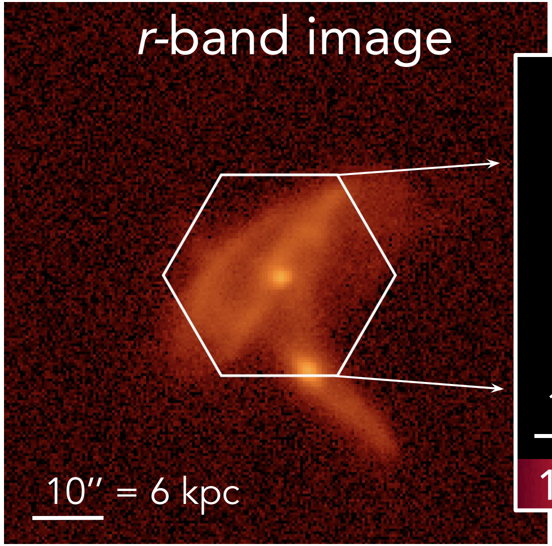
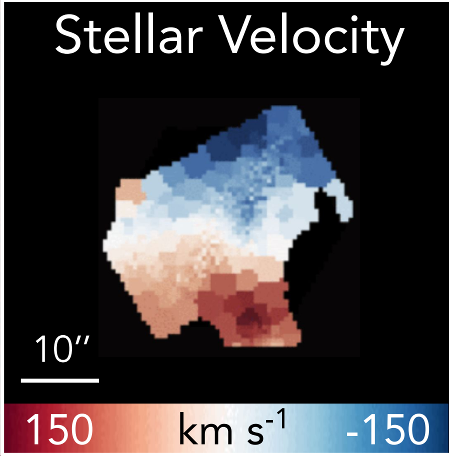
How do we best identify merging galaxies?:
Expanding the toolkit to include stellar kinematics and *HTST* NIRCam imaging

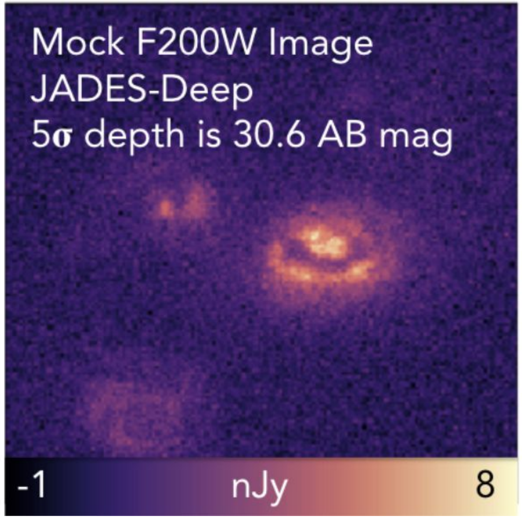
SDSS-ized



MaNGA-ized



HTST-ized



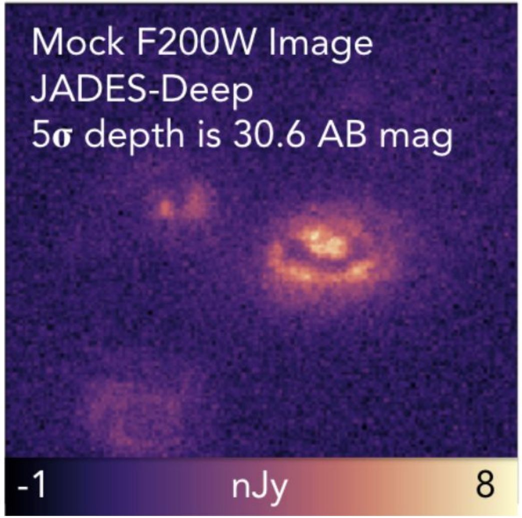
Becky Nevin | beckynevin.github.io

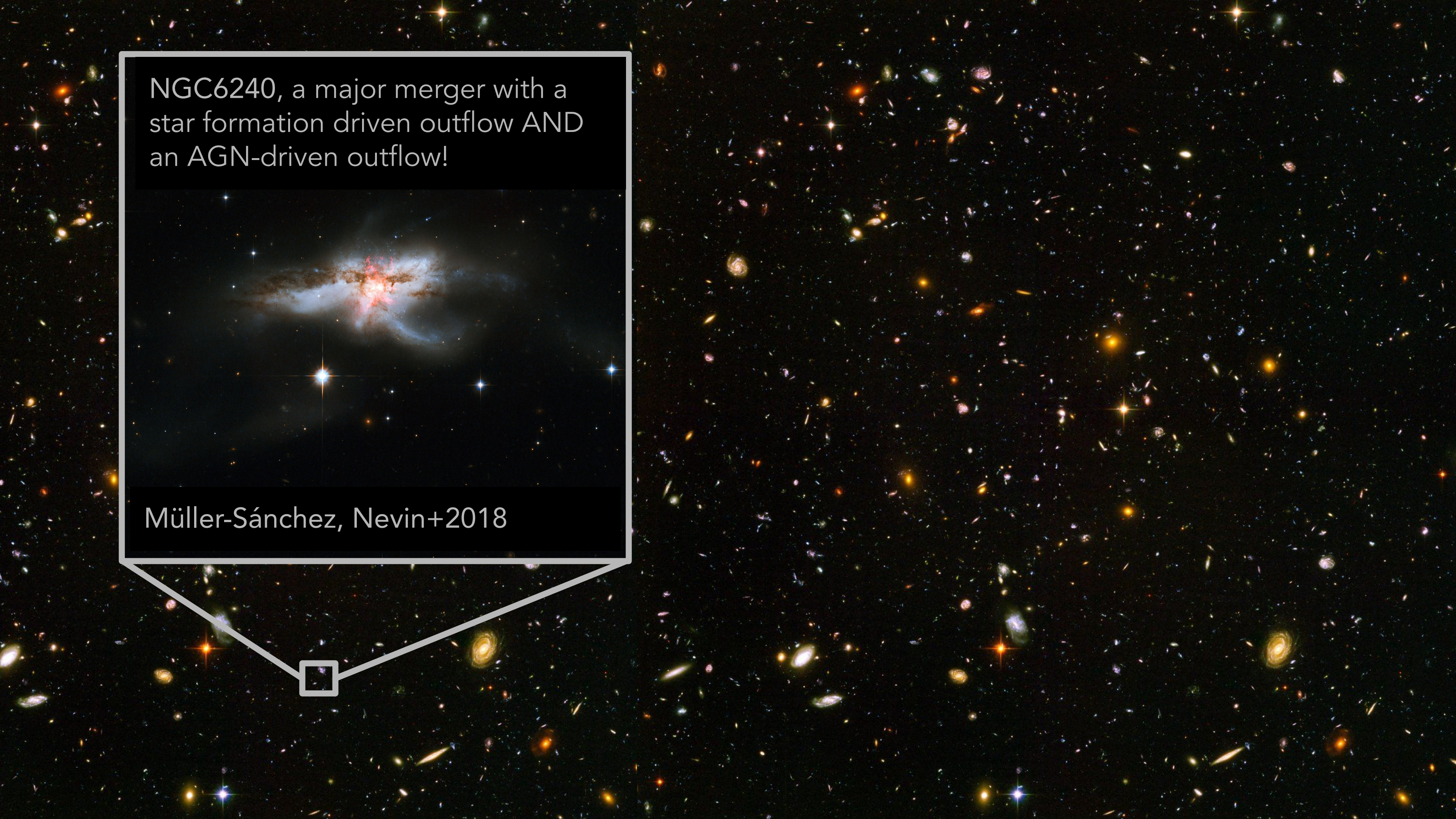
CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Harriet Tubman Space Telescope as a much better name for *JWST*: https://www.scientificamerican.com/article/nasa-needs-to-rename-the-james-webb-space-telescope/

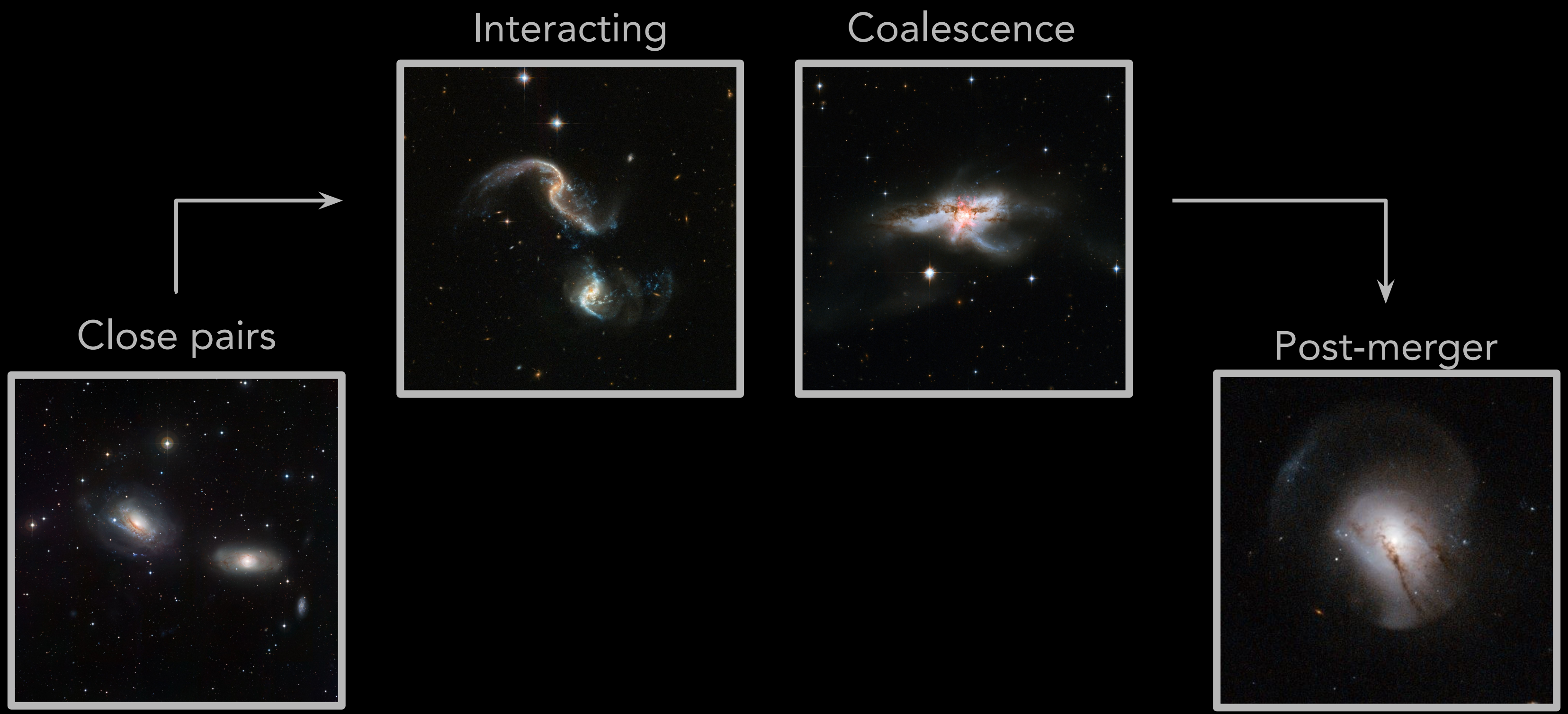
Petition to rename: http://bit.ly/RenameJWST

HTST-ized





NGC6240, a major merger with a star formation driven outflow AND an AGN-driven outflow!
Müller-Sánchez, Nevin+2018



*Huge caveat: This is one example of a possible evolutionary sequence. Not all galaxies go through all of these steps in this order, and by the way, this is a gas rich major merger.

Interacting

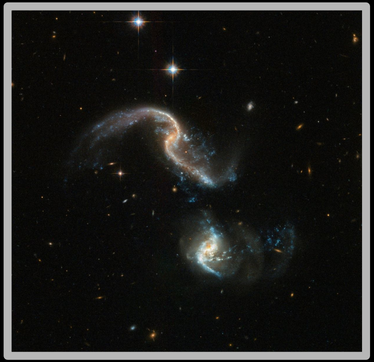
Coalescence

Close pairs

Post-merger

Close pairs

Interacting



Coalescence

Post-merger

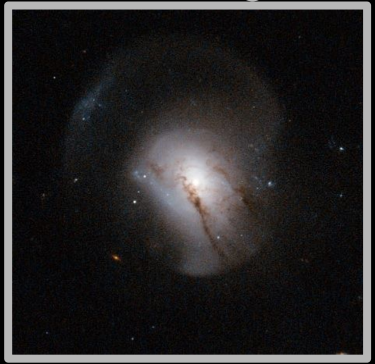
Close pairs

Interacting

Coalescence

Post-starburst

Post-merger



See cool work by Decker
French
Kate Rowlands

How do we identify a diversity of galaxy mergers?

Interacting

Coalescence

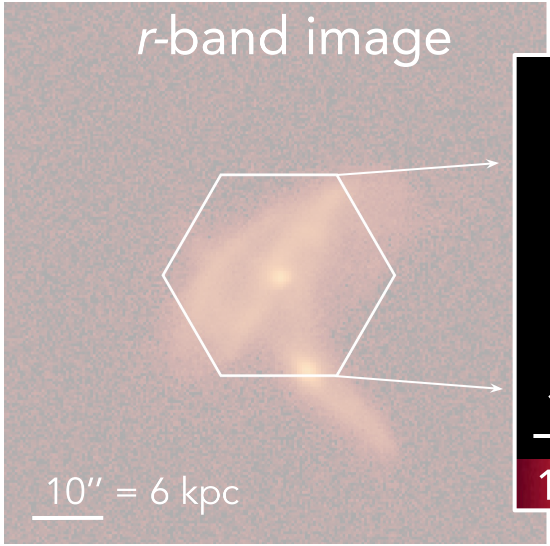
Close pairs

Post-merger

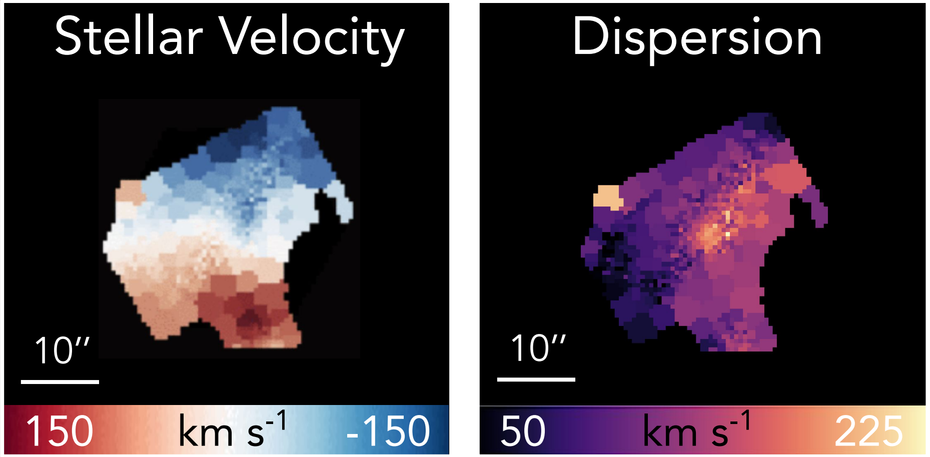
Different stages

Minor mergers

How do we best identify merging galaxies?: Expanding the toolkit to include stellar kinematics



Nevin+2019



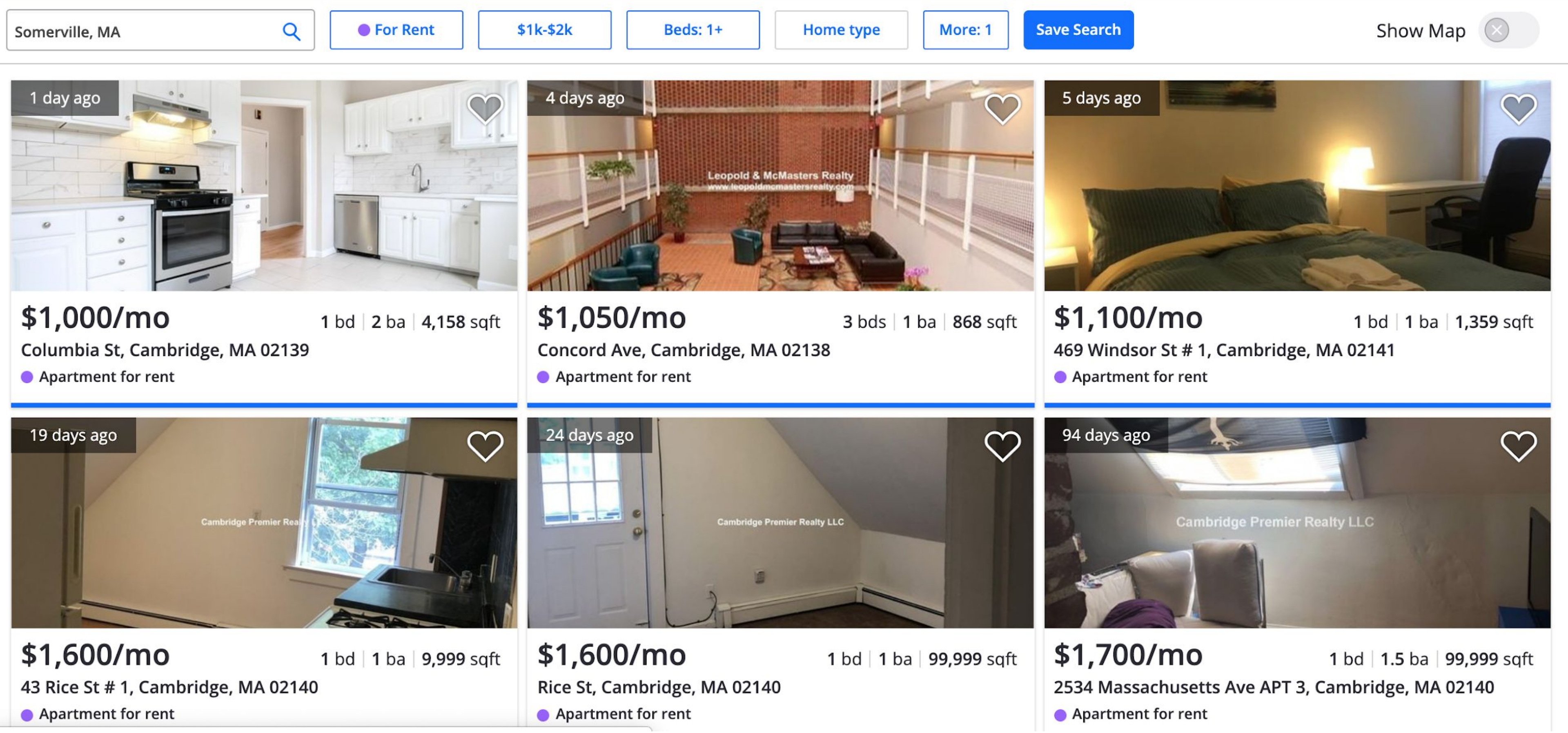
Nevin+2021;
https://arxiv.org/abs/2102.02208

