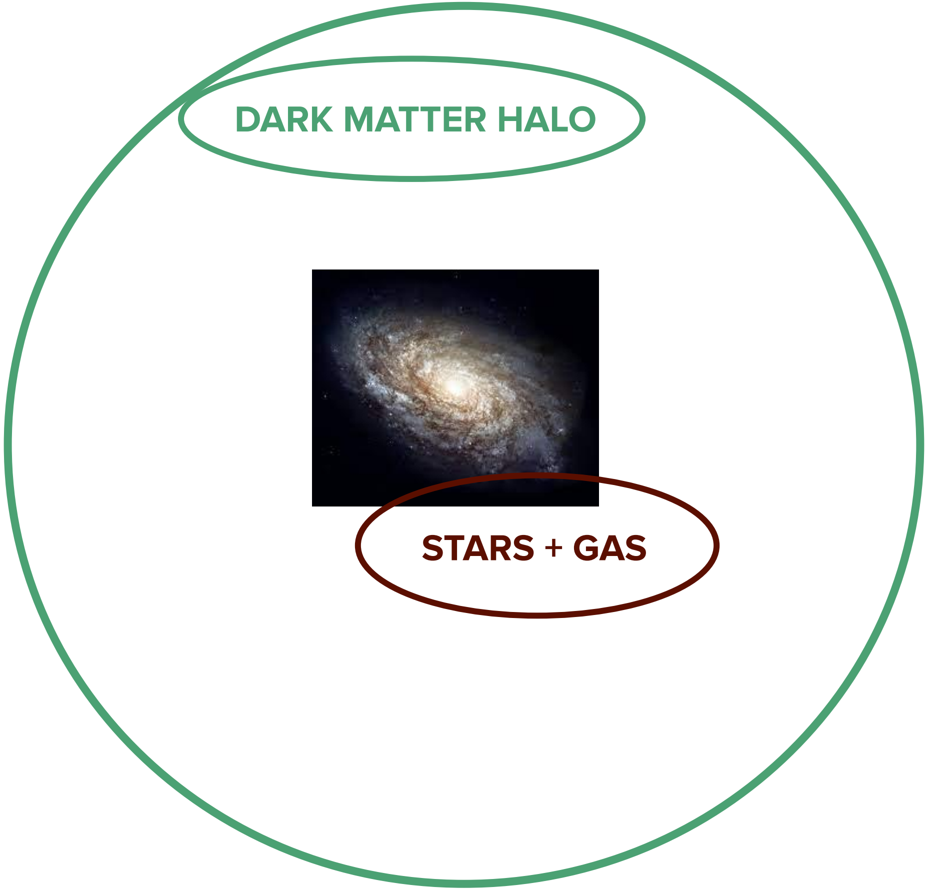
Markevitch et al.
https://arxiv.org/abs/astro-ph/0309303