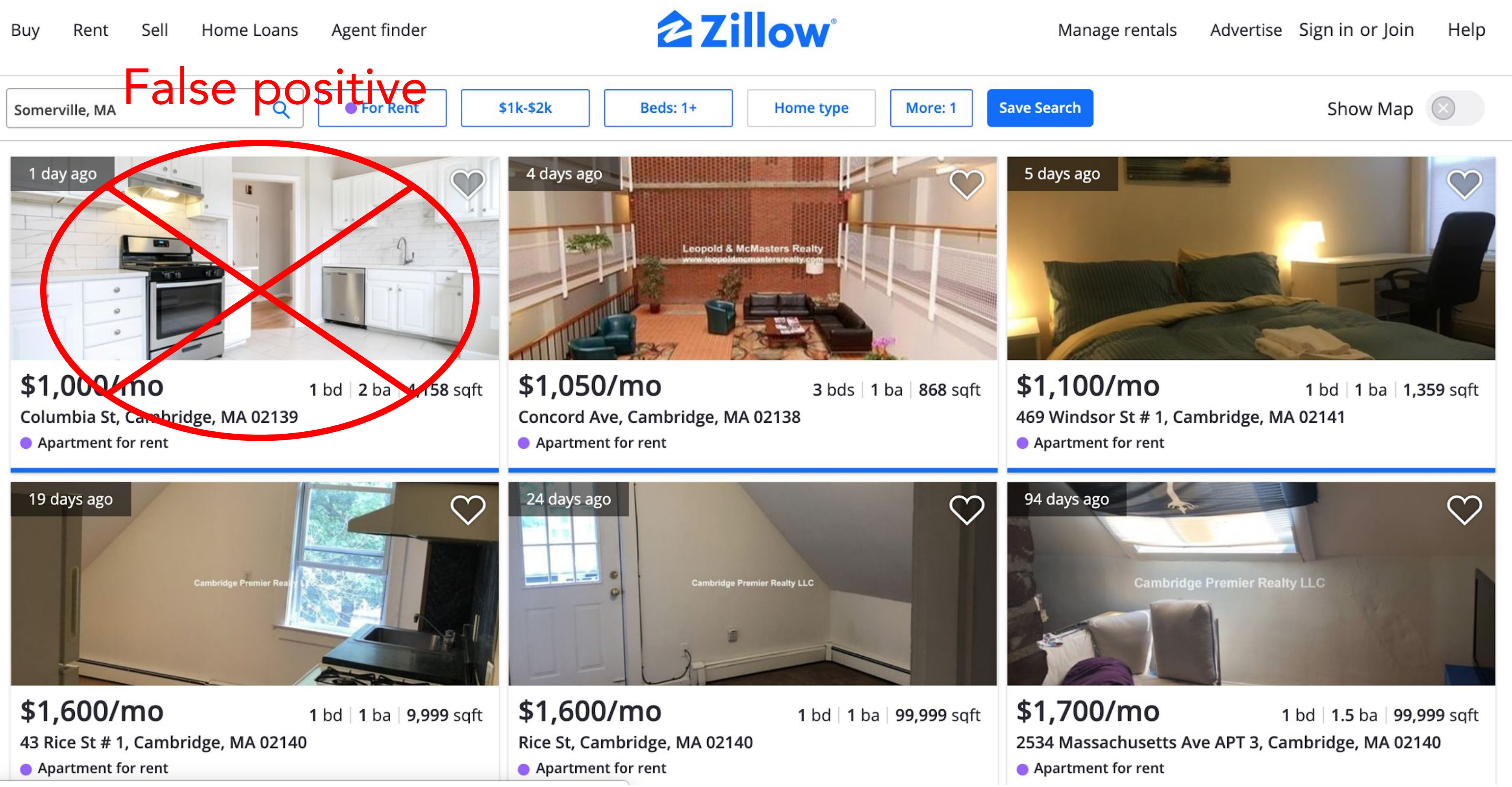
How do we identify a diversity of galaxy mergers?

How do we identify a diversity of galaxy mergers?

Search engine matters

Zillow   Manage rentals   Advertise   Sign in or Join   Help




Buy
Rent
Sell
Home Loans
Agent finder
Zillow
Manage rentals
Advertise
Sign in or Join
Help
False positive
Somerville, MA
For Rent
$1k-$2k
Beds: 1+
Home type
More: 1
Save Search
Show Map
1 day ago
$1,000/mo 1 bd | 2 ba | 1,158 sqft
Columbia St, Cambridge, MA 02139
Apartment for rent
4 days ago
Leopold & McMasters Realty
www.leopoldmcmastersrealty.com
$1,050/mo 3 bds | 1 ba | 868 sqft
Concord Ave, Cambridge, MA 02138
Apartment for rent
5 days ago
$1,100/mo 1 bd | 1 ba | 1,359 sqft
469 Windsor St # 1, Cambridge, MA 02141
Apartment for rent
19 days ago
$1,600/mo 1 bd | 1 ba | 9,999 sqft
43 Rice St # 1, Cambridge, MA 02140
Apartment for rent
24 days ago
Cambridge Premier Realty LLC
$1,600/mo 1 bd | 1 ba | 99,999 sqft
Rice St, Cambridge, MA 02140
Apartment for rent
94 days ago
Cambridge Premier Realty LLC
$1,700/mo 1 bd | 1.5 ba | 99,999 sqft
2534 Massachusetts Ave APT 3, Cambridge, MA 02140
Apartment for rent

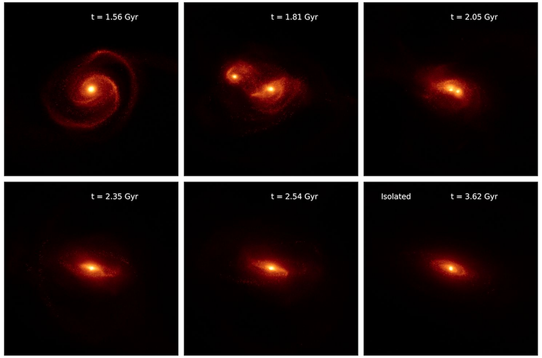
What can we learn from apartment hunting?

- The tool matters
- Combining tools can be great
- Intuition is helpful

How do we identify a greater diversity of galaxy mergers?

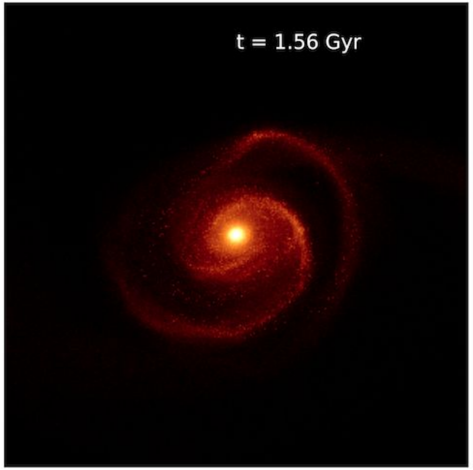
Simulations by
Laura Blecha :)

My solution is use simulations of galaxy mergers to create a merger identification tool



A suite of (five) N-body/SPH simulations with radiative transfer

Seven different viewing angles →



Advantage: These are high spatial resolution simulations at a high time cadence

Disadvantage: These are not cosmological simulations, these are disk-dominated intermediate mass galaxies

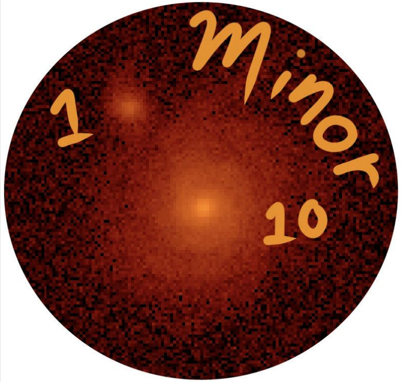
Major merger combined

1:2

1:3

1:3

Minor merger combined

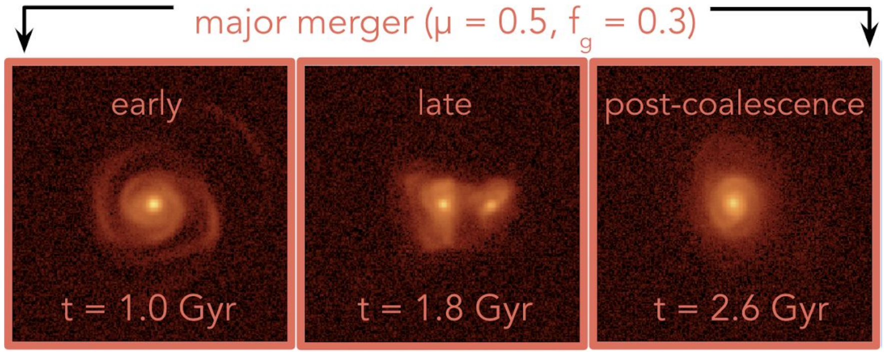


1:5

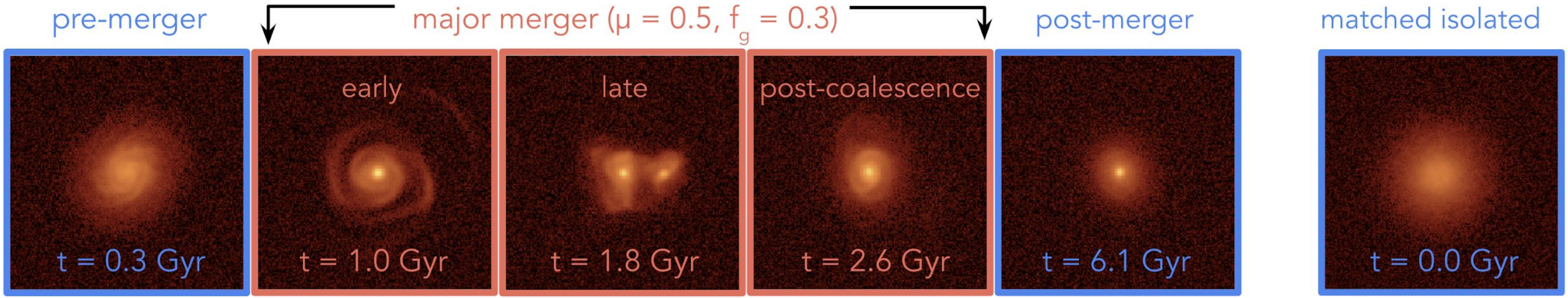
1:10

Spoiler alert! Mass ratio is the most impactful merger parameter.

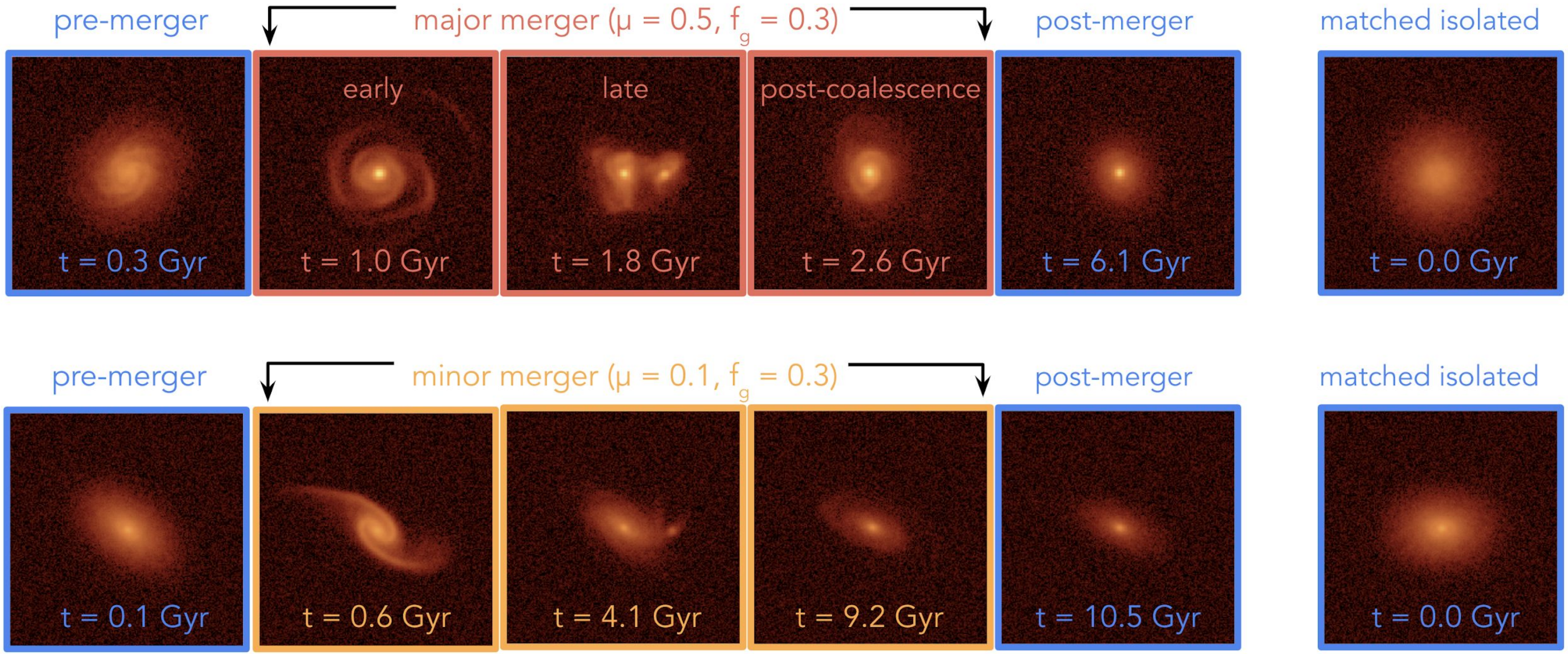
There are merging and nonmerging snapshots



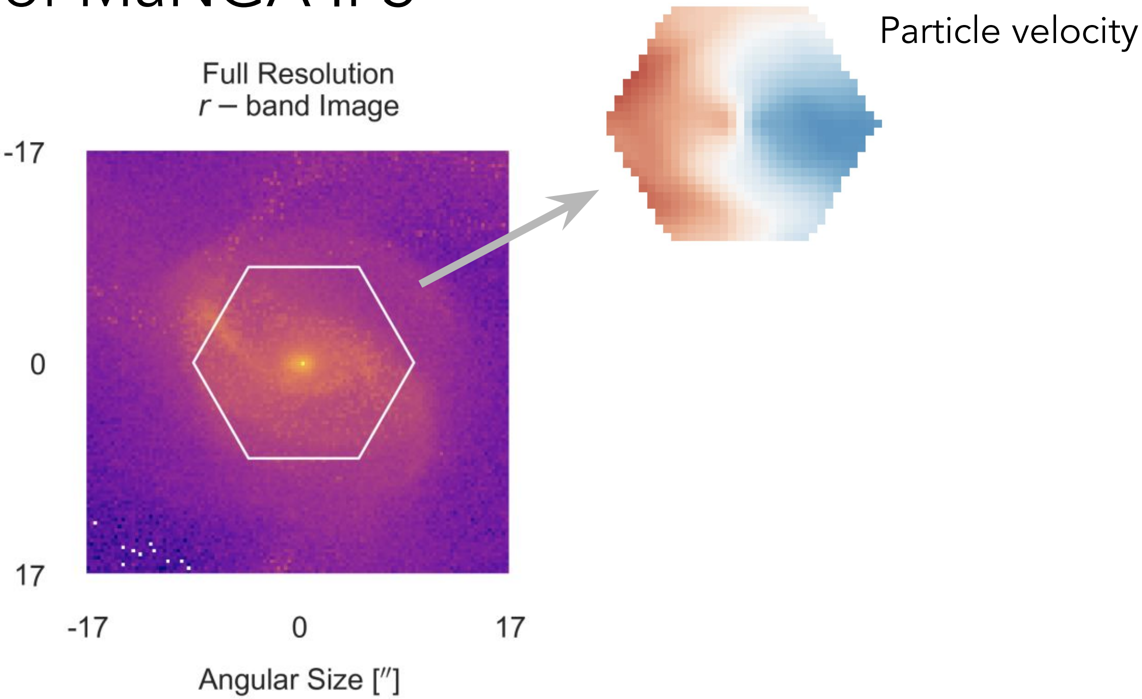
There are merging and nonmerging snapshots



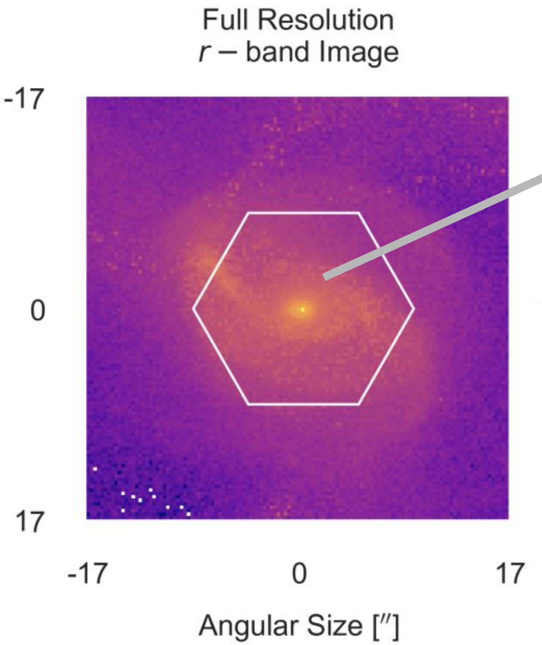
There are merging and nonmerging snapshots

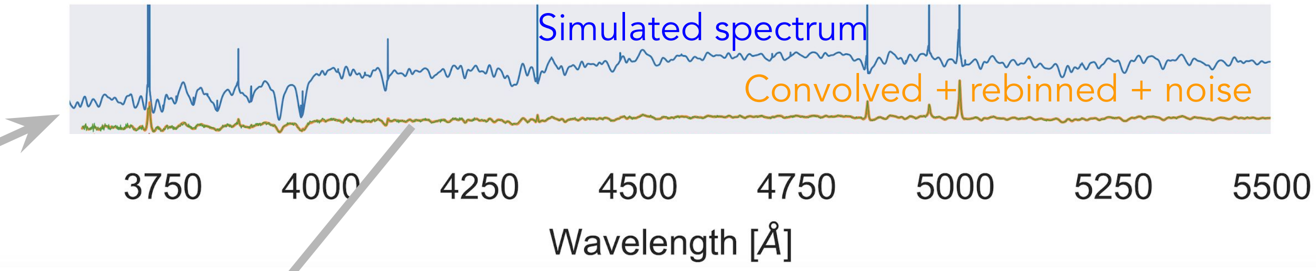


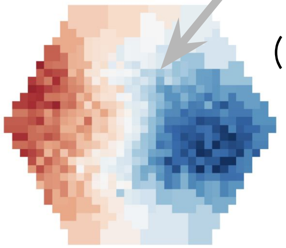
I create mock stellar kinematic maps to match the specifications of MaNGA IFS



I create mock stellar kinematic maps to match the specifications of MaNGA IFS



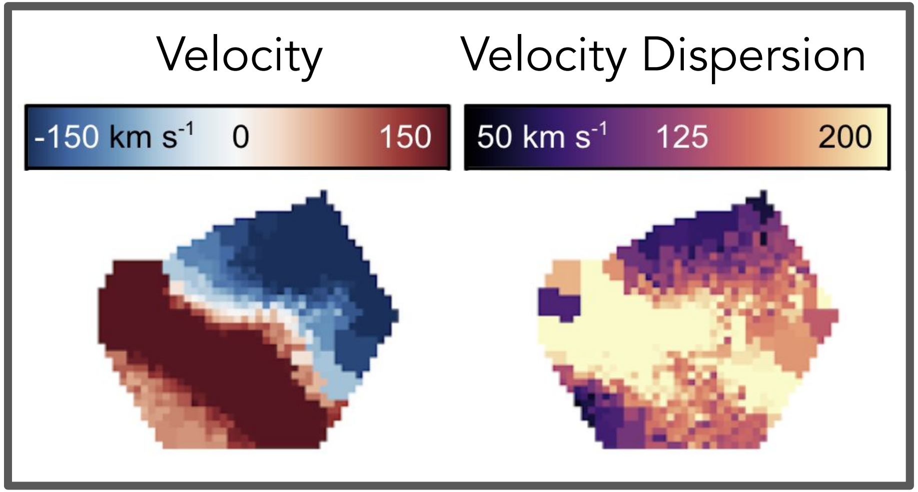




(Mock) Observed Stellar velocity

1. Radiative transfer
2. Convolve and rebin
3. Introduce characteristic noise
4. Mask and voronoi bin
5. ppxf fit absorption lines

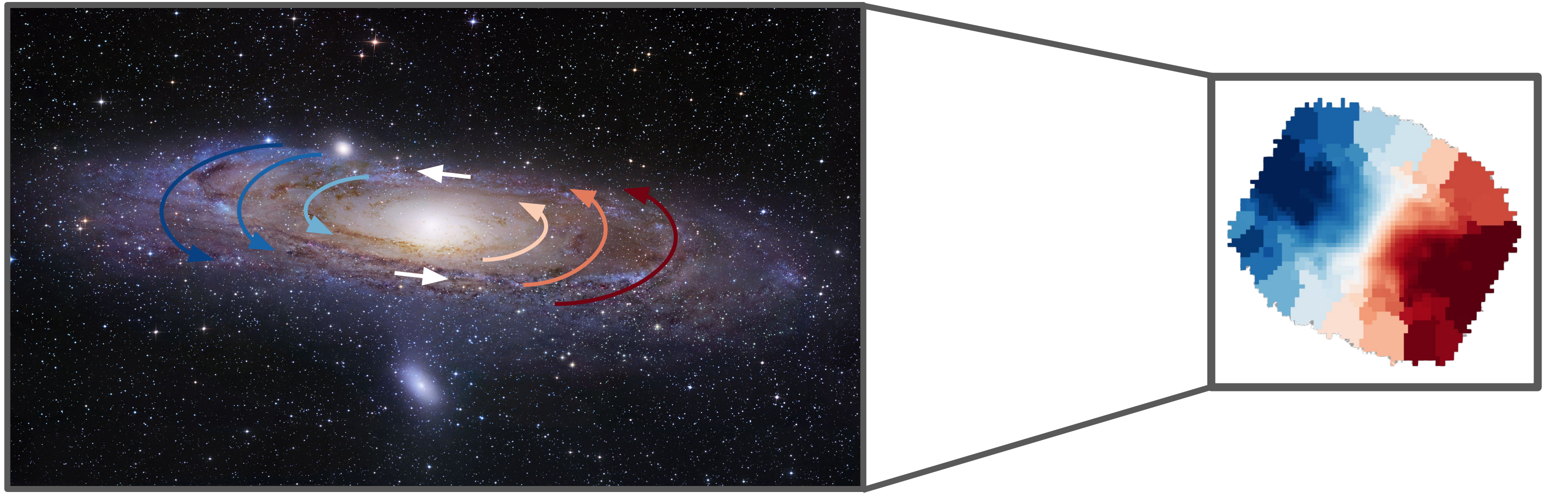
I'm happy to talk more about the details of how to create realistic kinematic maps.

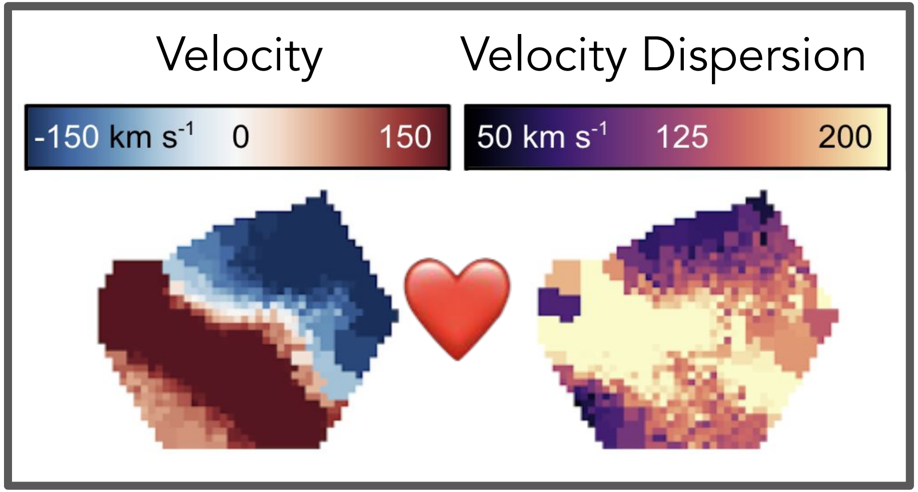


‘Who wants to be a merger?’

❤️ = Merger

😂 = Nonmerger

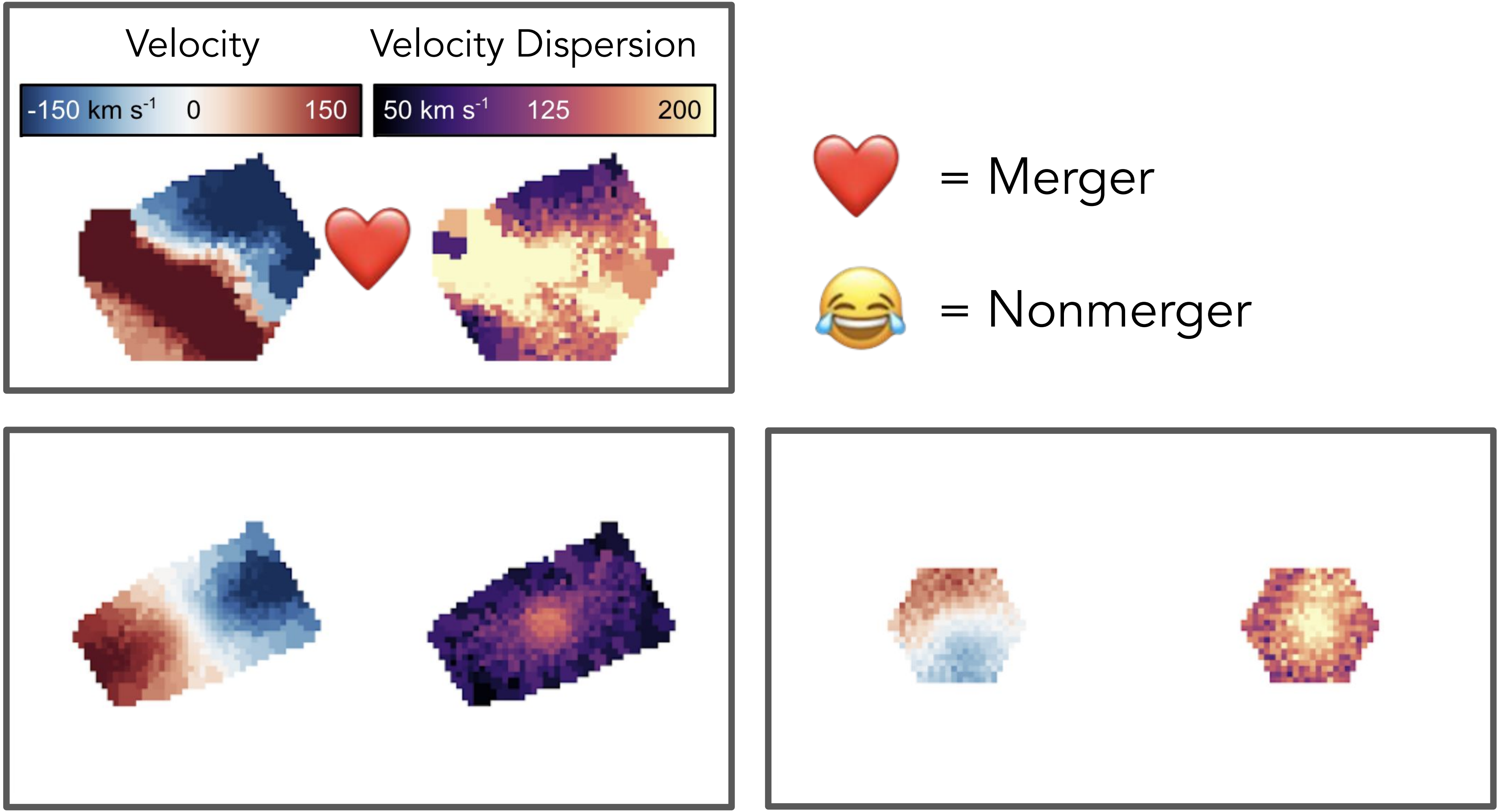




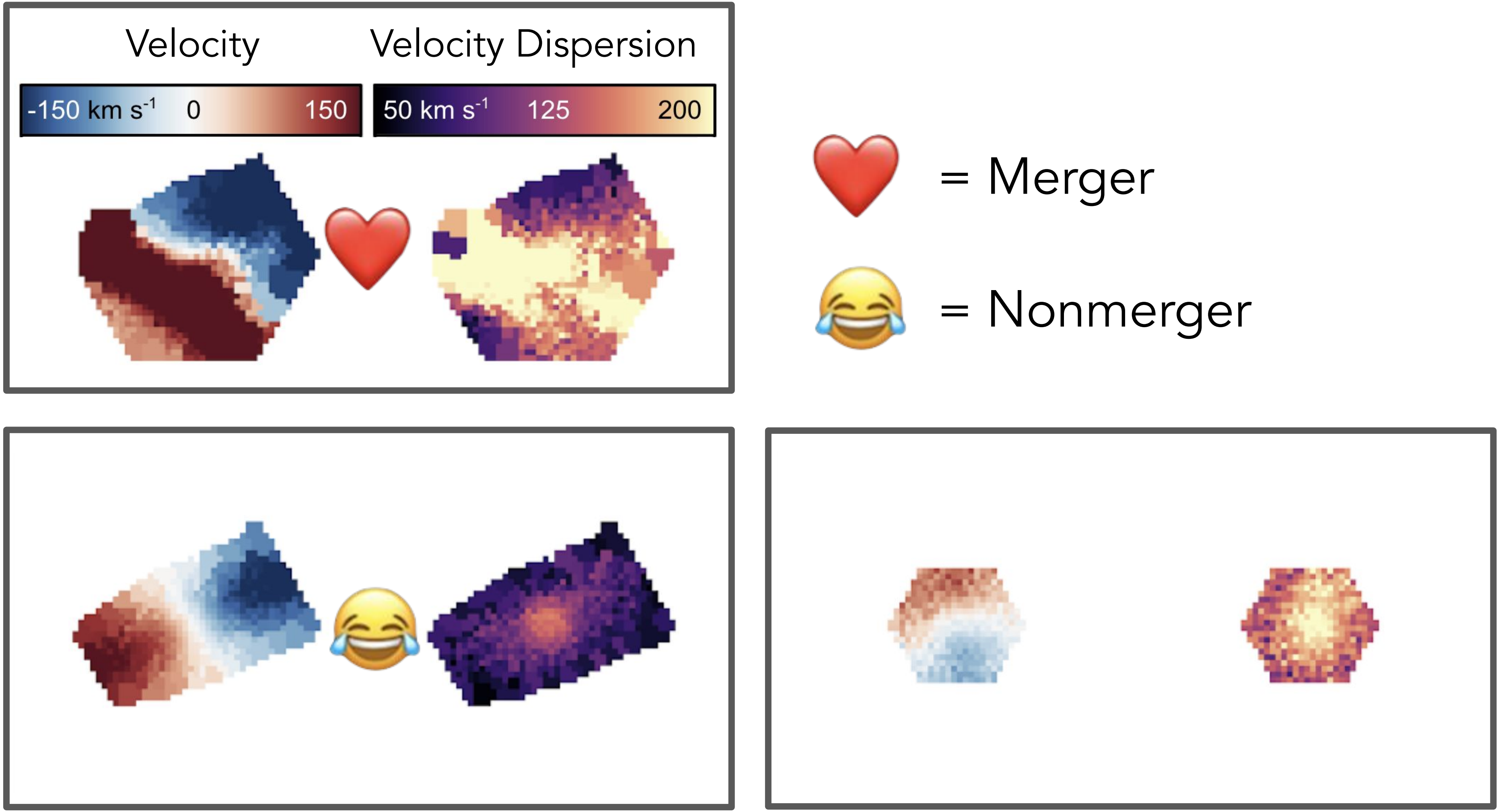
‘Who wants to be a merger?’