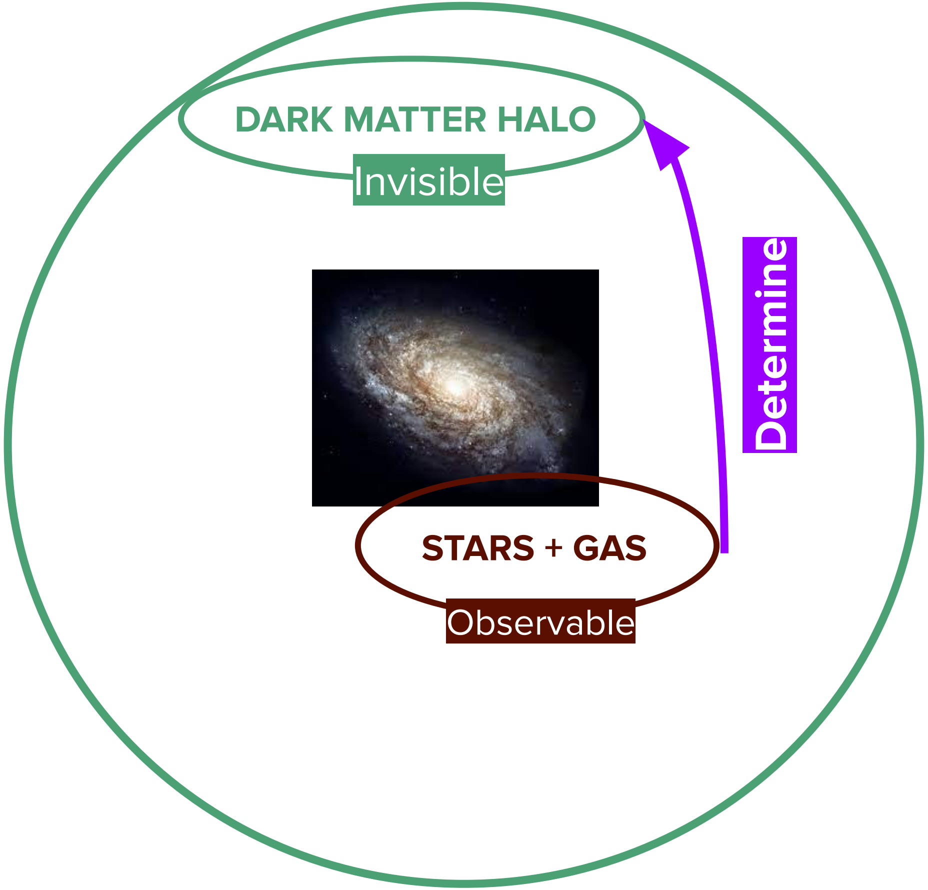


http://background.uchicago.edu/~whu/animbut/anim2.html

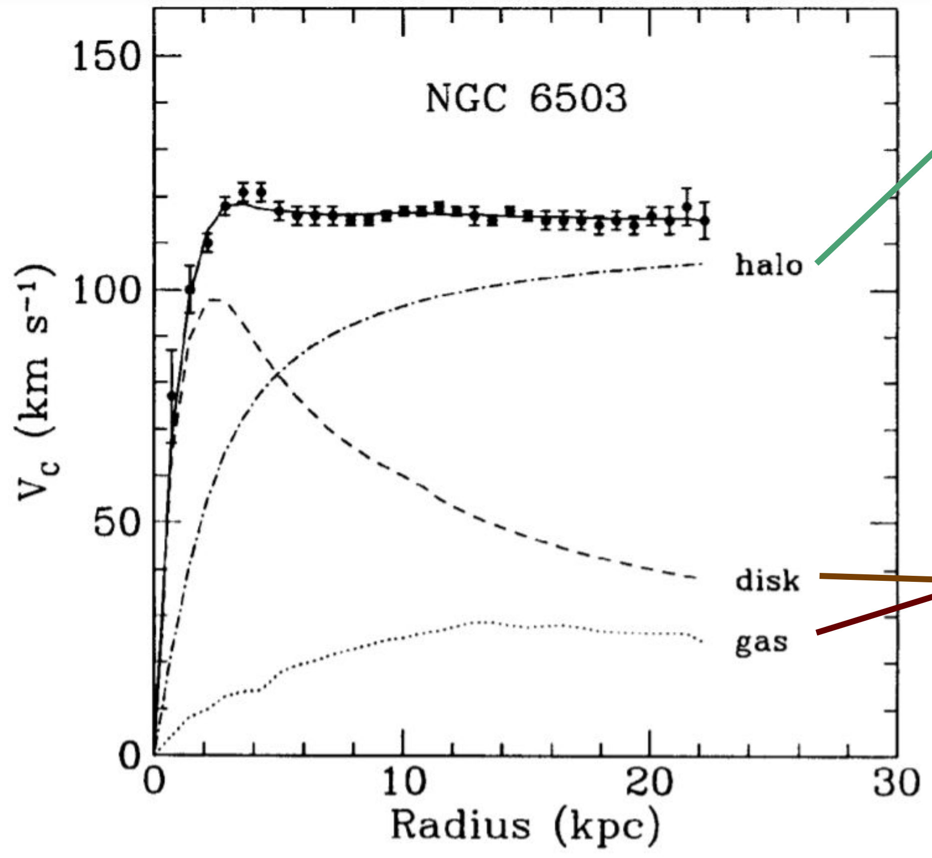


Planck Collaboration
https://arxiv.org/abs/1807.06209


DARK MATTER HALO
STARS + GAS


DARK MATTER HALO
Invisible
Determine
STARS + GAS
Observable

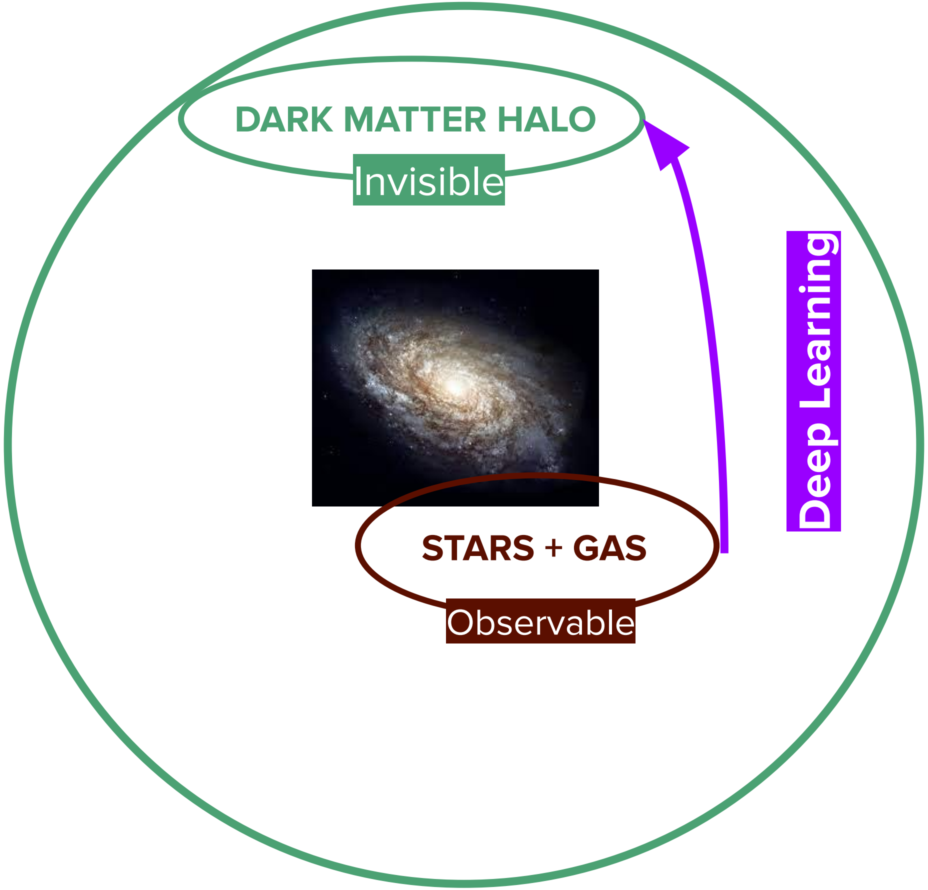
NGC 6503 Rotation Curve



Assume an NFW Profile

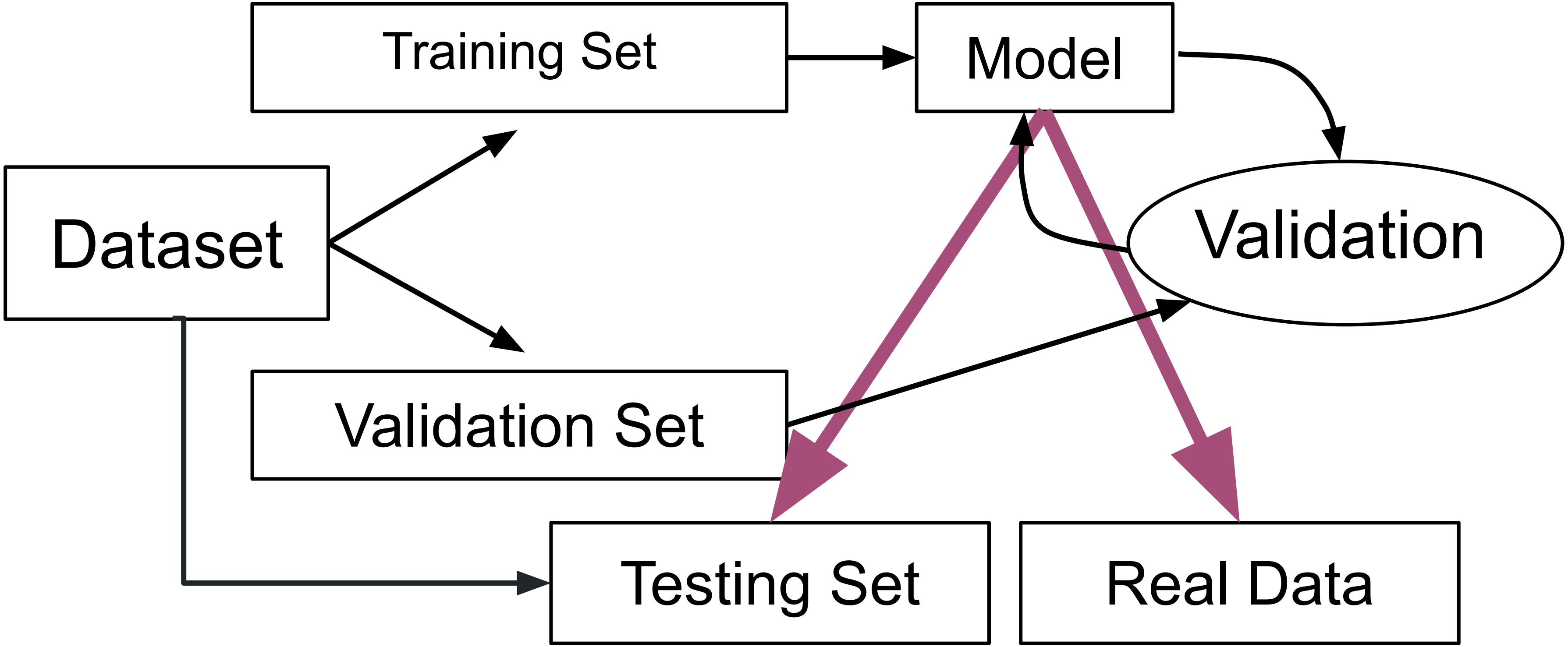
Estimated best-fit parameters

Inferred from observations

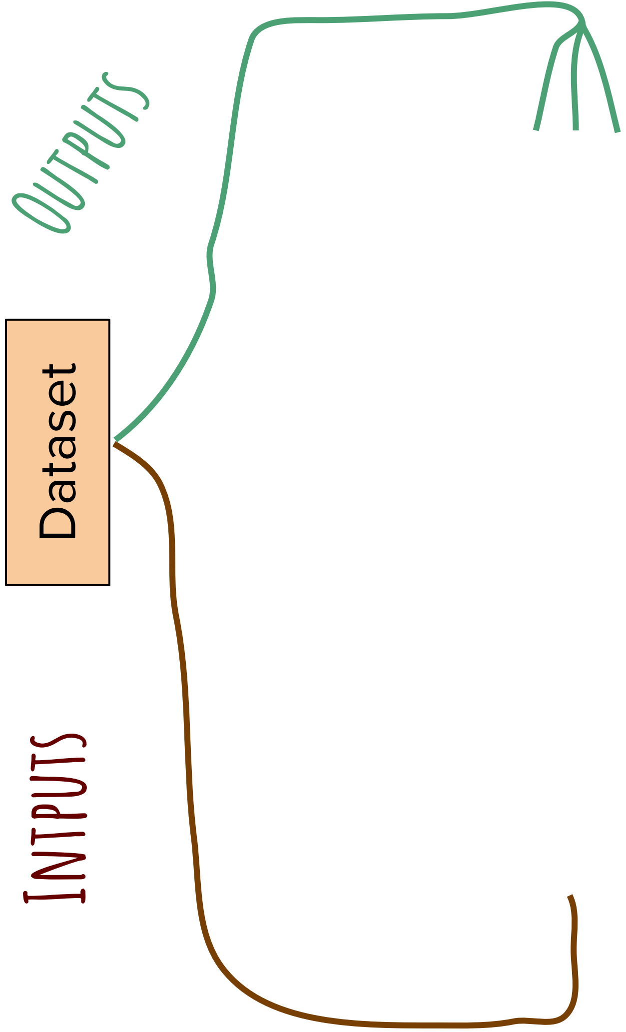

DARK MATTER HALO
Invisible
Deep Learning
STARS + GAS
Observable

Brief (1 slide) Introduction to Supervised Learning