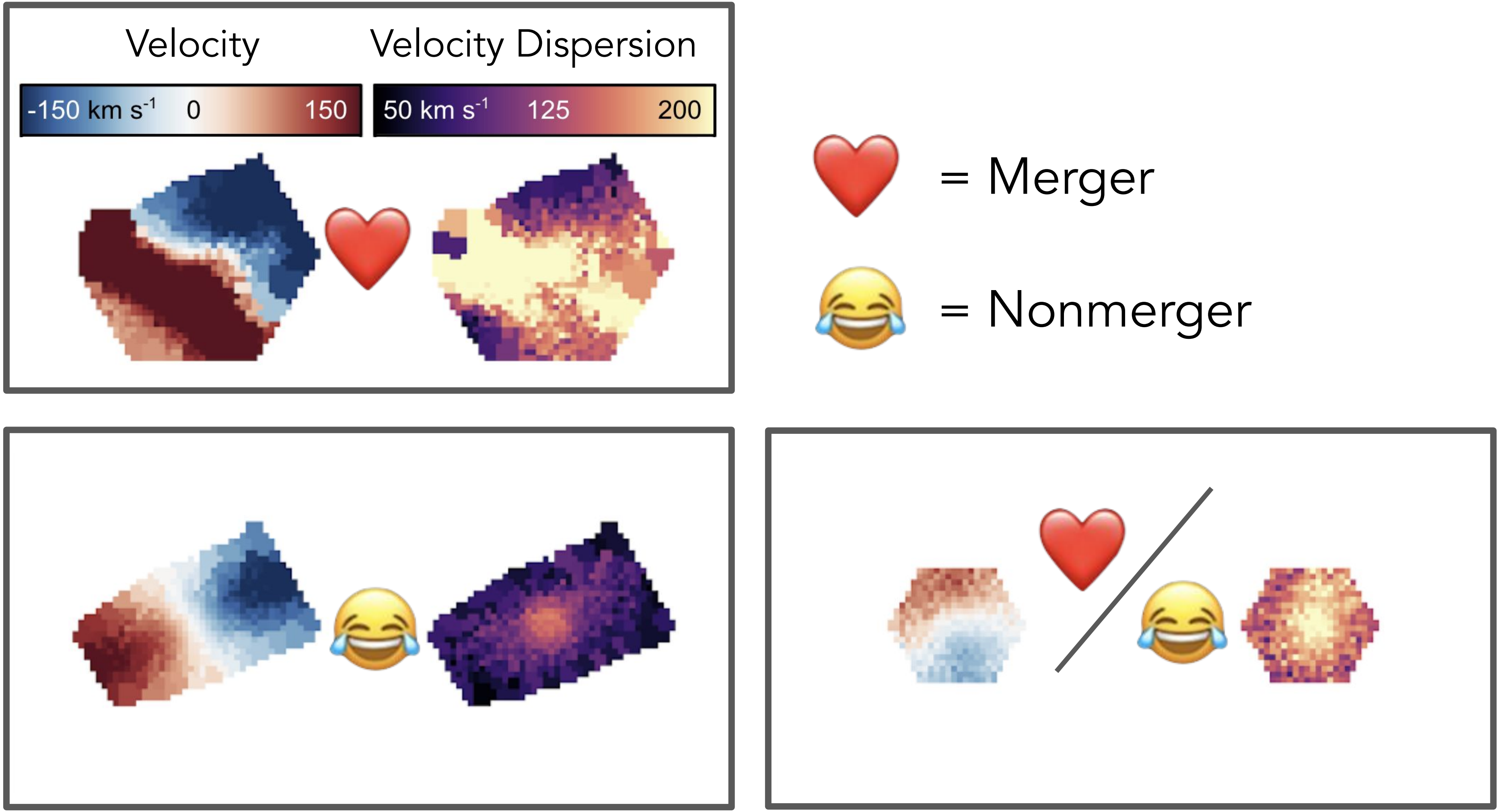
❤️ = Merger

😂 = Nonmerger

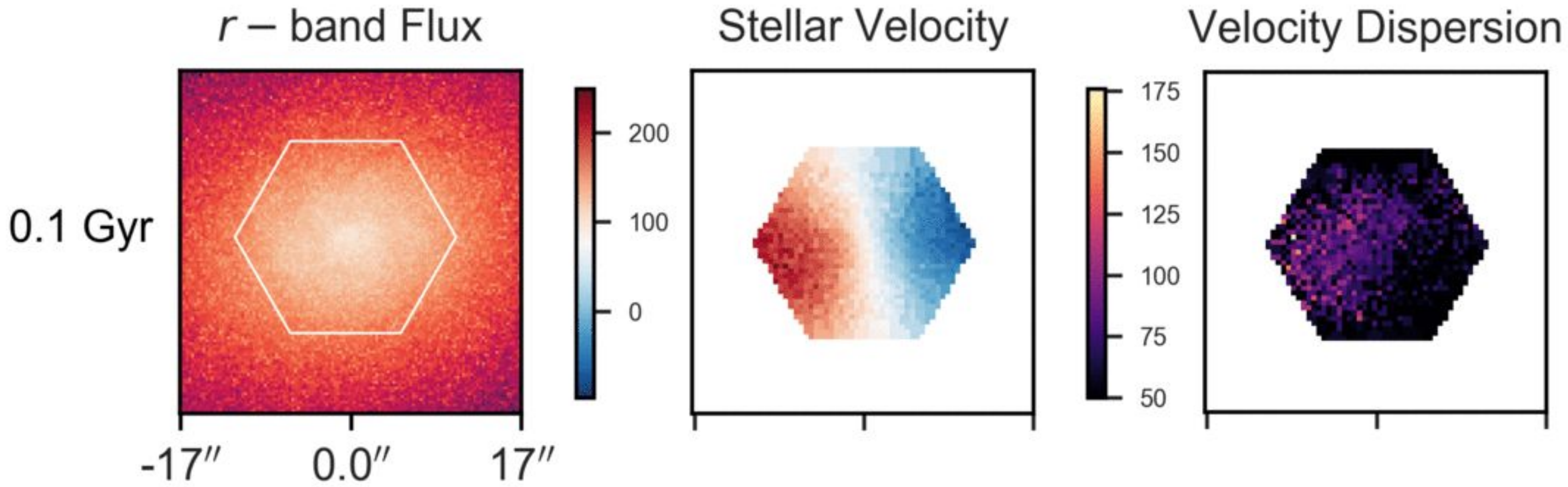


Velocity
Velocity Dispersion
-150 km s-1
0
150
50 km s-1
125
200
= Merger
= Nonmerger


Velocity
Velocity Dispersion
-150 km s-1 0 150
50 km s-1 125 200
= Merger
= Nonmerger

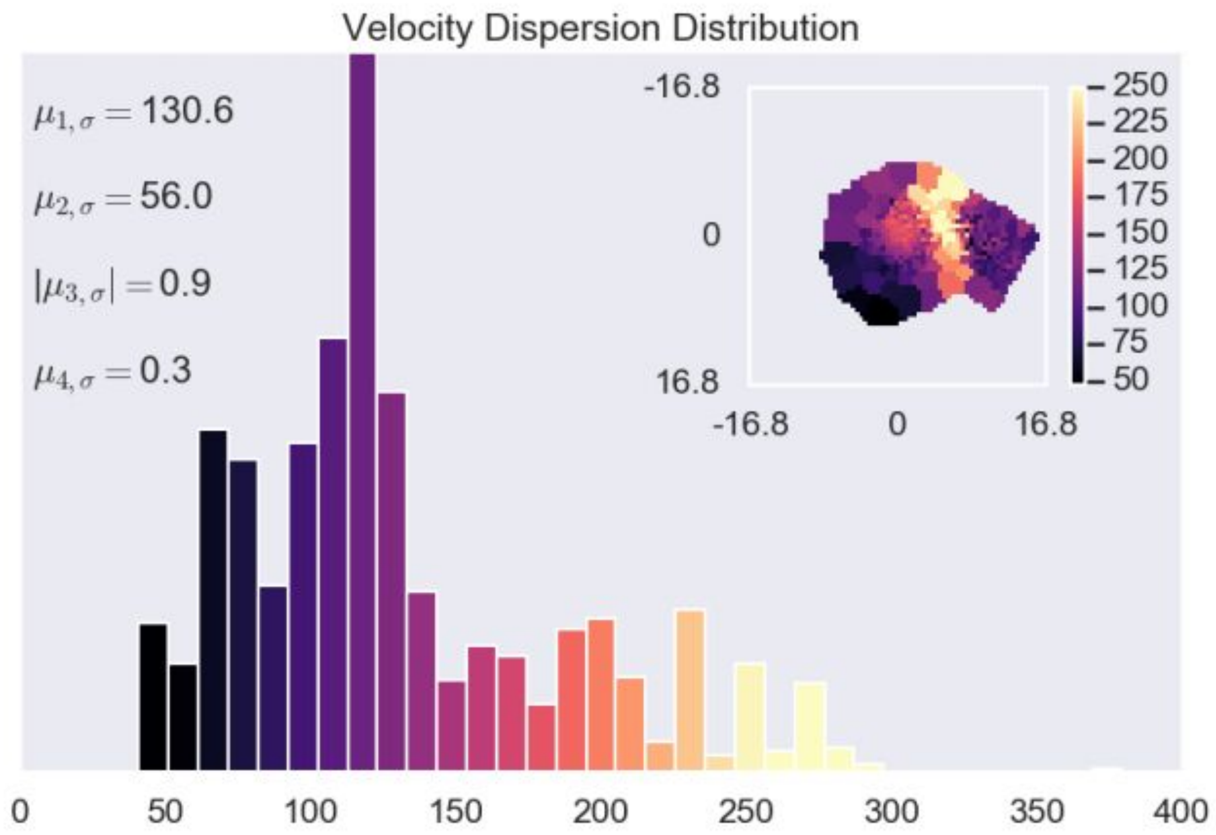

Velocity
Velocity Dispersion
-150 km s-1 0 150
50 km s-1 125 200
= Merger
= Nonmerger

I measure kinematic predictors that quantify the features in the kinematic maps over all stages



Radon Transform (A, A_2)
ΔPA
kinemetry (v_{asym}, σ_{asym}, resid)

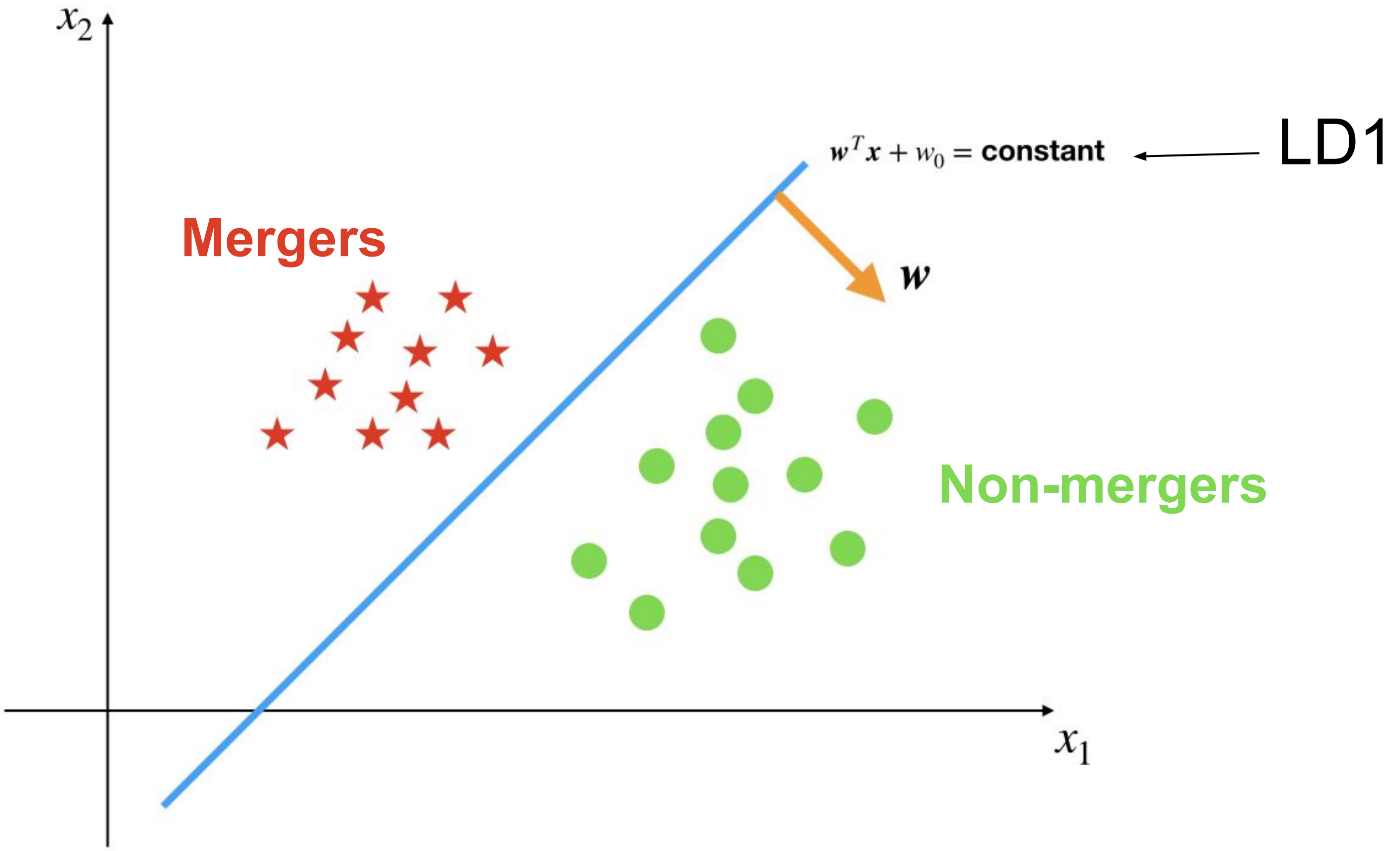
λ_{Re}
Δx_v, Δx_σ
moments of v and σ distributions ($\mu_{1,v}$, $\mu_{2,v}$, $\mu_{3,v}$, $\mu_{4,v}$, $\mu_{1,\sigma}$, $\mu_{2,\sigma}$, $\mu_{3,\sigma}$, $\mu_{4,\sigma}$)

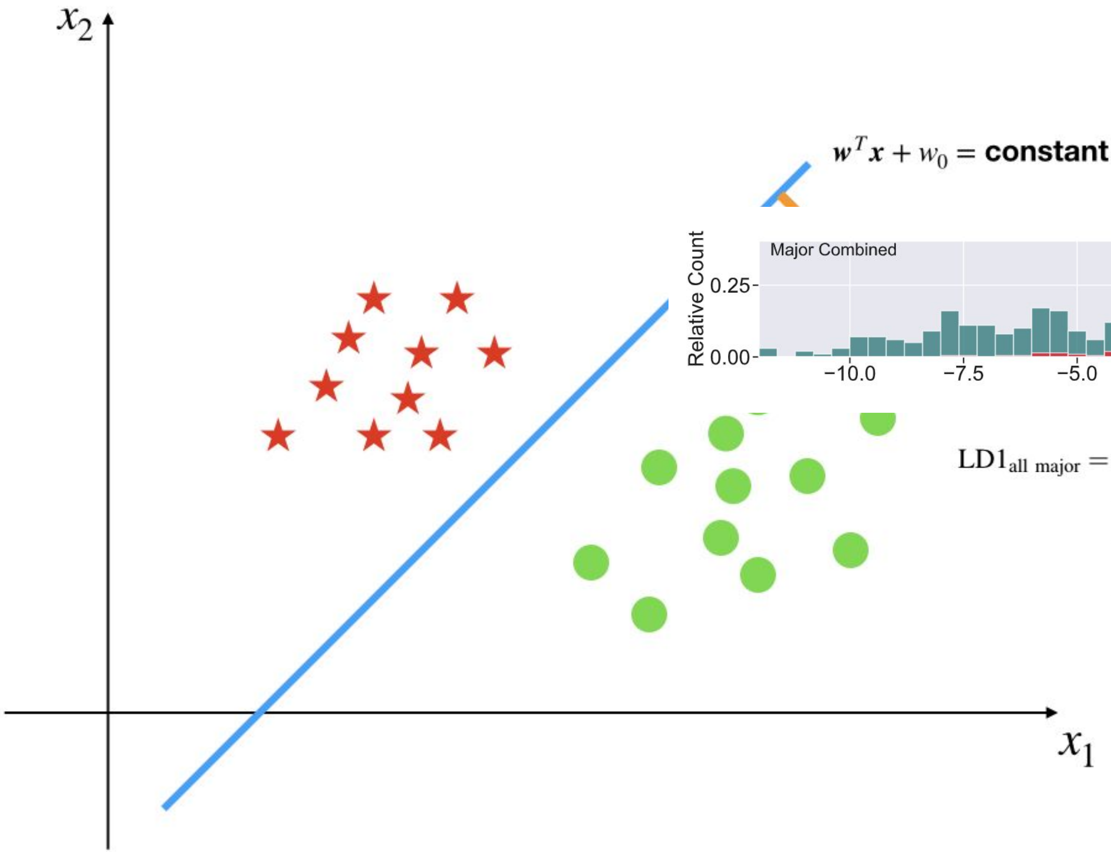


moments of v and σ distributions ($\mu_{1,v}$, $\mu_{2,v}$, $\mu_{3,v}$, $\mu_{4,v}$, $\mu_{1,\sigma}$, $\mu_{2,\sigma}$, $\mu_{3,\sigma}$, $\mu_{4,\sigma}$)

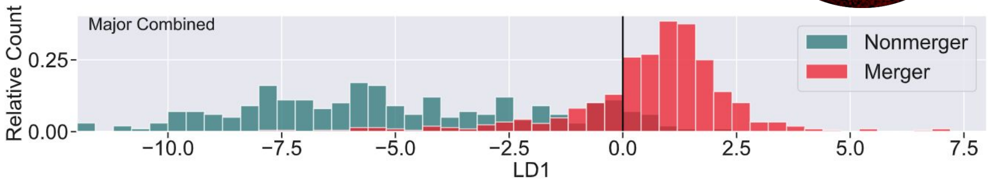
I combine all of the predictors into one statistical learning technique: linear discriminant analysis (LDA)

I combine all of the predictors into one statistical learning technique: linear discriminant analysis (LDA)



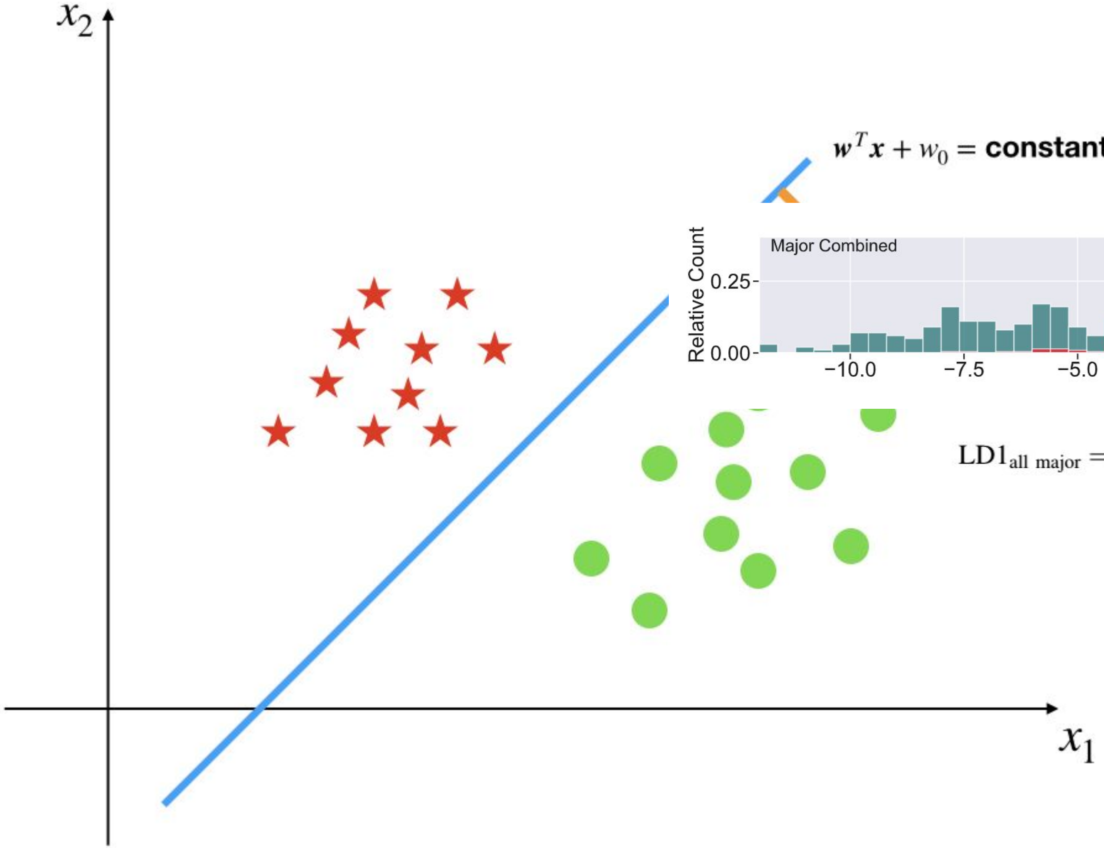


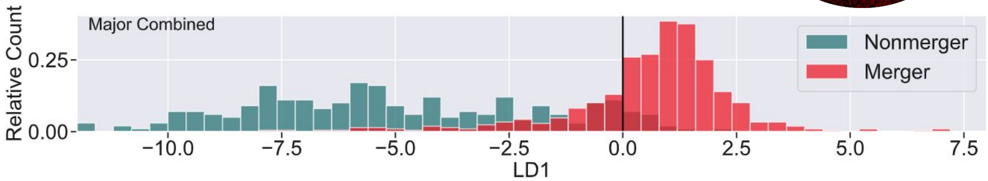
$\boldsymbol{w}^T\boldsymbol{x} + w_0 =$ **constant** ← LD1



$$
\begin{aligned}
\mathrm{LD1}_{\text{all major}} = & -6.8\lambda_{R_e} + 5.0|h_{3,\sigma}| + 4.5\mu_\sigma * \lambda_{R_e} \\
& - 4.4\mu_\sigma * |h_{3,\sigma}| - 1.0\mu_\sigma * \text{resids} \\
& + 1.7\mu_V * \lambda_{R_e} + 1.7h_{4,\sigma} * h_{4,V} \\
& + \ldots - 1.2
\end{aligned}
$$

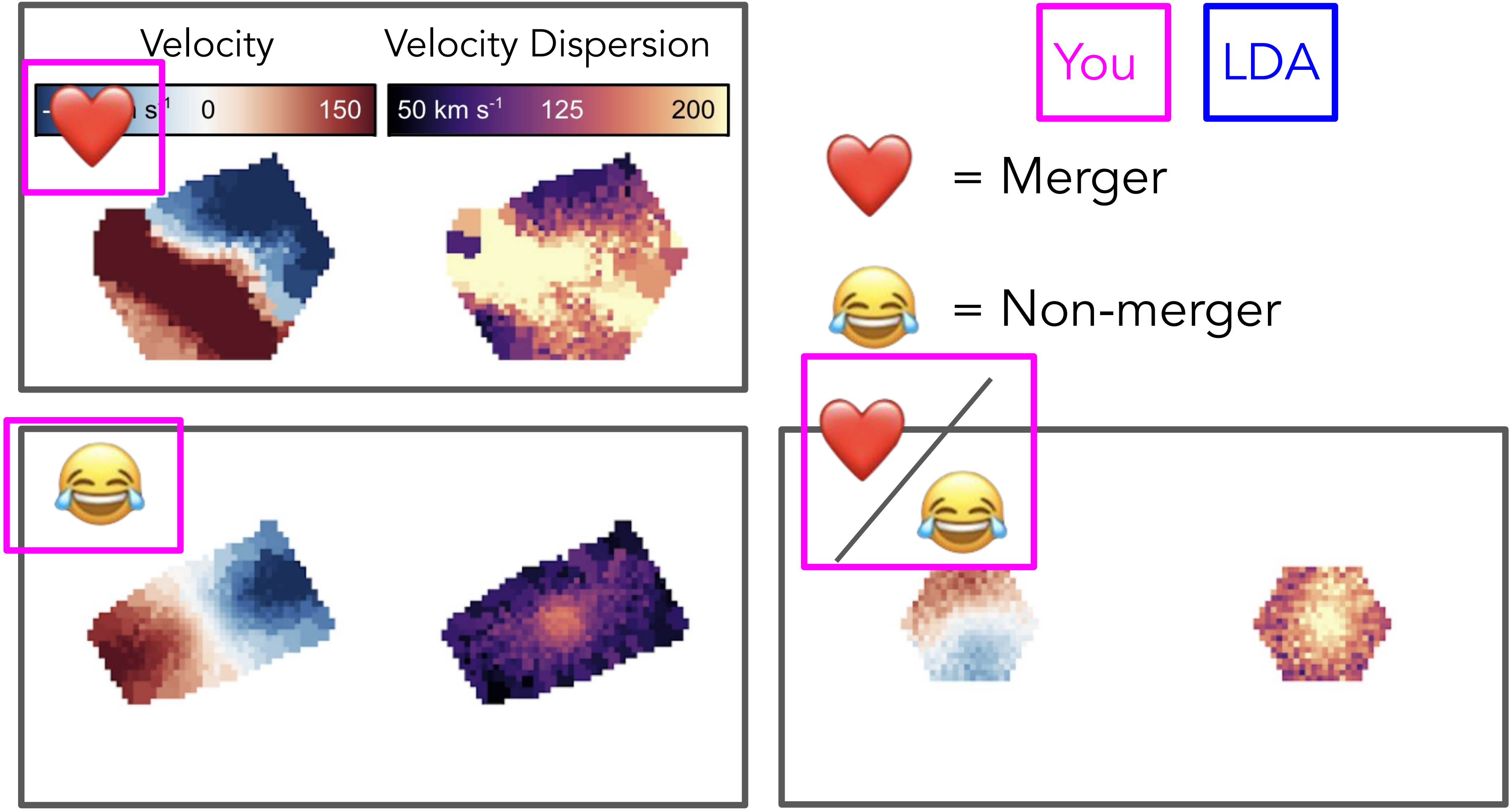
$w^T x + w_0 =$ **constant** ← LD1

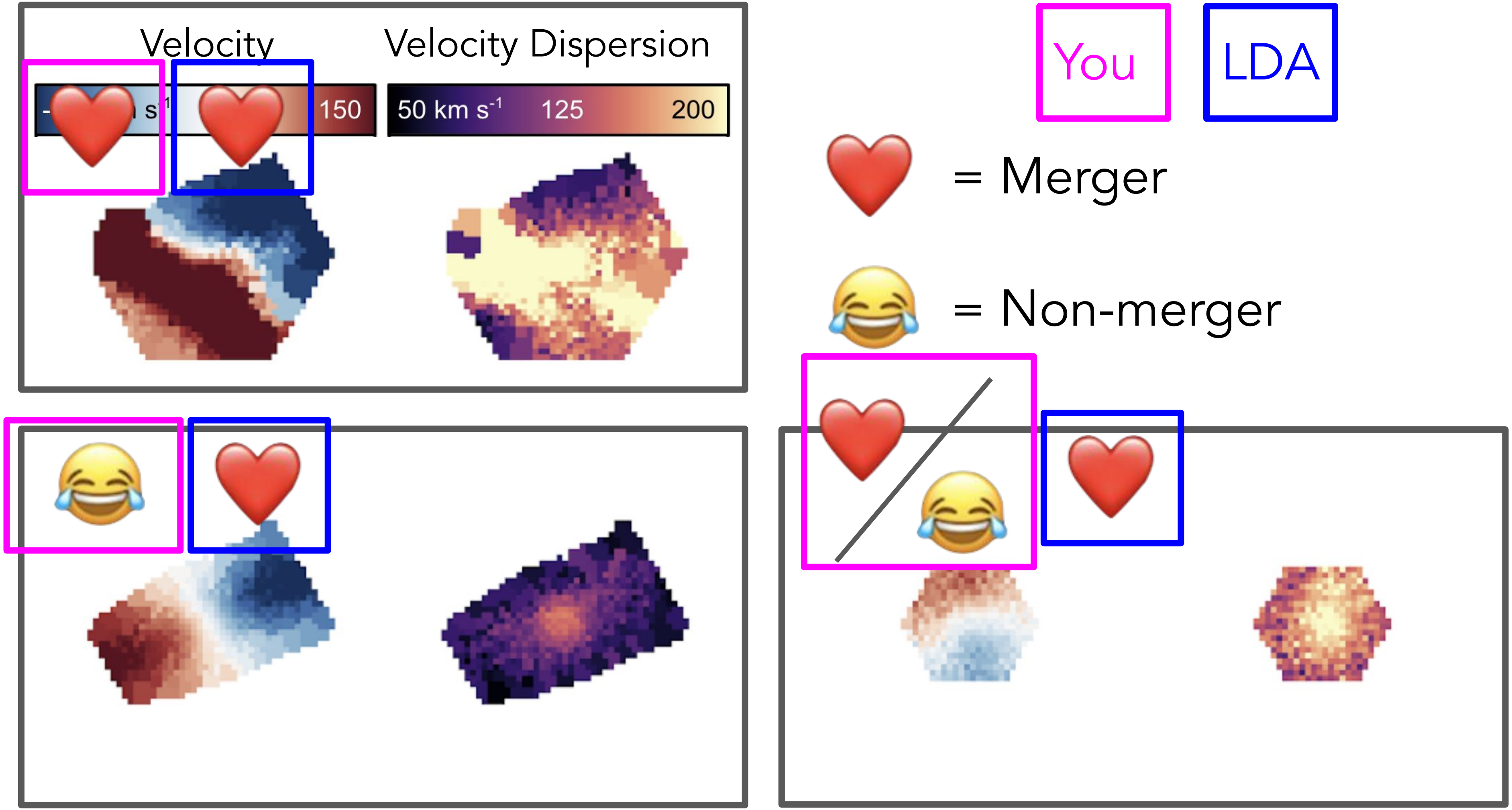




$$\begin{aligned}\mathrm{LD1}_{\text{all major}} = & -6.8\lambda_{R_e} + 5.0|h_{3,\sigma}| + 4.5\mu_\sigma * \lambda_{R_e} \\ & -4.4\mu_\sigma * |h_{3,\sigma}| - 1.0\mu_\sigma * \text{resids} \\ & +1.7\mu_V * \lambda_{R_e} + 1.7h_{4,\sigma} * h_{4,V} \\ & + \ldots - 1.2\end{aligned}$$

Which kinematic predictors are most informative?

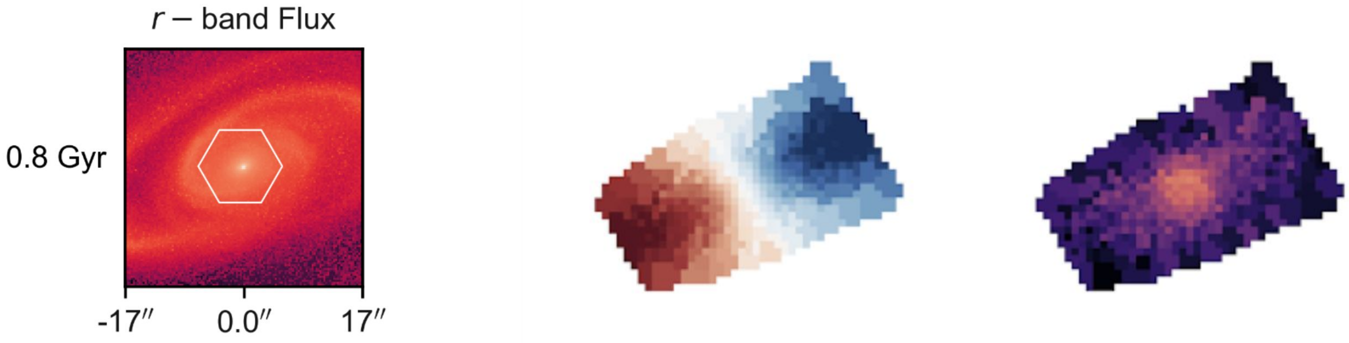

Velocity
Velocity Dispersion
0
150
50 km s-1
125
200
You
LDA
= Merger
= Non-merger



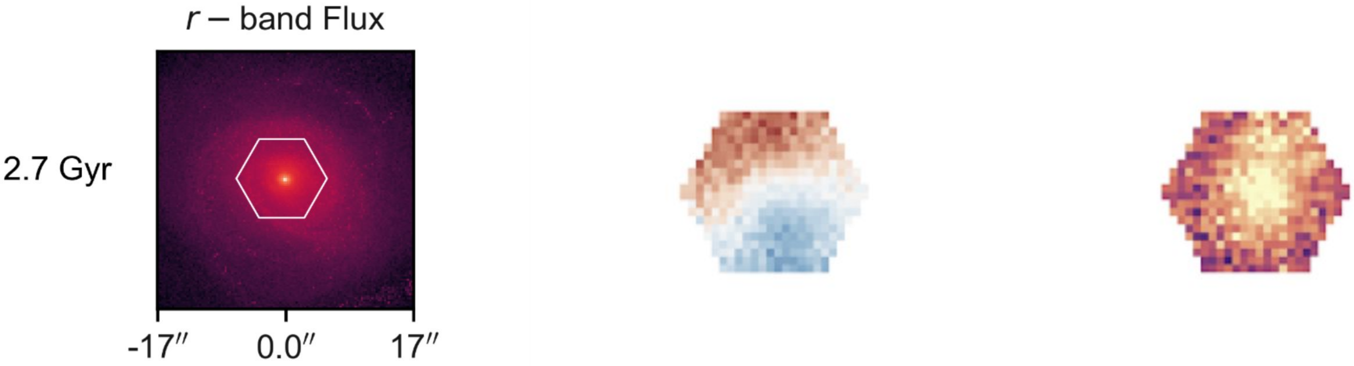
Velocity
Velocity Dispersion
150
50 km s-1
125
200
You
LDA
= Merger
= Non-merger

How does the performance of the classification vary with time?

1) Pre-coalescence mergers are disky

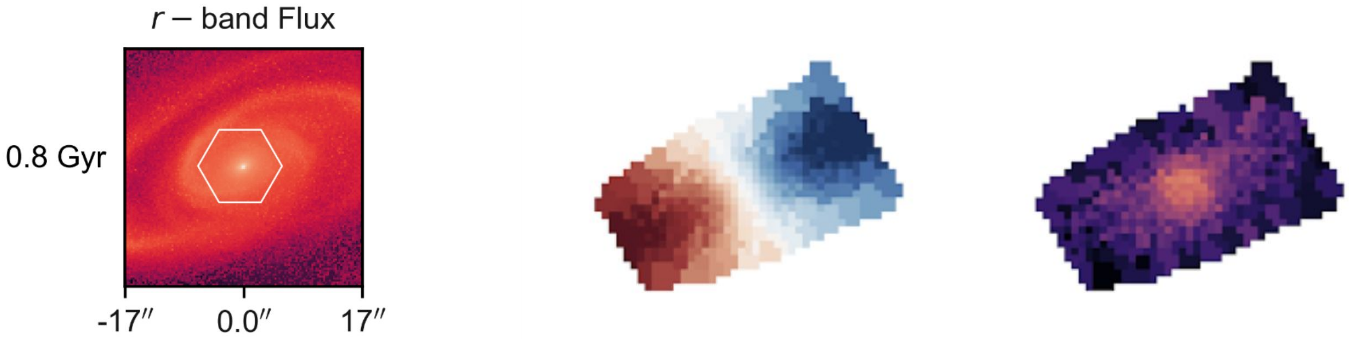


2) Post-coalescence mergers have long-lived features

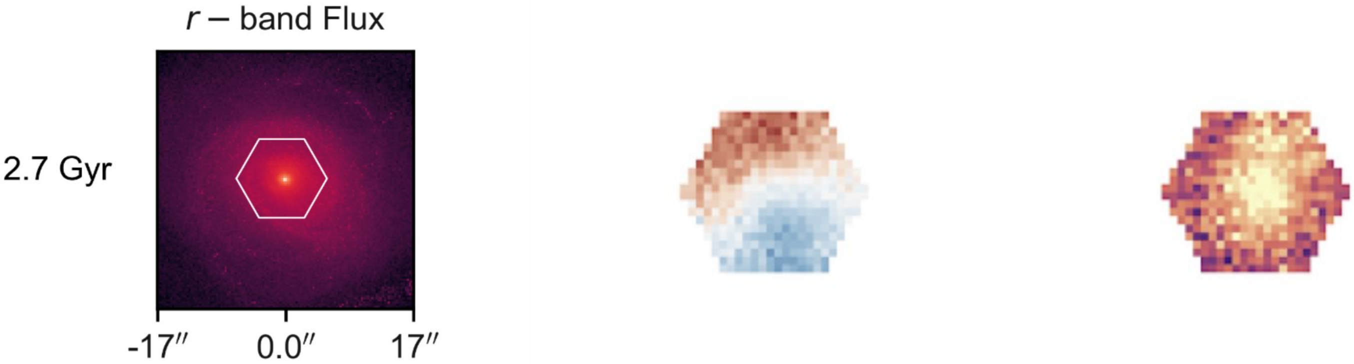


How does the performance of the classification vary with time?

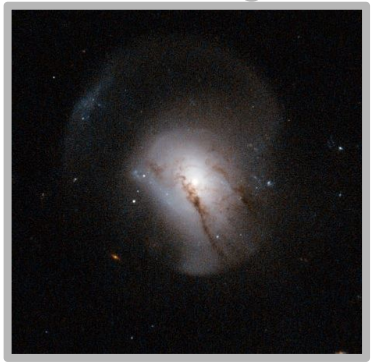
1) Pre-coalescence mergers are disky



2) Post-coalescence mergers have long-lived features



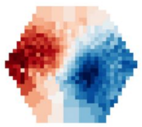
Post-merger

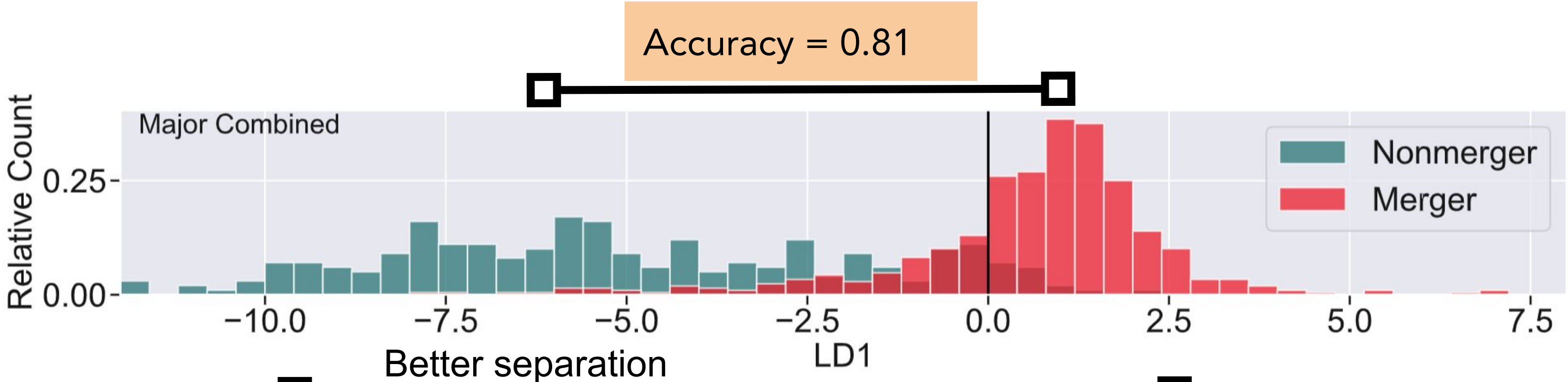


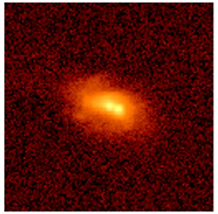
When does a merger end?