Supervised Learning



Construction of the Dataset


OUTPUTS
Dataset
INTPUTS

TNG100 Cosmological Hydrodynamical Simulation
(https://www.tng-project.org/)

- Planck cosmology
- 106.5 Mpc by side
- 1820^3 DM particles
- 1820^3 hydrodynamic cells
- DM resolution 7.5 *10^6 M☉
- Baryon resolution 1.4*10^6 M☉
- 136 snapshots from z=127 to z=0

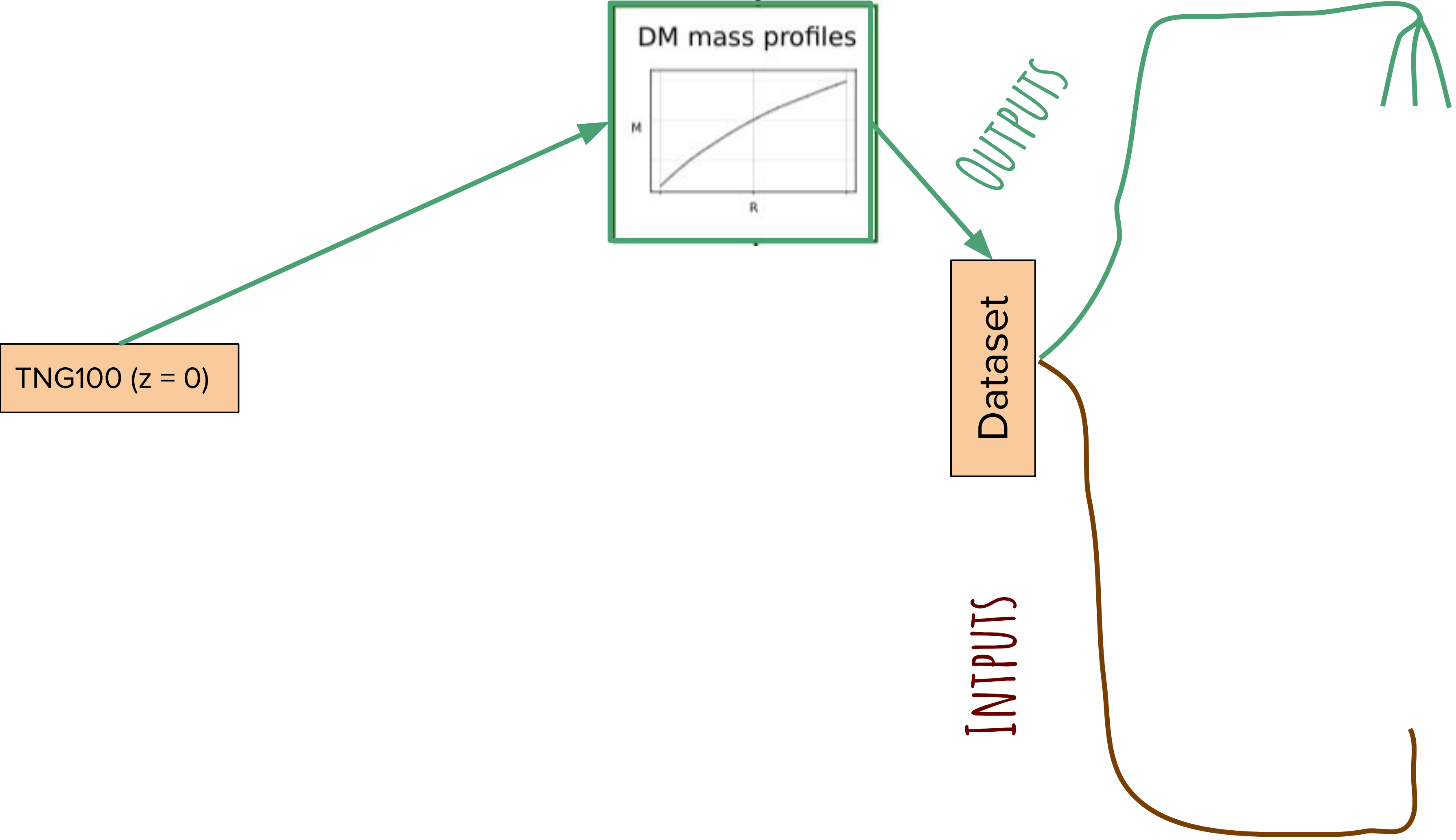
Property	Criterium
Simulation snapshot	99 $(z = 0)$
Stellar mass	$10^{10}\ M_\odot \leq M_\star \leq 10^{12}\ M_\odot$
Star formation rate	$\mathrm{SFR} \geq 0.1\ M_\odot/\mathrm{yr}$
Central galaxy	SubhaloParent $= 0$
Cosmological origin	SubhaloFlag $= 1$