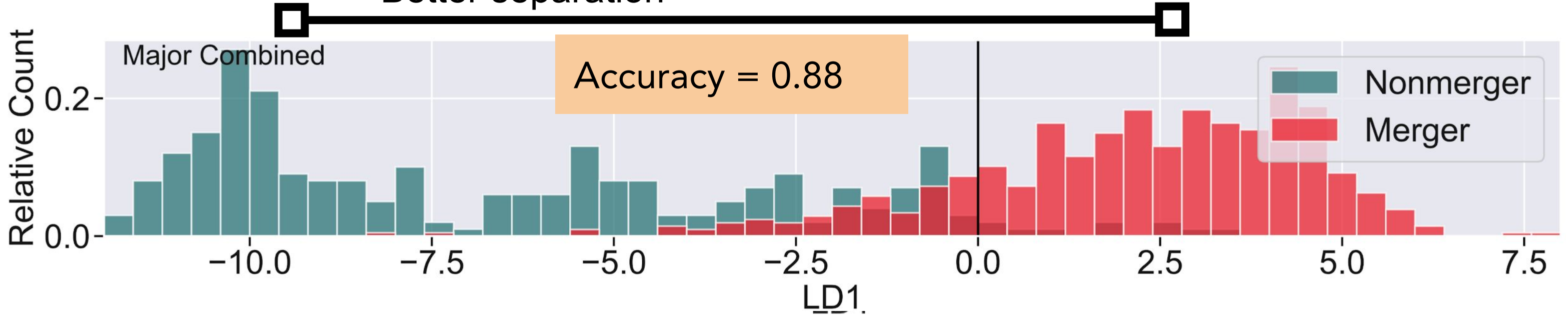
How does the kinematic classification compare to the imaging classification?

The imaging LDA is more accurate



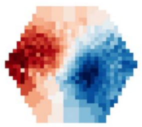


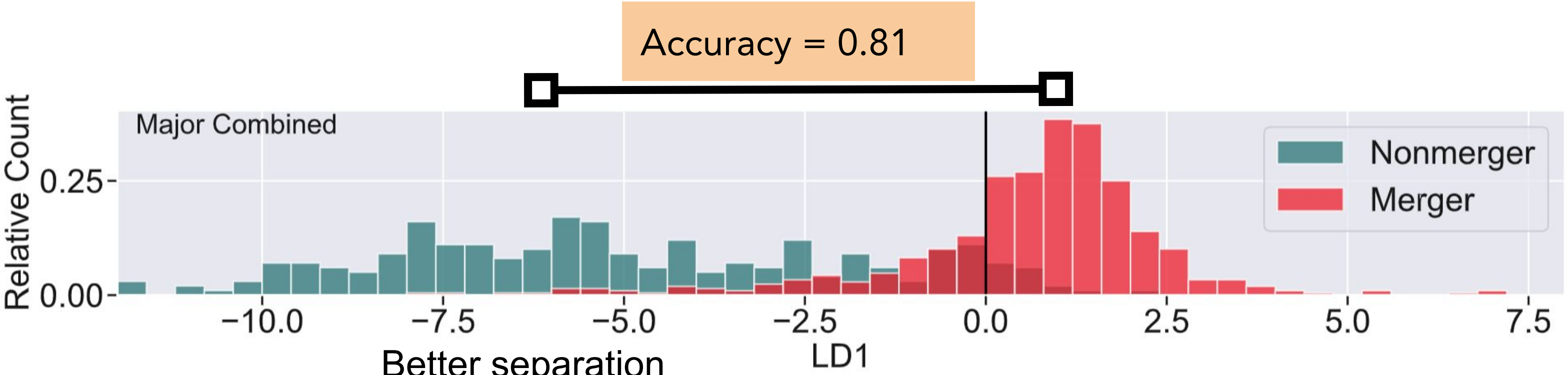




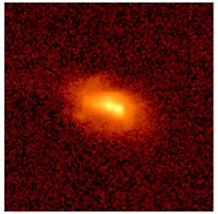
How does the kinematic classification compare to the imaging classification?

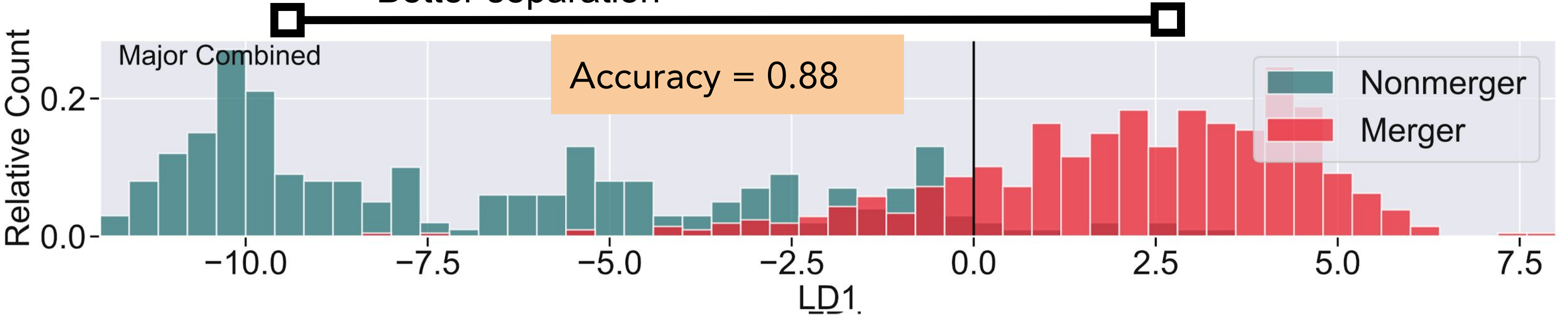
The imaging LDA is more accurate





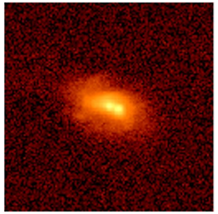
Better separation



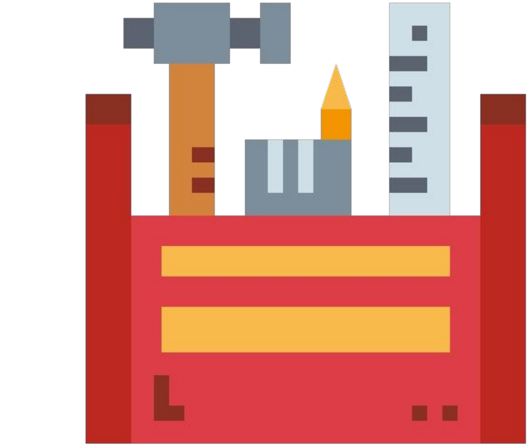


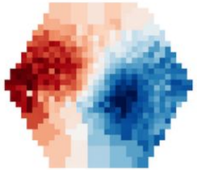
What does the kinematic classification add to our toolkit?

What does the kinematic classification add to our toolkit?



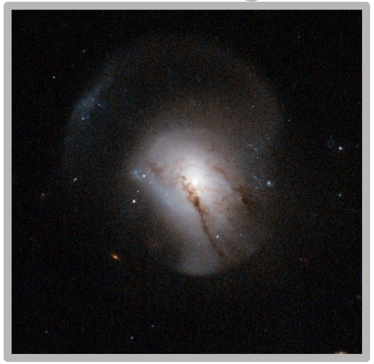
Better at finding pre-coalescence mergers





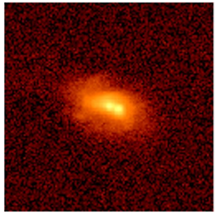
Better at finding post-coalescence mergers

Post-merger

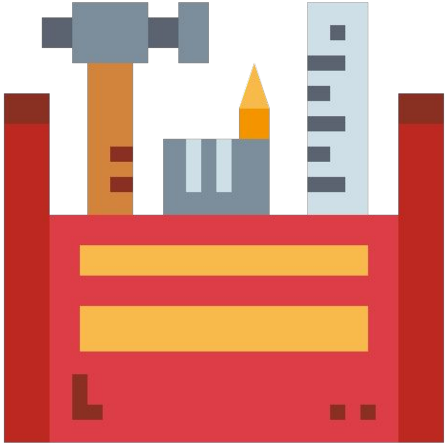


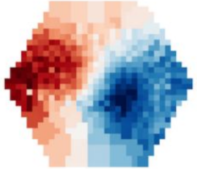
These are complementary methods, combining them shows an improvement in performance

What does the kinematic classification add to our toolkit?



Better at finding
pre-coalescence mergers





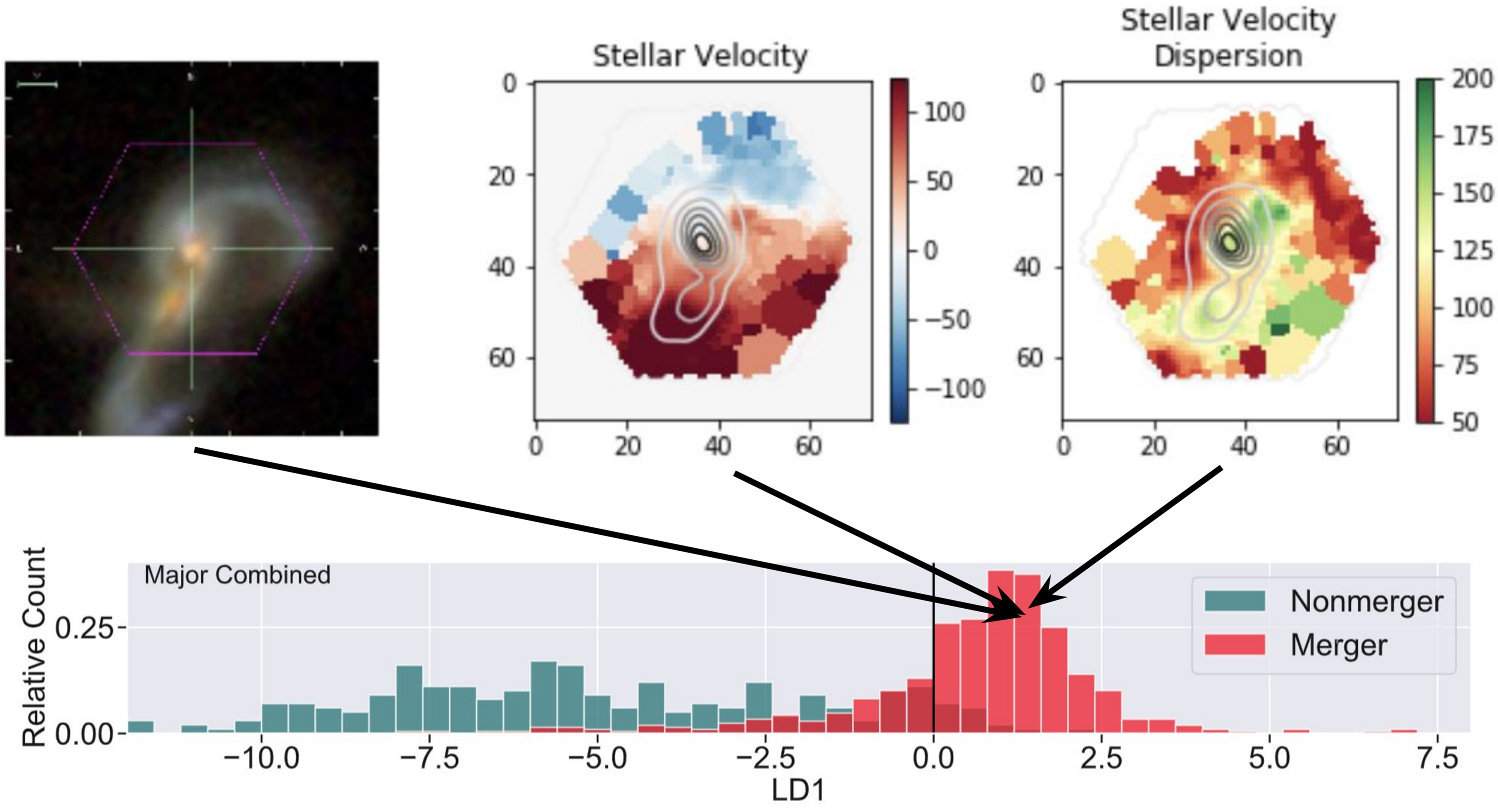
Better at finding
post-coalescence
mergers

These are complementary methods, combining them shows an improvement in performance

Next steps:

1) Apply the classification to galaxies in SDSS/MaNGA

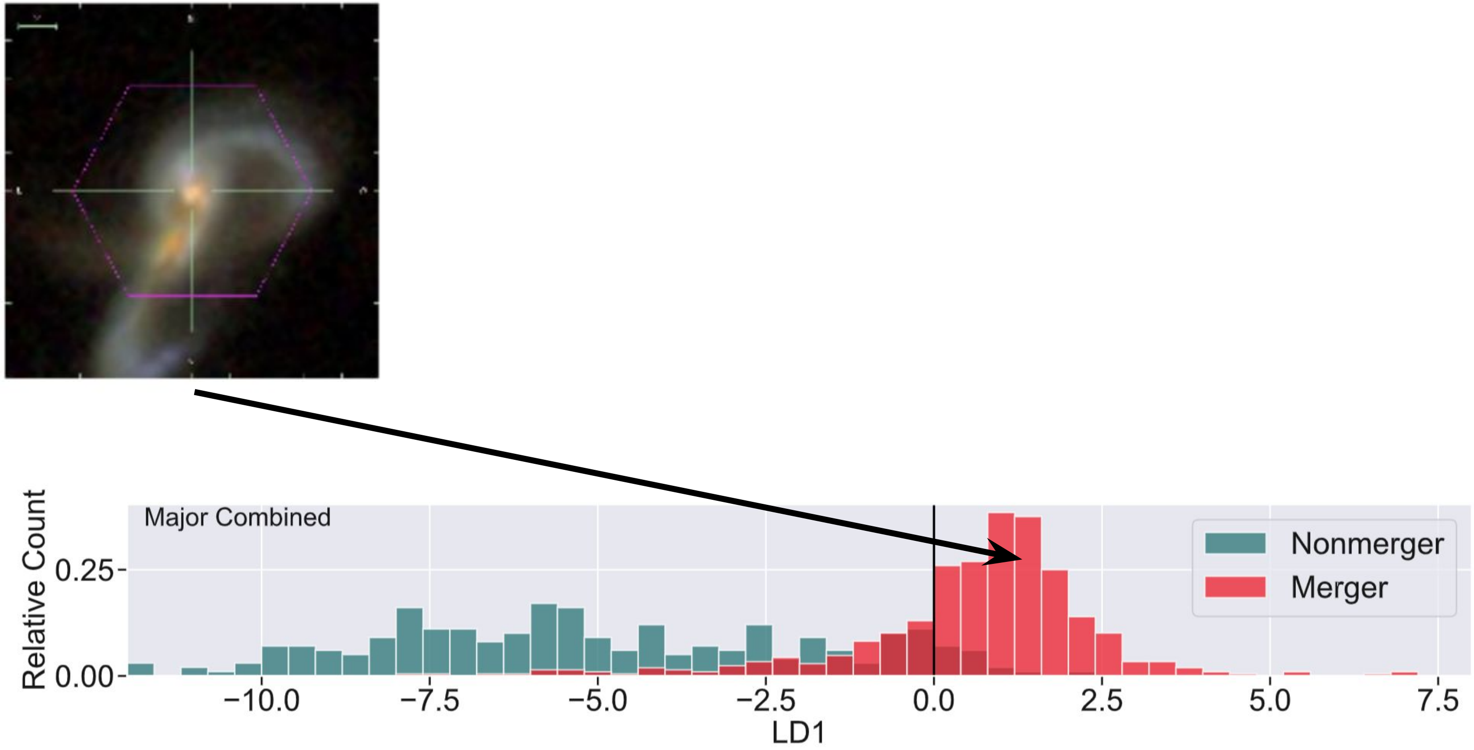
2) Further split by merger stage



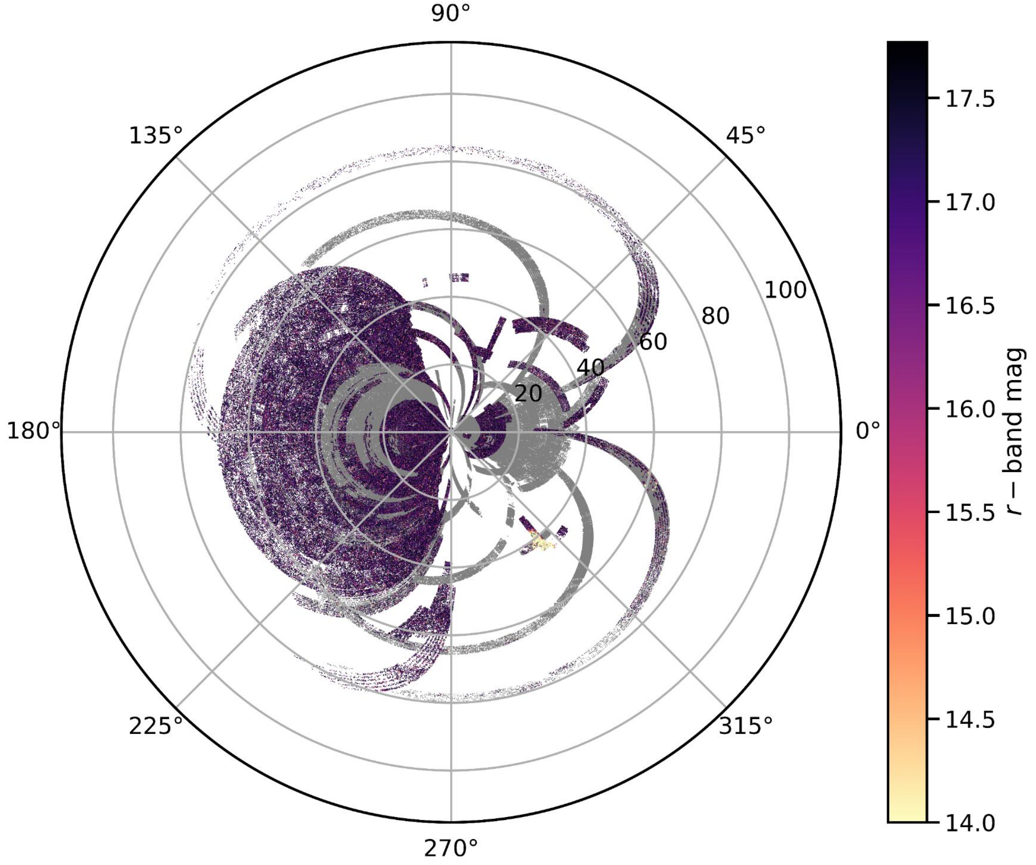
Next steps:

1) Apply the classification to galaxies in SDSS/MaNGA

2) Further split by merger stage



I measure photometric properties (Gini, asymmetry, etc) for the 1.3 million galaxy DR16 photometric sample:



ObjID:1237665329864114245

Step 1: Measure predictor values

Gini = 0.47

M_{20} = -0.96

C = 1.83

A = 0.49

S = 0.04

n = 0.42

A_s = 0.16

ObjID:1237665329864114245

Step 1: Measure predictor values

Gini = 0.47

M_{20} = -0.96

C = 1.83

A = 0.49

S = 0.04

n = 0.42

A_s = 0.16

Step 2: Standardize the predictor values and plug into the LD1 formulae, e.g.:

$$LD1_{major,\ pre} = 11.66\ A_S - 7.76\ A_S{*}C - 6.5\ A_S{*}A + 5.72\ A + 4.51\ C + 0.41\ S - 0.91$$

$$= 11.66{*}(-0.68) - 7.76{*}(-0.98) - 6.5{*}(0.31) + 5.72{*}(2.69) + 4.51{*}(-2.18) + 0.41{*}(0.90) - 0.91$$

$$= -7.9 + 7.6 - 2.0 + 15.4 - 9.8 + 0.37 - 0.91$$

$$= 2.76$$

ObjID:1237665329864114245

Step 1: Measure predictor values

Gini = 0.47

M_{20} = -0.96

C = 1.83

A = 0.49

S = 0.04

n = 0.42

A_s = 0.16

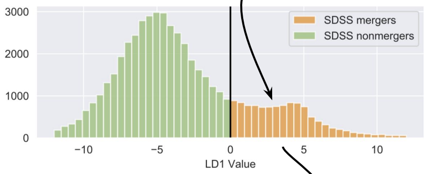
Step 2: Standardize the predictor values and plug into the LD1 formulae, e.g.:

$$LD1_{major,\ pre} = 11.66\ A_S - 7.76\ A_S{*}C - 6.5\ A_S{*}A + 5.72\ A + 4.51\ C + 0.41\ S - 0.91$$

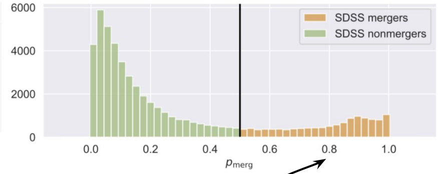
$$= 11.66{*}(-0.68) - 7.76{*}(-0.98) - 6.5{*}(0.31) + 5.72{*}(2.69) + 4.51{*}(-2.18) + 0.41{*}(0.90) - 0.91$$

$$= -7.9 + 7.6 - 2.0 + 15.4 - 9.8 + 0.37 - 0.91$$

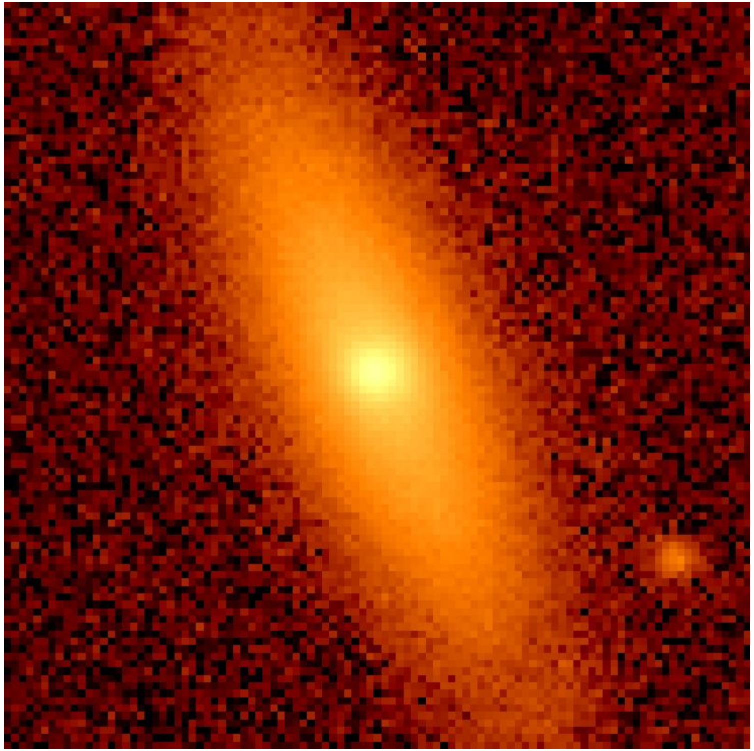
$$= 2.76$$



Step 3: Solve for p_{merg}



$$p_{merg} = 1/(1+e^{-LD1}) = 1/(1+e^{-(0.5)*2.76}) = 0.8$$



Probability of:

Minor merger = 0.71

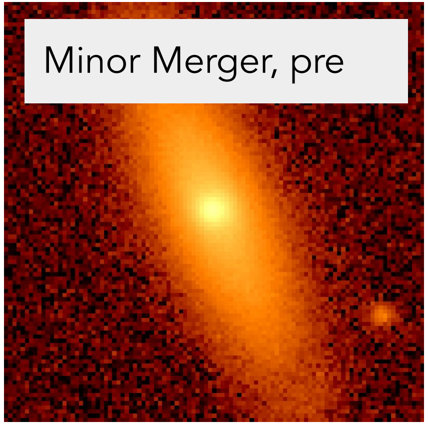
pre = 0.65

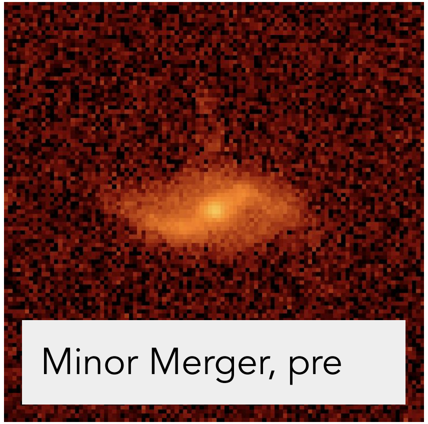
post = 0.16

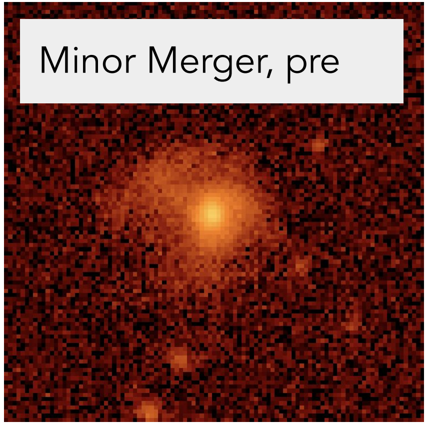
Major merger = 0.16

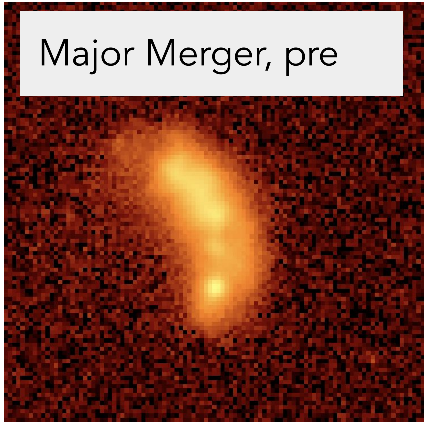
pre = 0.11

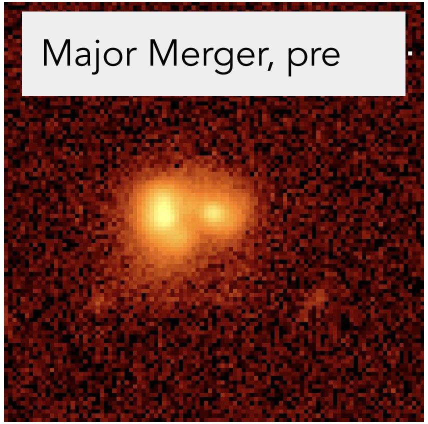
post = 0.06


Minor Merger, pre

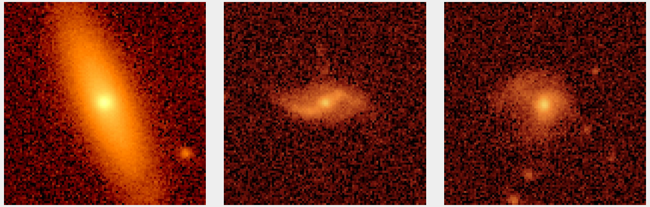

Minor Merger, pre


Minor Merger, pre

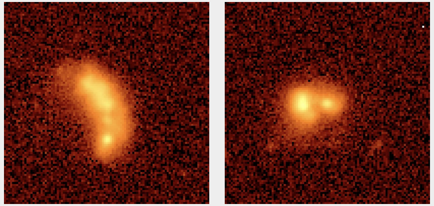

Major Merger, pre


Major Merger, pre

Measure star formation rate and AGN fraction for the different samples of mergers



Most likely minor mergers, pre-merging

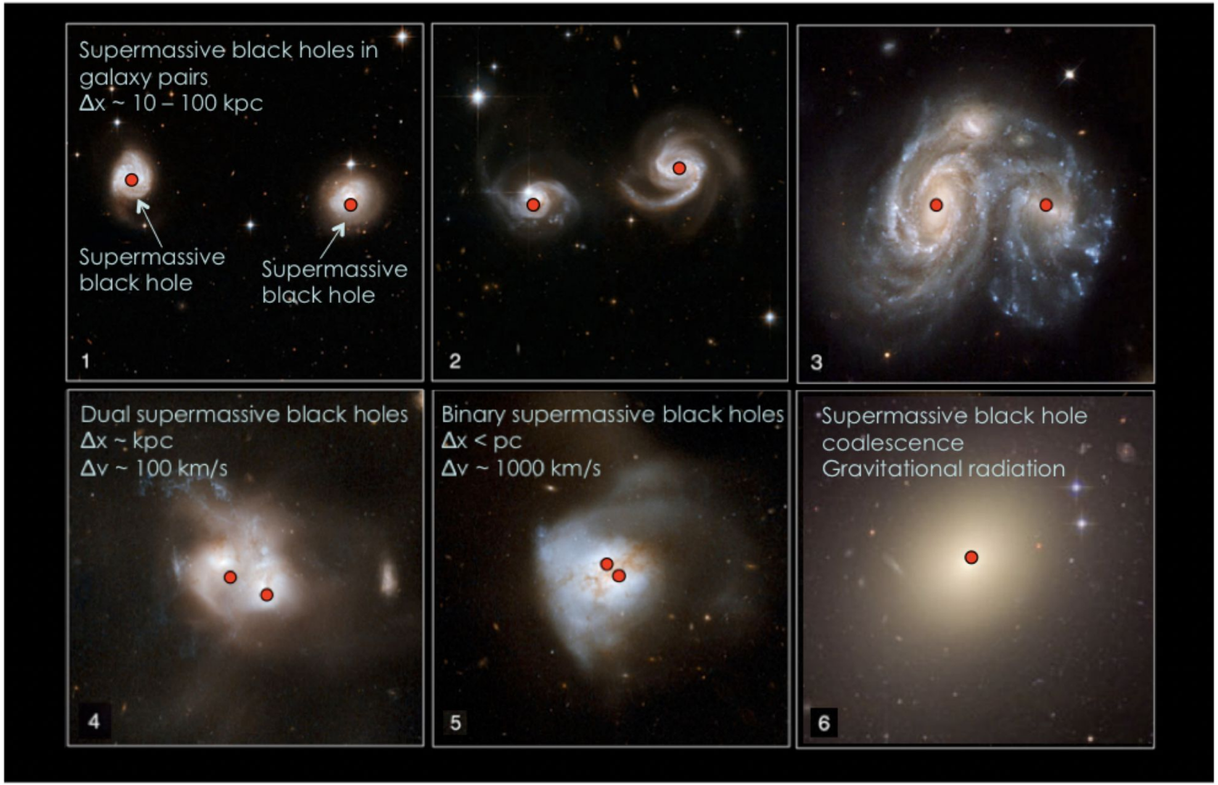


Most likely major mergers, pre-merging

Next step: Use the classification technique to determine the local (and non-local) galaxy merger rate → supermassive black hole merging rate → amplitude of the gravitational wave background

Joe Simon

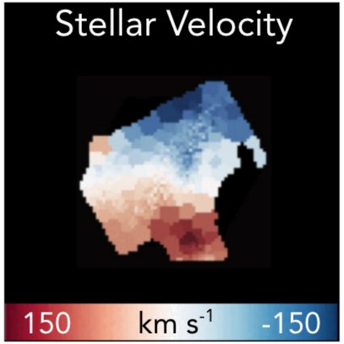
Julie Comerford

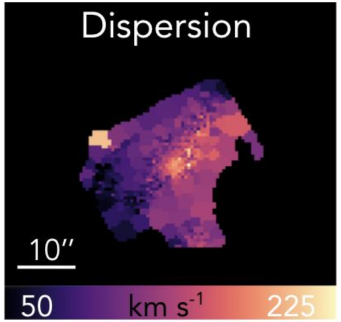


HST/NASA/STScI/ESA

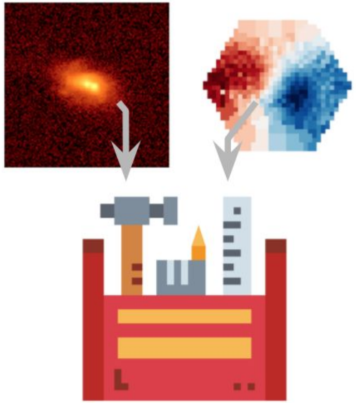
Main takeaways

Create mock MaNGA-ized kinematic maps





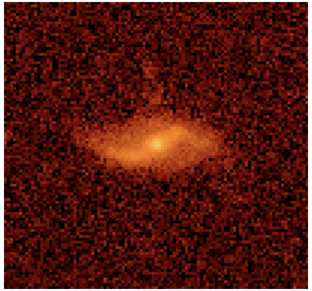
What does the kinematic classification add to our toolkit?