TNG100 (z = 0)

OUTPUTS

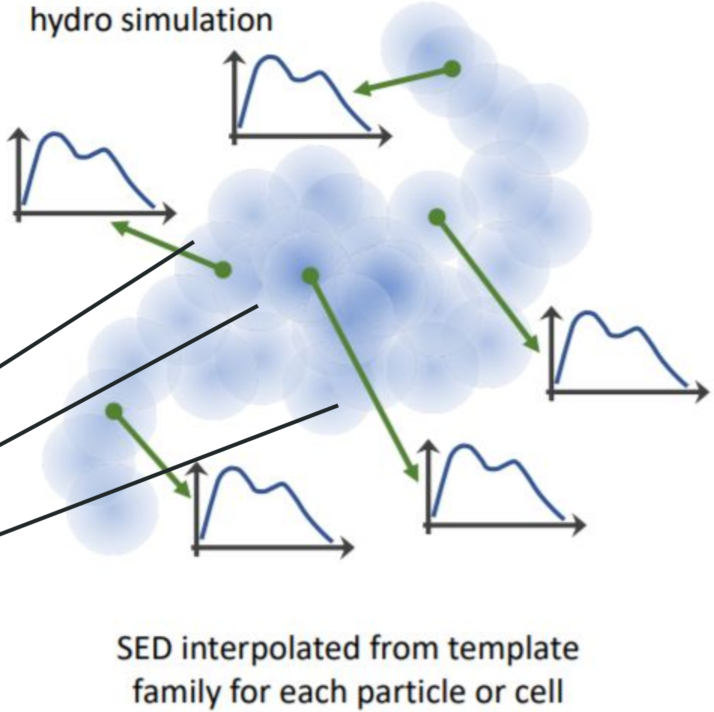
Dataset

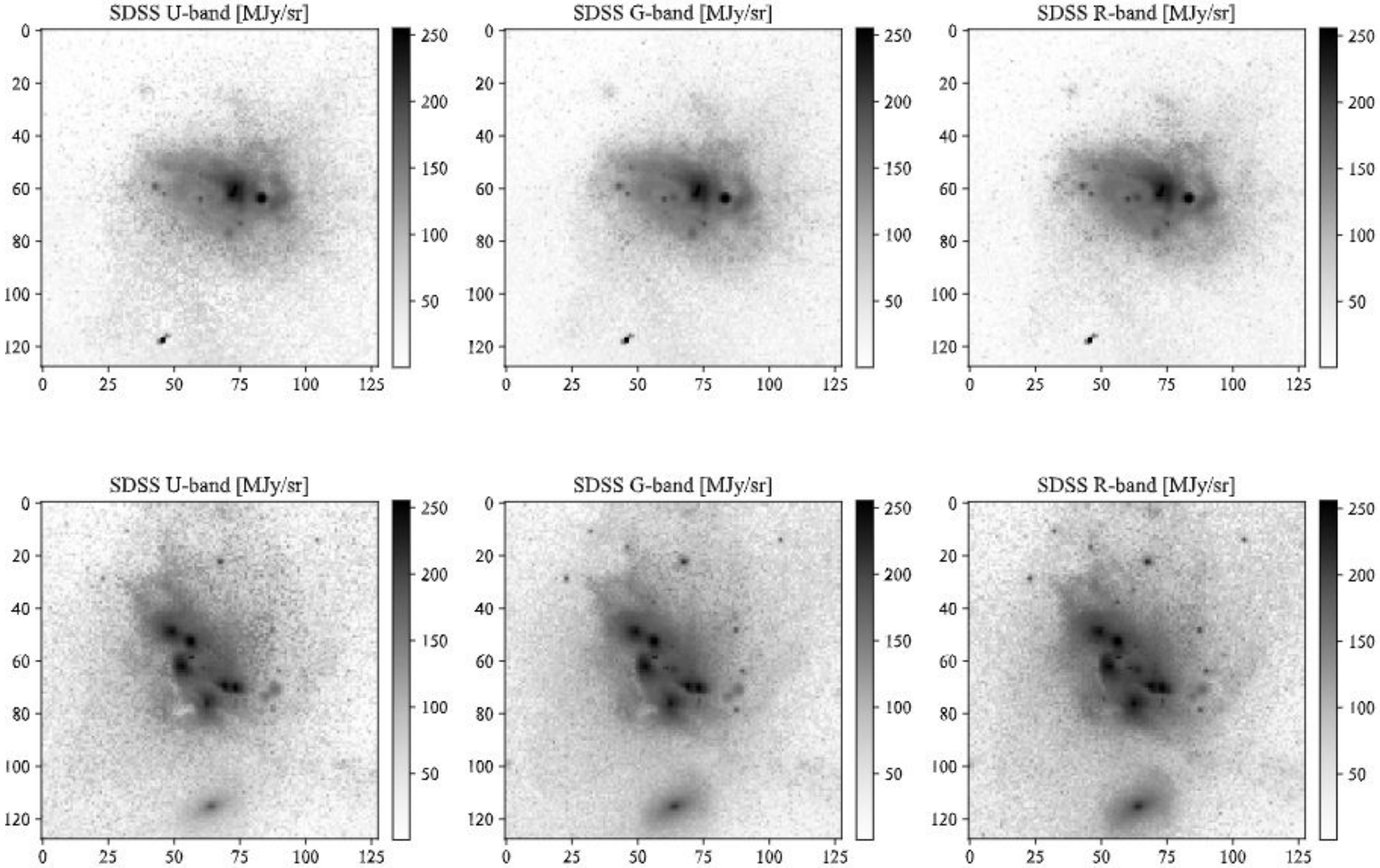
INTPUTS

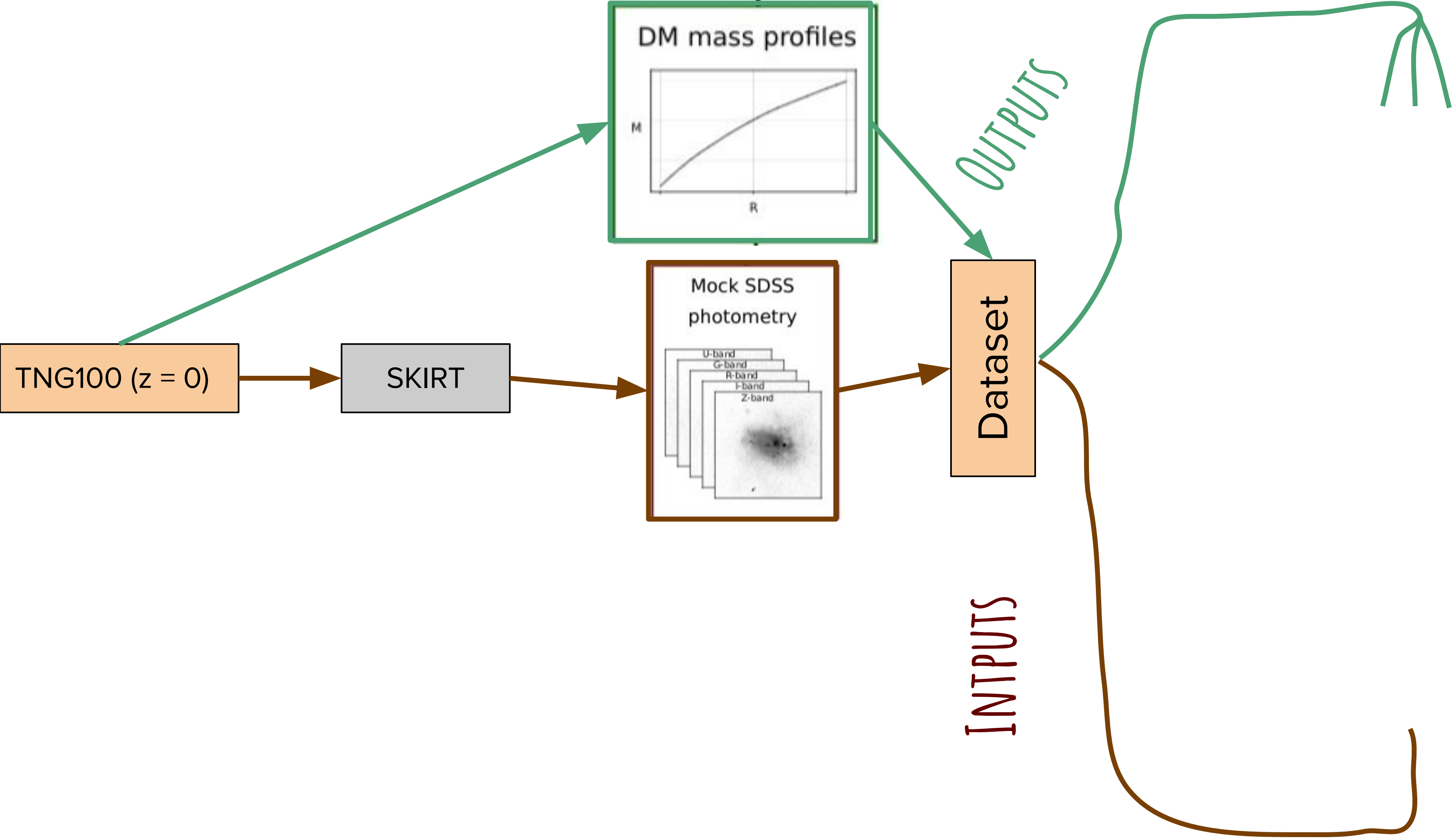

DM mass profiles
M
R
TNG100 (z = 0)
Dataset
OUTPUTS
INTPUTS

SKIRT* (2003.00721, skirt.ugent.be)

Radiative transfer code which emulates the stellar emissions and subsequent light-ray propagation to the observer, taking into account the absorption and re-emission by dust.



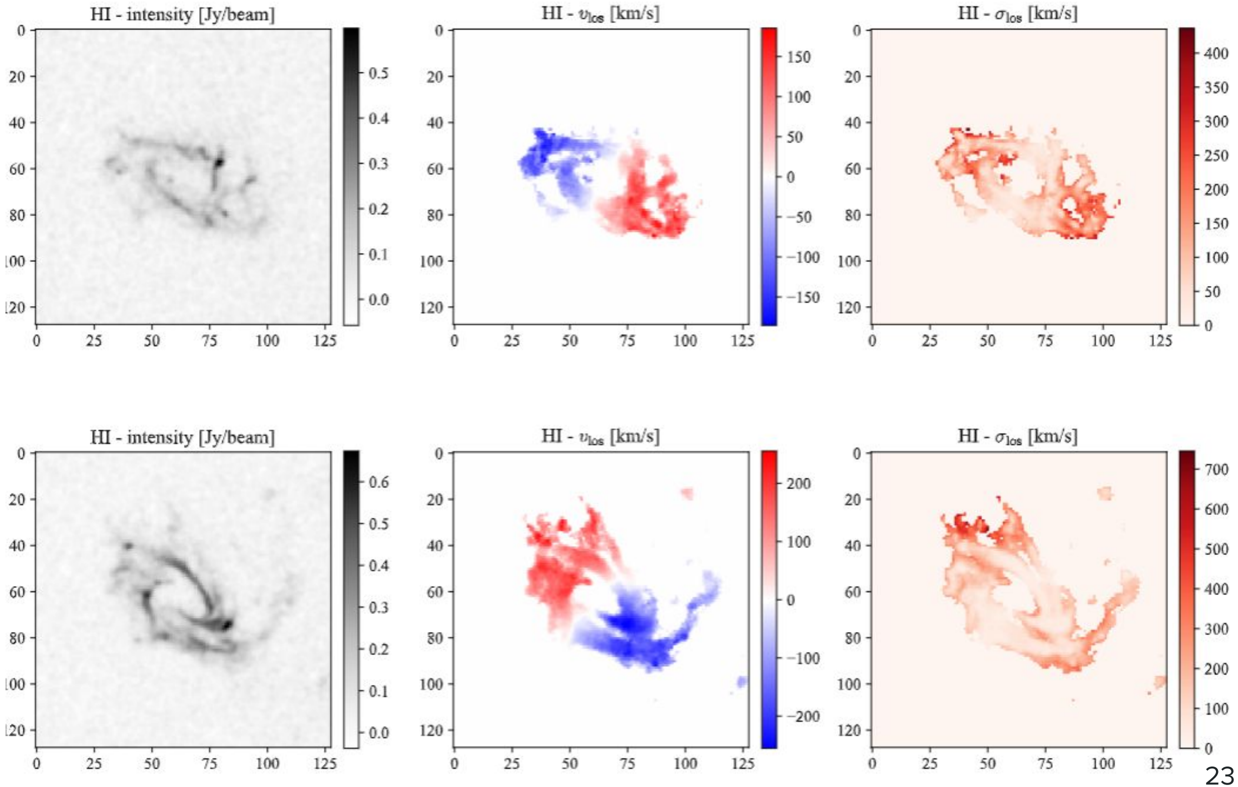

SDSS U-band [MJy/sr]
SDSS G-band [MJy/sr]
SDSS R-band [MJy/sr]
SDSS U-band [MJy/sr]
SDSS G-band [MJy/sr]
SDSS R-band [MJy/sr]
0 20 40 60 80 100 120
0 25 50 75 100 125
50 100 150 200 250

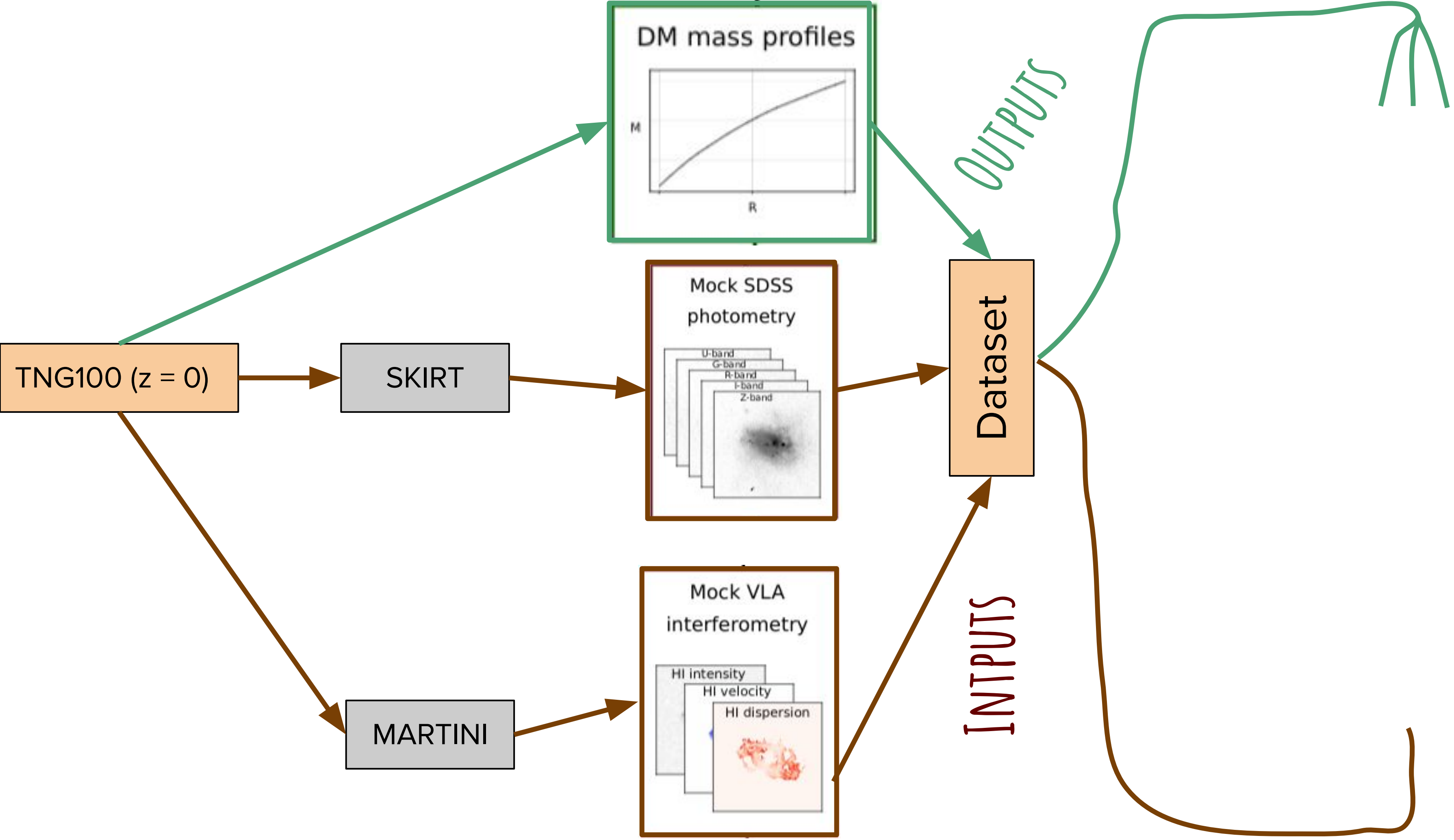


DM mass profiles
M
R
OUTPUTS
Mock SDSS
photometry
U-band
G-band
R-band
I-band
Z-band
TNG100 (z = 0)
SKIRT
Dataset
INTPUTS

MARTINI (1706.07478 ; https://github.com/kyleaoman/martini)

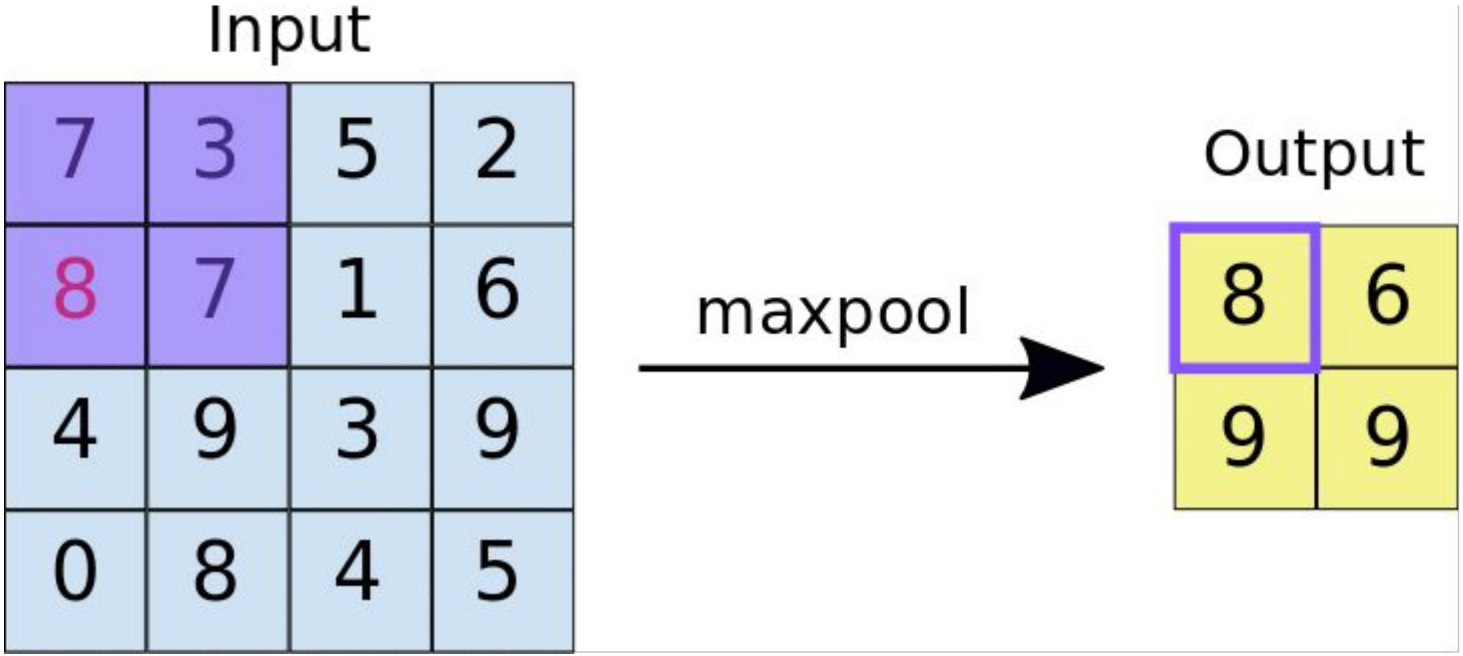
Allows for the creation of synthetic resolved HI line observations (i.e. data cubes) directly from the snapshot of a hydrodynamic simulation, and its posterior analysis.



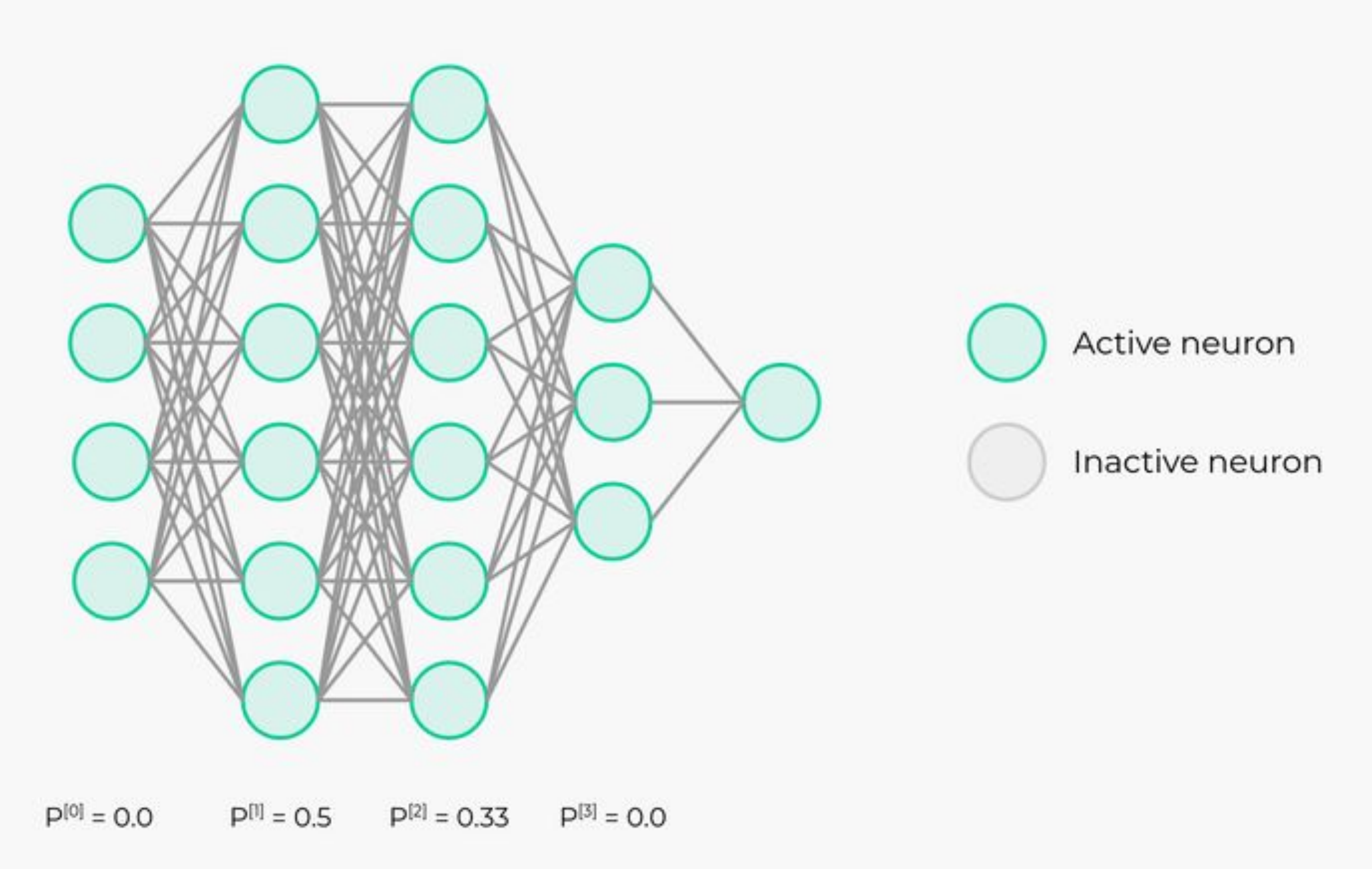

DM mass profiles
M
R
Mock SDSS
photometry
U-band
G-band
R-band
I-band
Z-band
Mock VLA
interferometry
HI intensity
HI velocity
HI dispersion
TNG100 (z = 0)
SKIRT
MARTINI
Dataset
OUTPUTS
INTPUTS

Convolutional layers

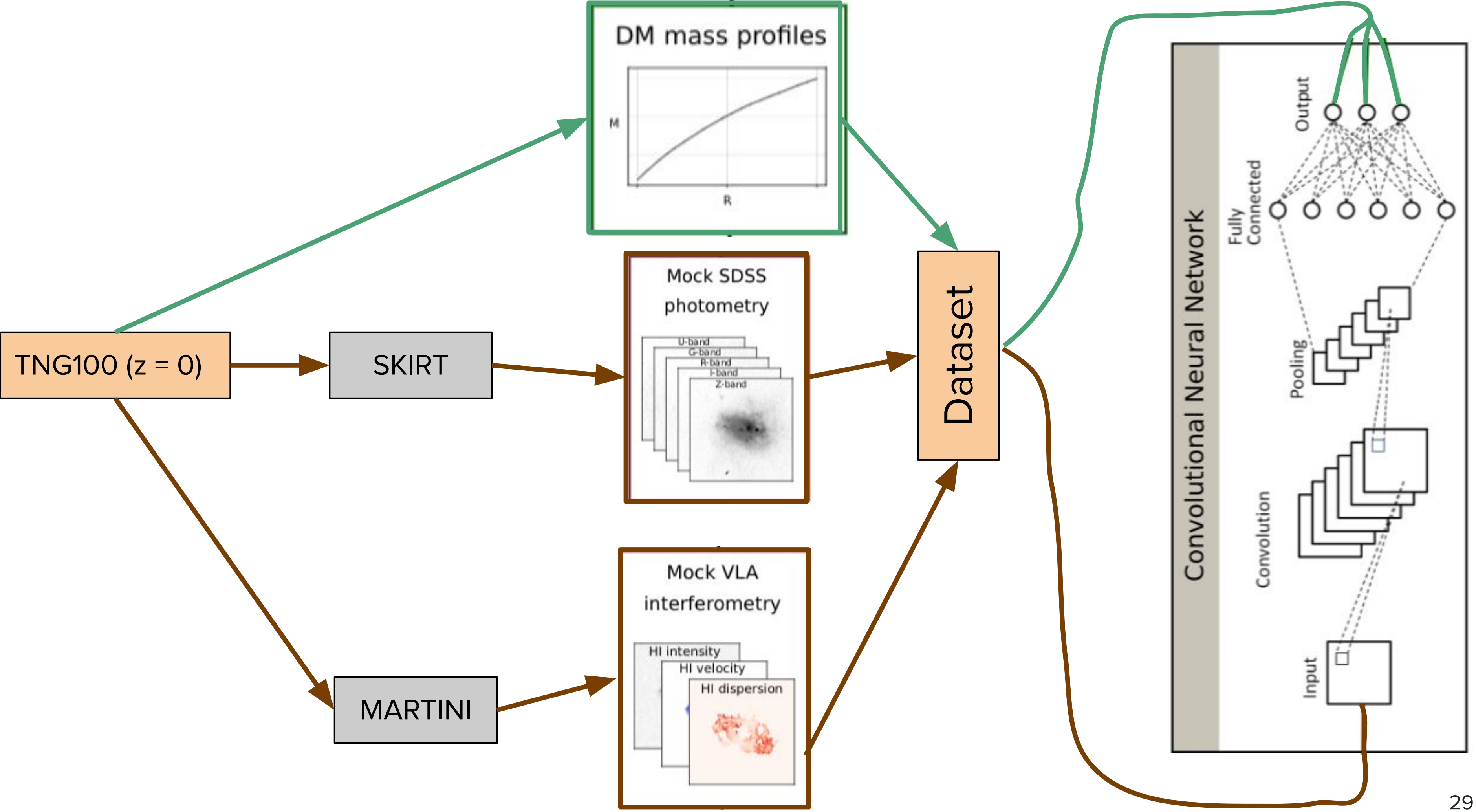
Pooling layers



Dropout



Layer	Details
2D convolution	64 kernels, 5×5 px kernel size, 2 px stride, ReLU activation
2D max pooling	2 px pooling
Dropout	50% dropout fraction
Batch normalization	
2D convolution	128 kernels, 5×5 px kernel size, 2 px stride, ReLU activation
2D max pooling	2 px pooling
Dropout	50% dropout fraction
Batch normalization	
2D convolution	256 kernels, 5×5 px kernel size, 2 px stride, ReLU activation
Batch normalization	
Dense	256 units, ReLU activation
Dropout	50% dropout fraction
Batch normalization	
Dense	128 units, ReLU activation
Dropout	50% dropout fraction
Batch normalization	
Dense	64 units, ReLU activation
Dropout	50% dropout fraction
Batch normalization	
Dense (output)	20 units, linear activation


DM mass profiles
M
R
TNG100 (z = 0)
SKIRT
MARTINI
Mock SDSS
photometry
U-band
G-band
R-band
I-band
Z-band
Mock VLA
interferometry
HI intensity
HI velocity
HI dispersion
Dataset
Convolutional Neural Network
Input
Convolution
Pooling
Fully
Connected
Output

Results

Comparison between different architectures

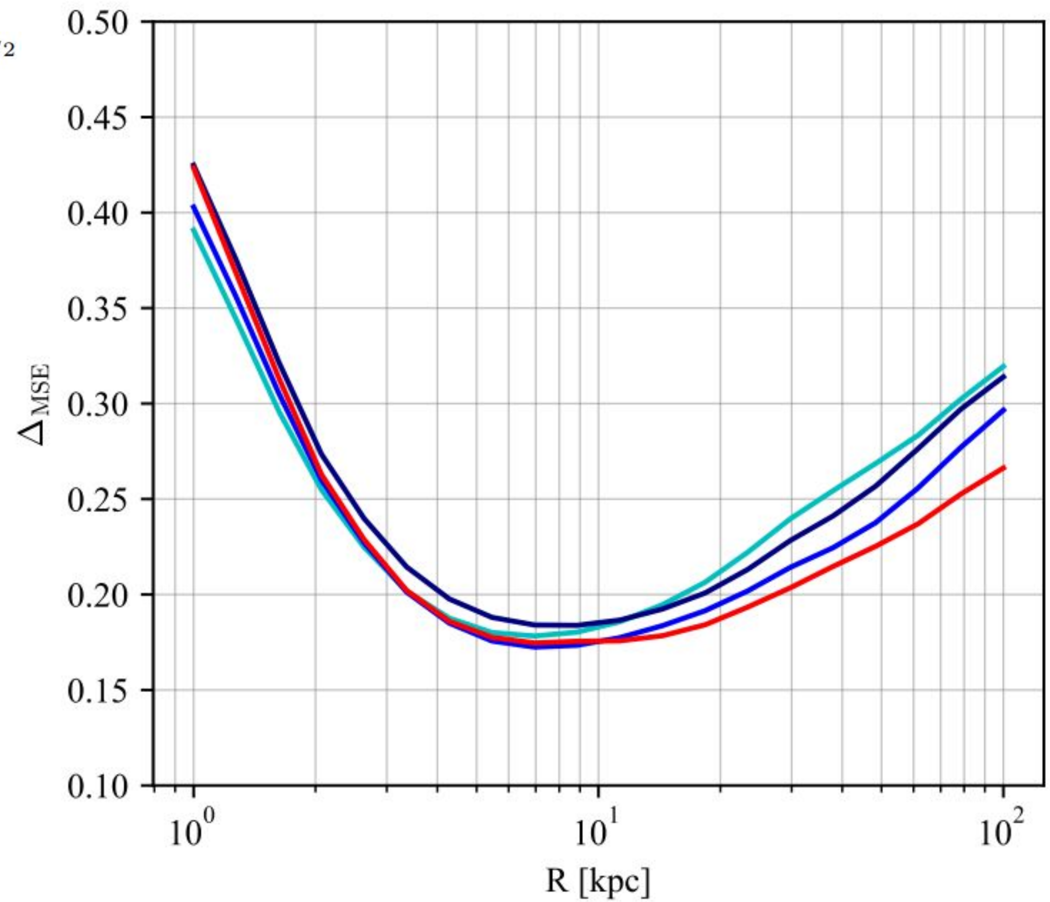
$$\Delta_{\mathrm{MSE}}(R_i) = \left[\frac{1}{N}\sum_{j=1}^{N}\left(\frac{\mu_j(R_i) - \hat{\mu}_j(R_i)}{\hat{\mu}_j(R_i)}\right)^2\right]^{1/2}$$

Architecture A

Architecture B

Architecture C

ResNet50



Comparison between different inputs

SDSS I

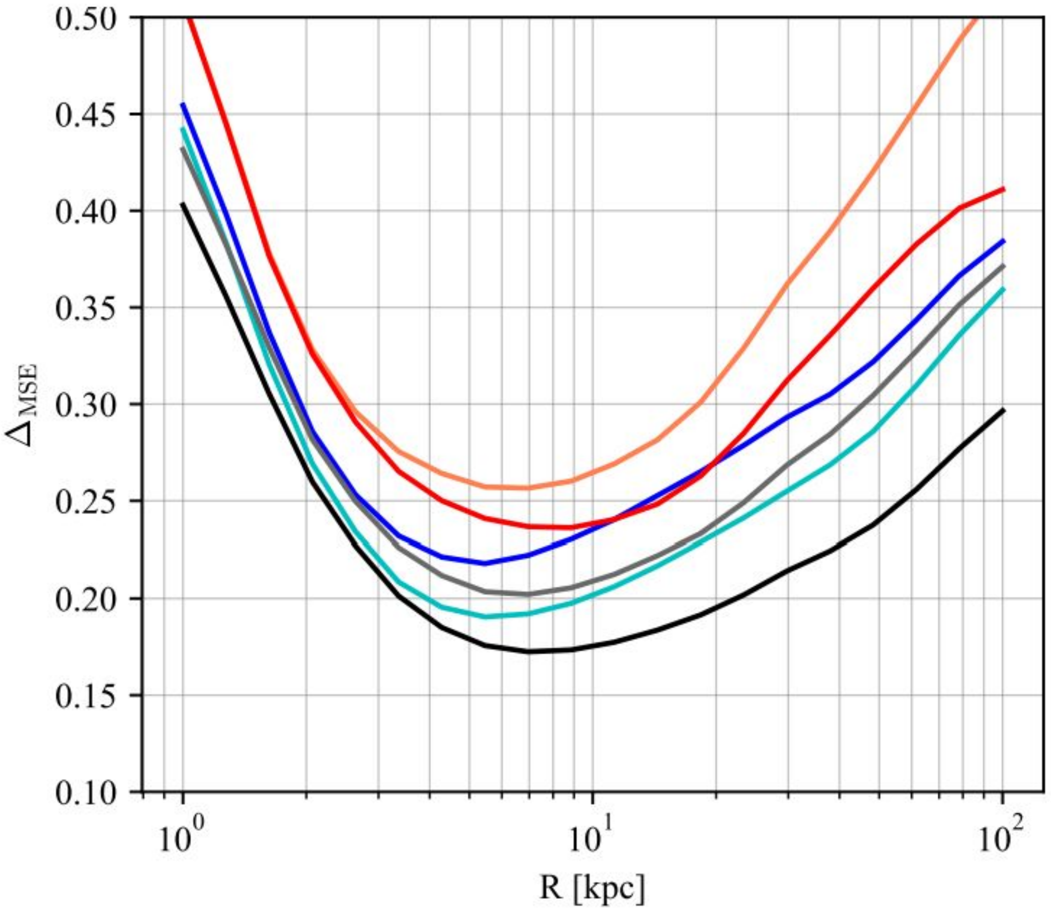
SDSS URZ

V_{HI}

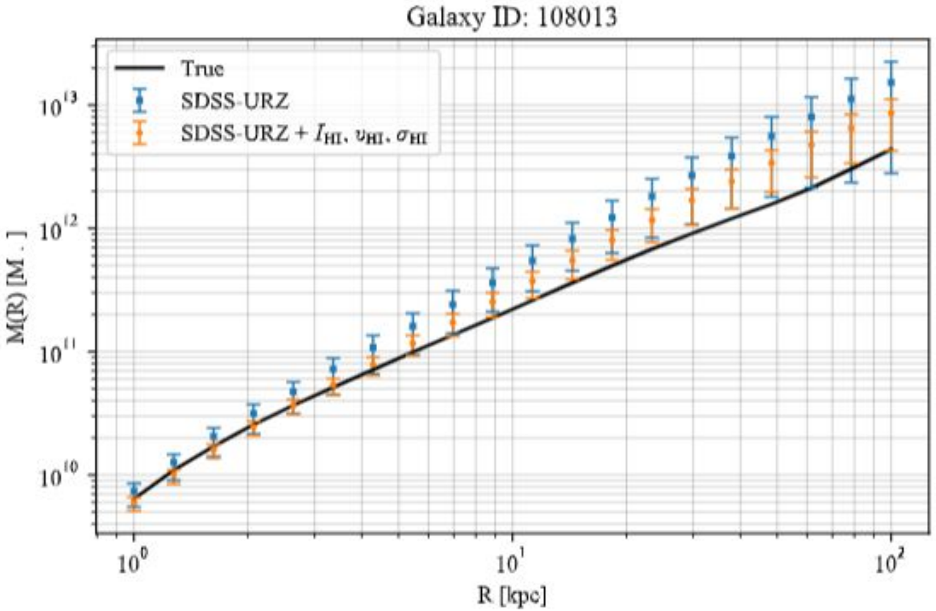
$I_{HI} + V_{HI} + \sigma_{HI}$

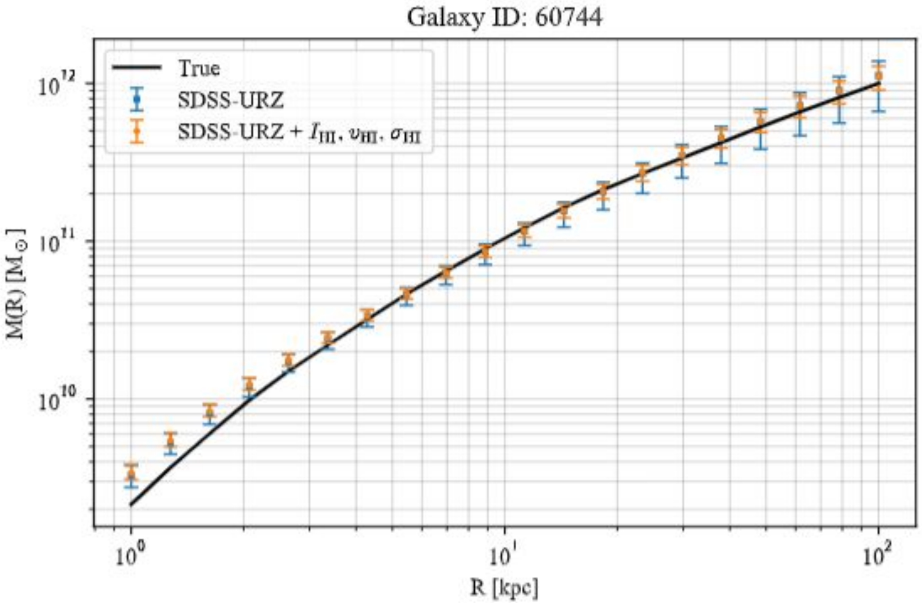
SDSS URZ + V_{HI}

SDSS URZ + $I_{HI} + V_{HI} + \sigma_{HI}$



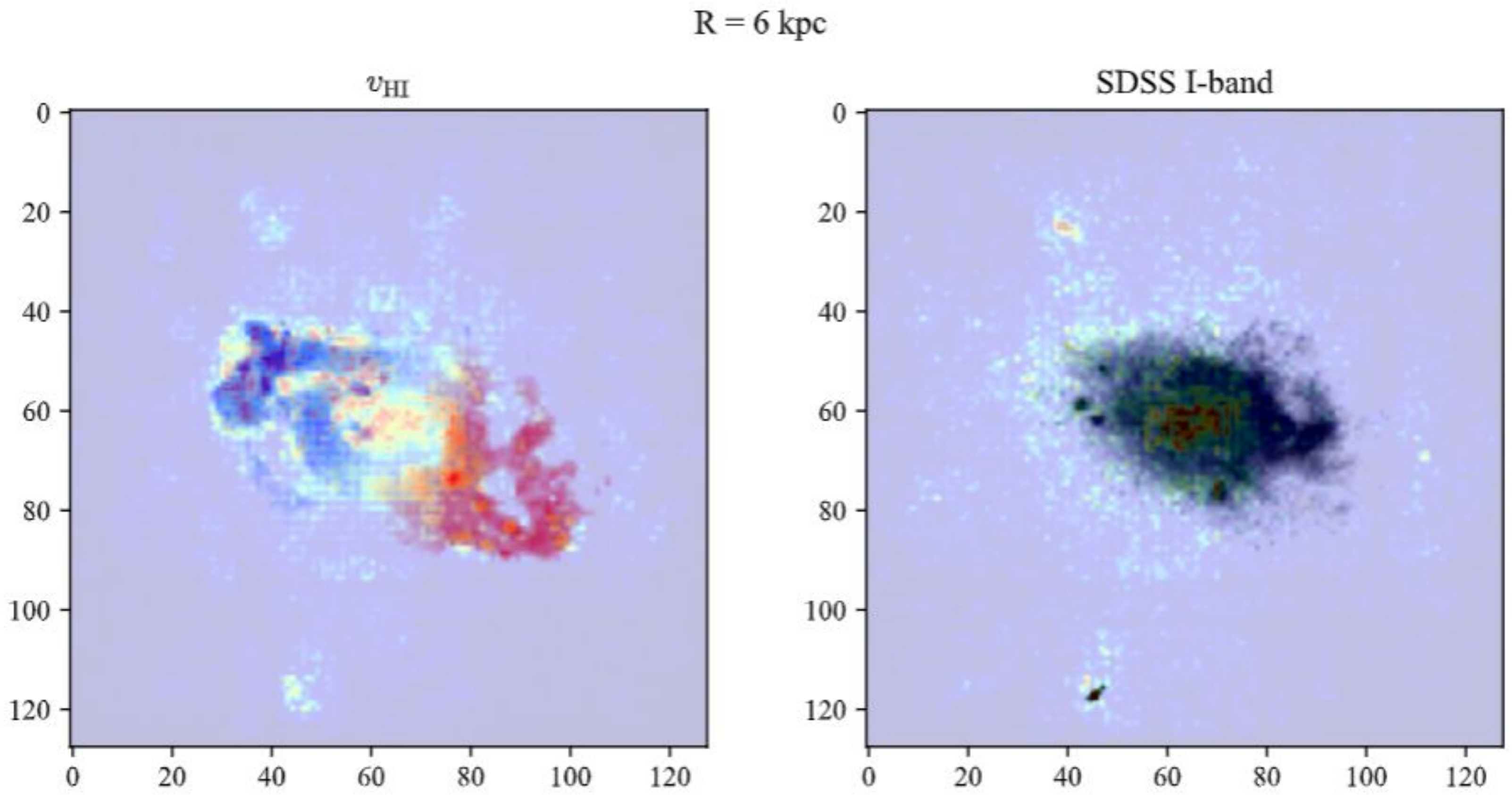
Prediction of the dark matter profile



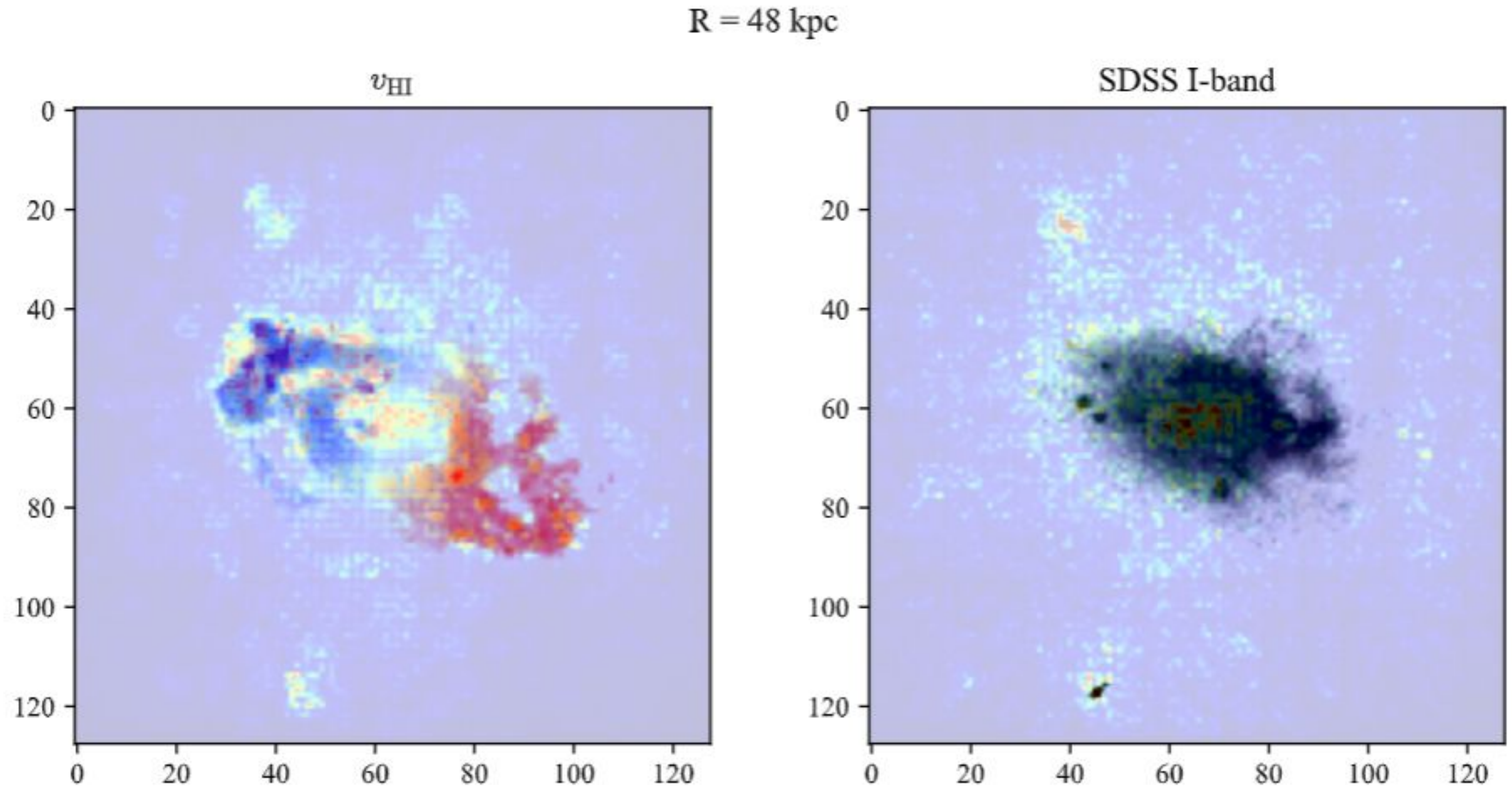


Understanding the results

$$S_{ij} \equiv \frac{\partial y}{\partial x_{ij}}$$

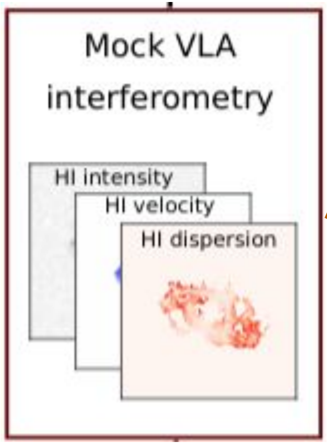


Understanding the results



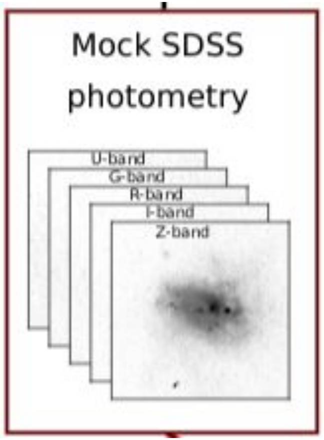
Comparison with RC analysis

(preliminary)



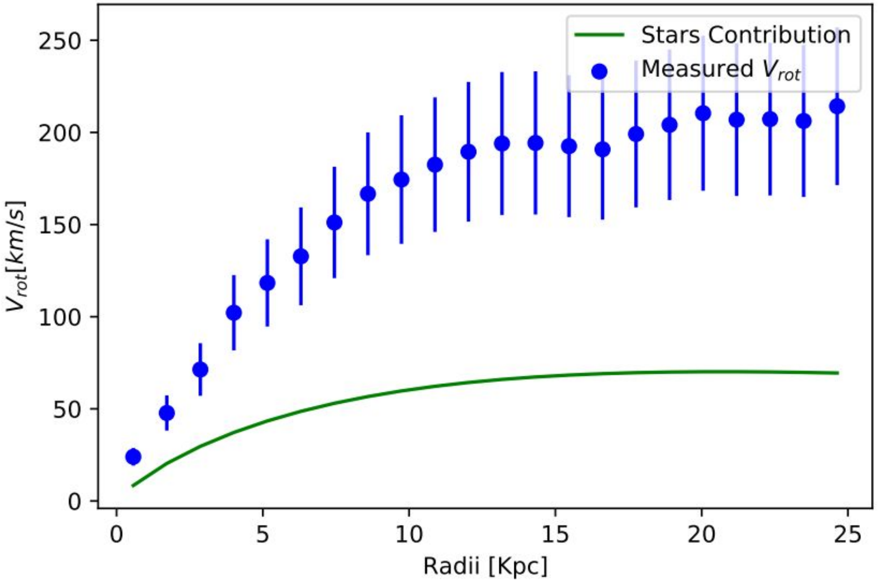
^{3D}BAROLO

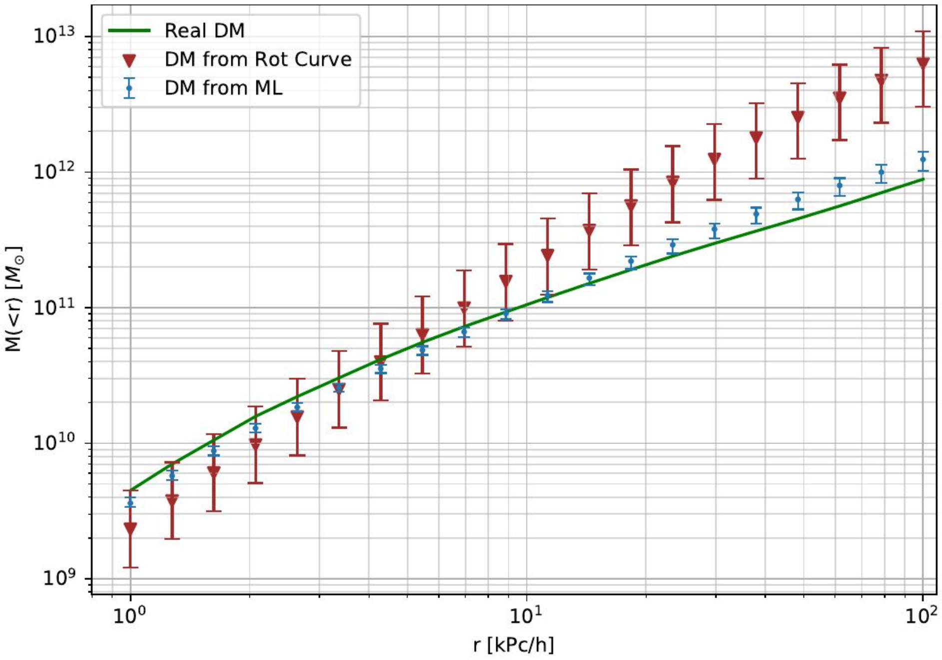
https://arxiv.org/abs/1505.07834

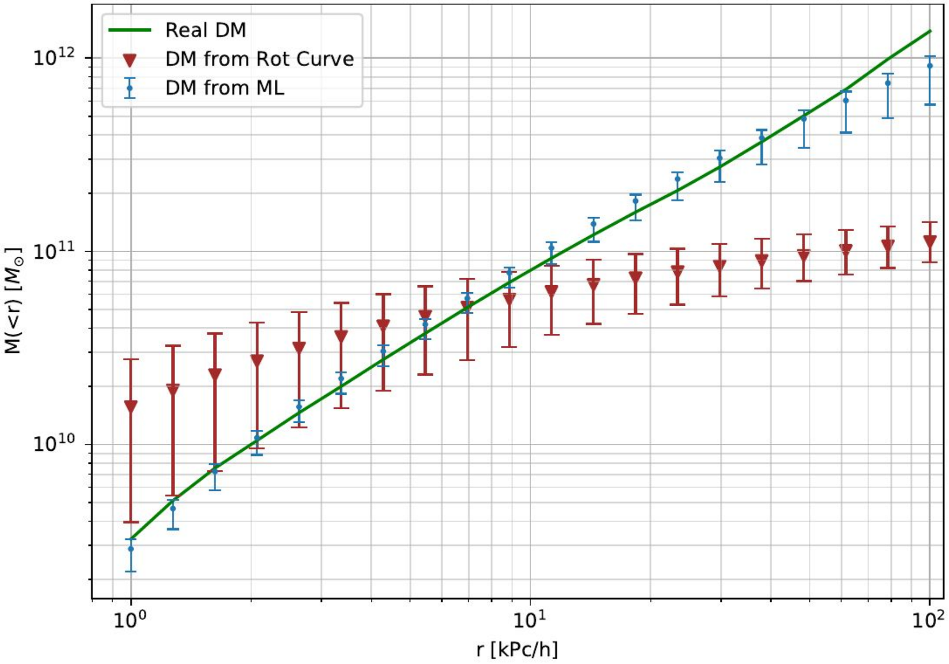


Autoprof

https://arxiv.org/abs/2106.13809




Real DM
DM from Rot Curve
DM from ML
10^{13}
10^{12}
10^{11}
10^{10}
10^{9}
$M(<r)\ [M_\odot]$
10^{0}
10^{1}
10^{2}
r [kPc/h]


Real DM
DM from Rot Curve
DM from ML
10^{12}
10^{11}
10^{10}
$M(<r)\ [M_\odot]$
10^{0}
10^{1}
10^{2}
r [kPc/h]

Conclusion and Future Work

- Our algorithm is able to reconstruct the DM distribution profile with high performance throughout the extension of the galaxy.

- The highest performance is achieved in the intermediate regions with a mean square error below 0.2 using all the photometric and spectroscopic information.

- Even in the absence of spectroscopic information, our method is able to recover the dark matter profile with a mean square error below 0.3 in the intermediate regions.

- Our reconstruction of the DM distribution is completely data-driven, and does not need any assumption on the shape nor the functional form of the DM profile.

- The method developed here is applicable to different types of galaxies since it does not rely on explicit physical assumptions regarding the dynamical state of the system.

- The results achieved have been obtained for galaxies with masses in the range ~10^{10}-10^{12} $M_\odot$ but the methodology can be extended to a broader mass range.

- We will make a comparison with the dark matter profile obtained through the traditional rotation curve analysis for the simulated galaxies.
- Study the robustness of our results to the hydrodynamical cosmological simulation.
- Apply our method to real galaxies and compare the results with other estimations.

THANK YOU