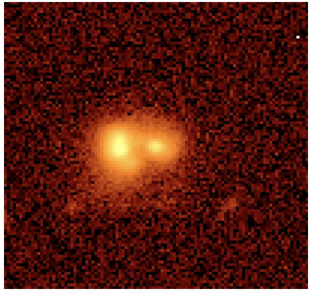
These are complementary methods, combining them shows an improvement in performance

Next step: Apply the classification to galaxies in SDSS/MaNGA

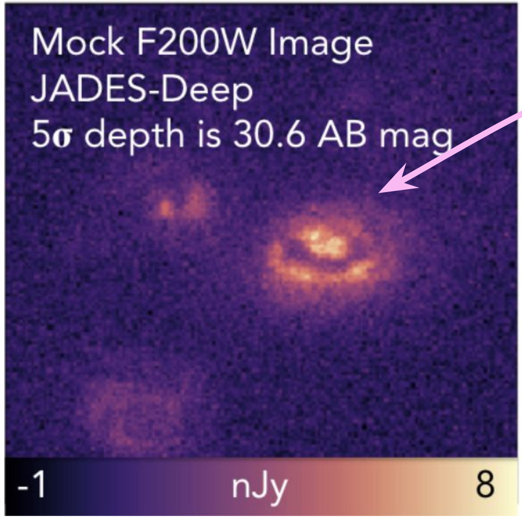
Minor, pre



Major, pre



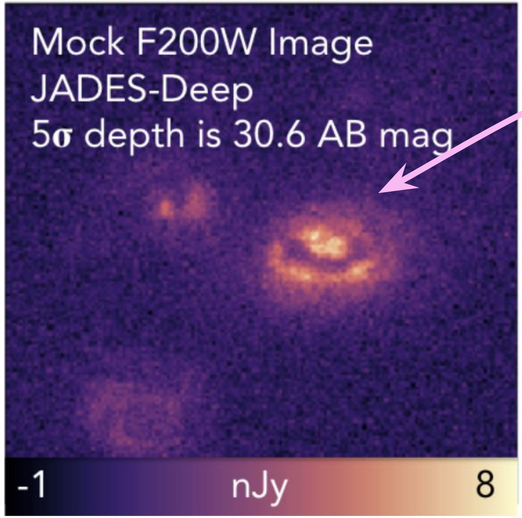
HTST-ized



CiNNamonroll:

A convolutional neural network framework to identify mergers during cosmic noon and brunch

HTST-ized

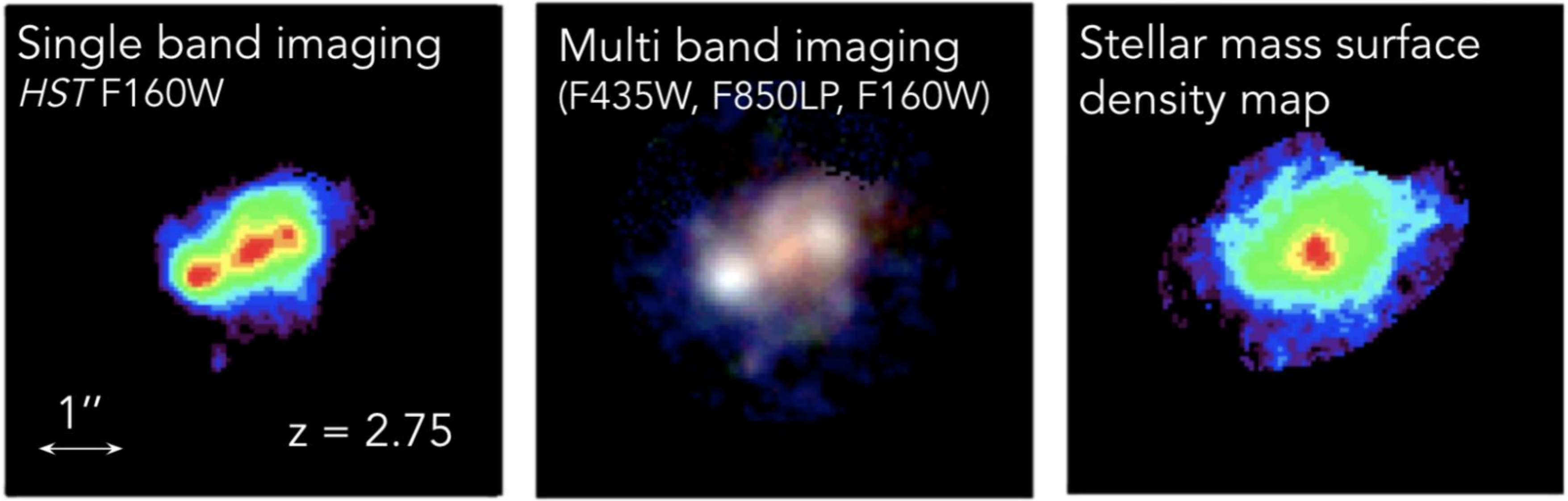


A sweet suite of CNNs and Illustris TNG50 simulated mergers

CiNNamonroll:

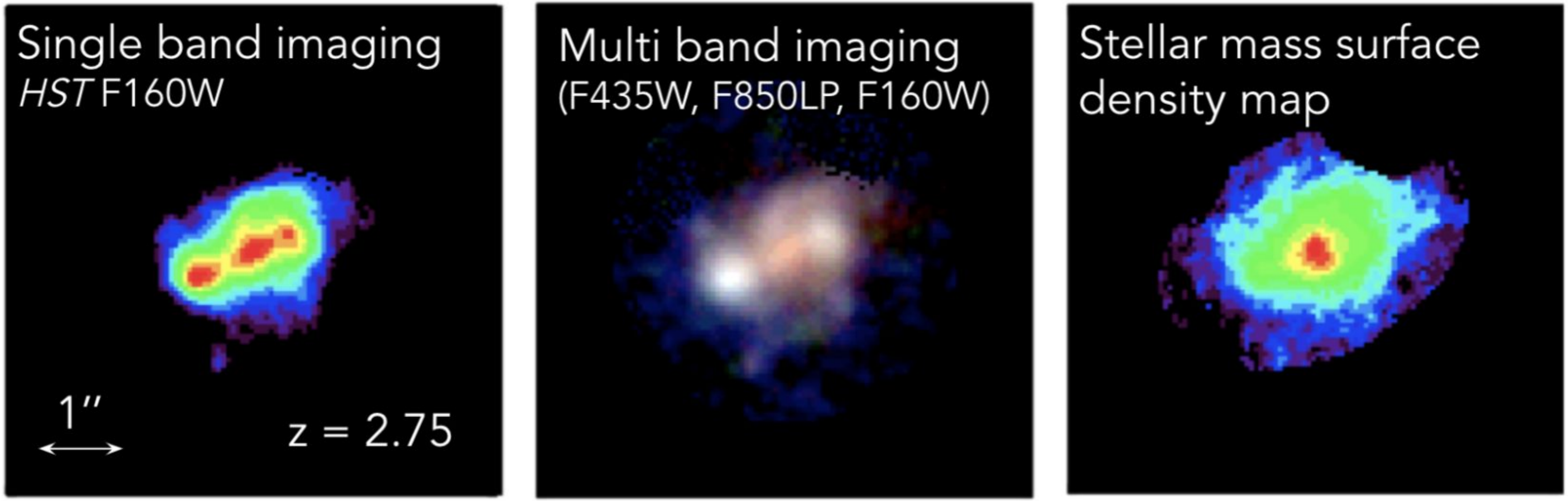
A convolutional neural network framework to identify mergers during cosmic noon and brunch

High redshift galaxies are inherently clumpy and mergers are harder to identify



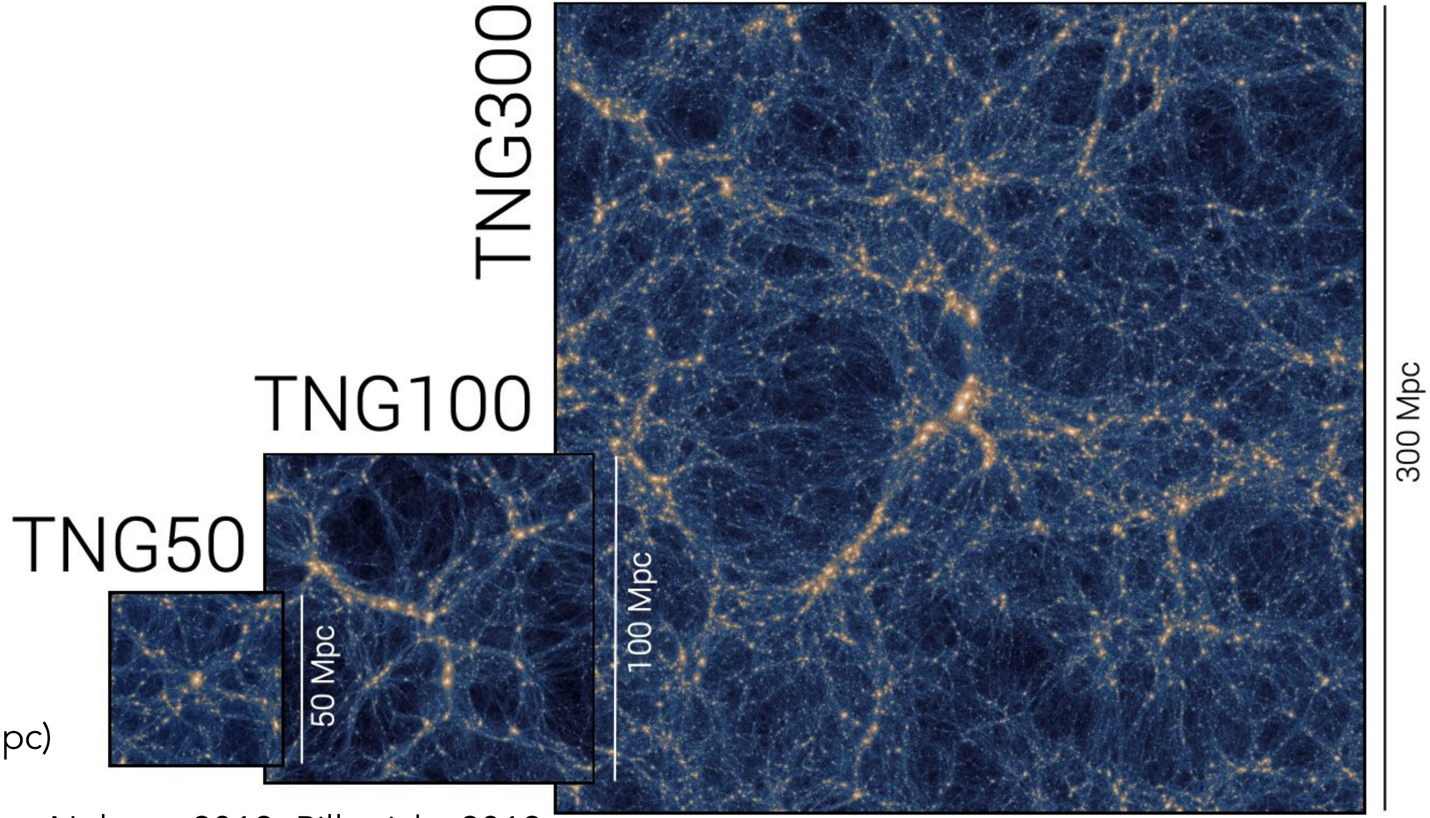
Cibinel+2015

Tools derived from multiple filters can enable more accurate merger identification



Cibinel+2015

New training set! → Illustris TNG50

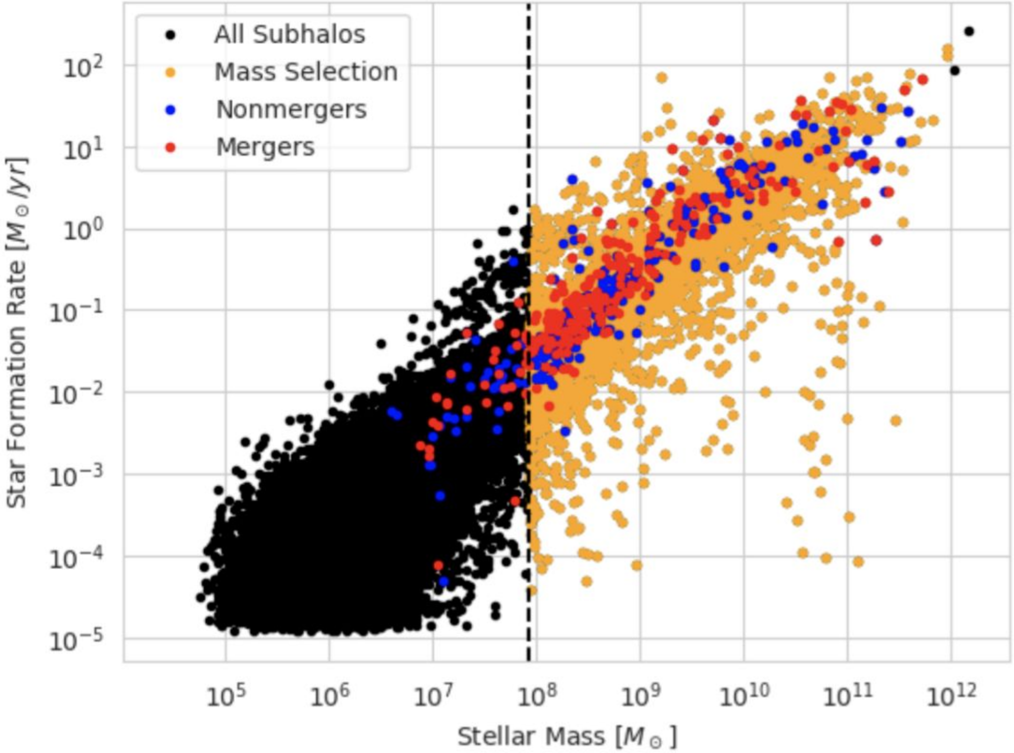


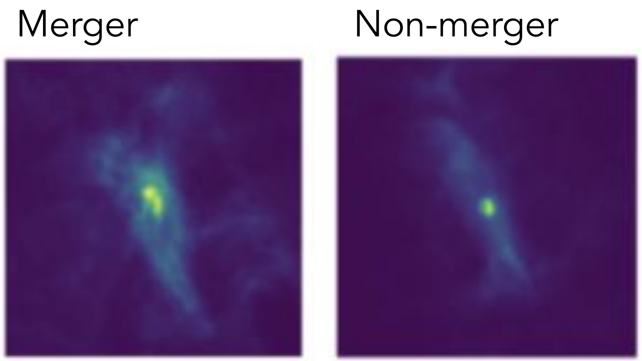
~72pc resolution
(TNG100 is about ~190pc)

TNG50 presentation papers: Nelson+2019, Pillepich+2019

I identify merging and nonmerging galaxies in TNG50

There are ~300 merging galaxies for z=1

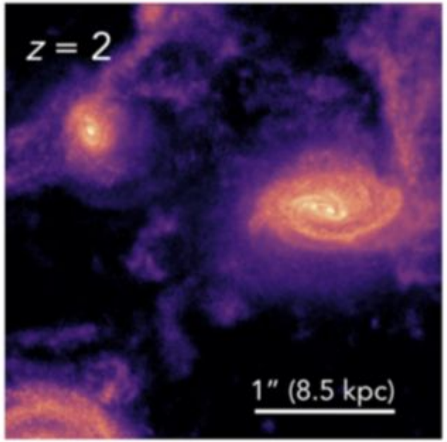




Gas density

To create realistic mock images, we run SKIRT radiative transfer on the full sample of mergers and non-mergers

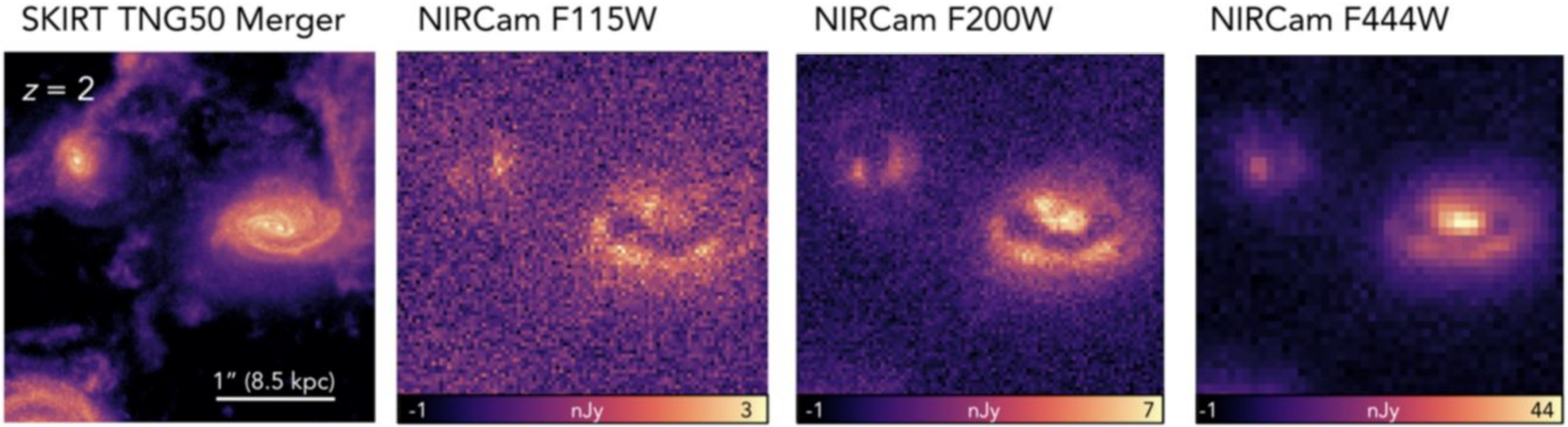
SKIRT TNG50 Merger



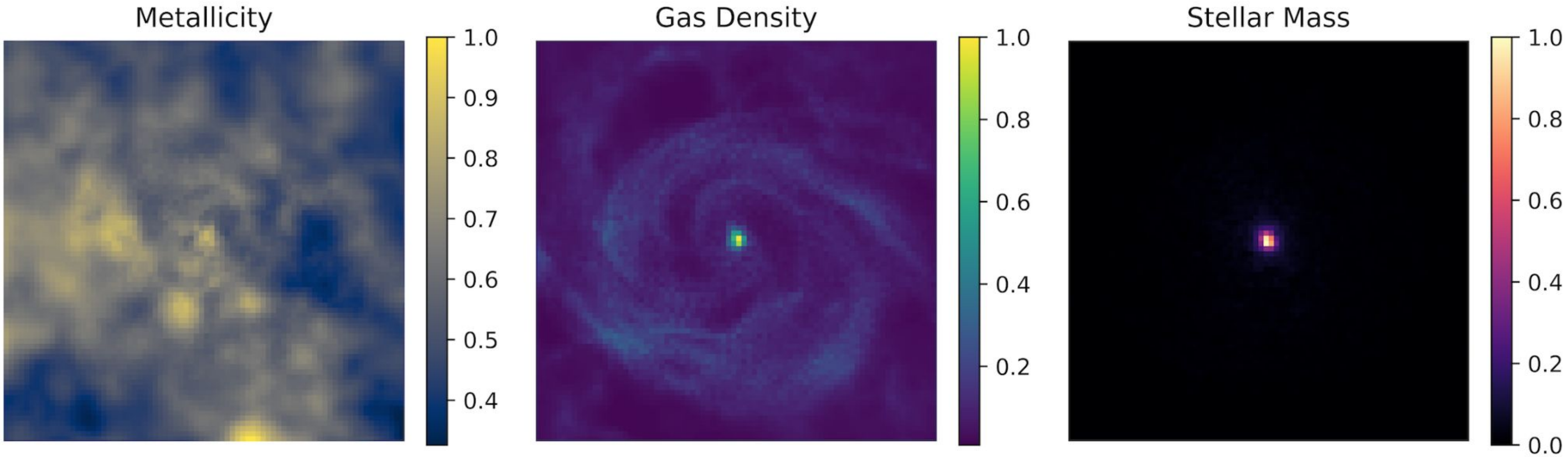
Jacob Shen

The final step is to create observationally realistic images by introducing noise + background sources

Aimee Schechter

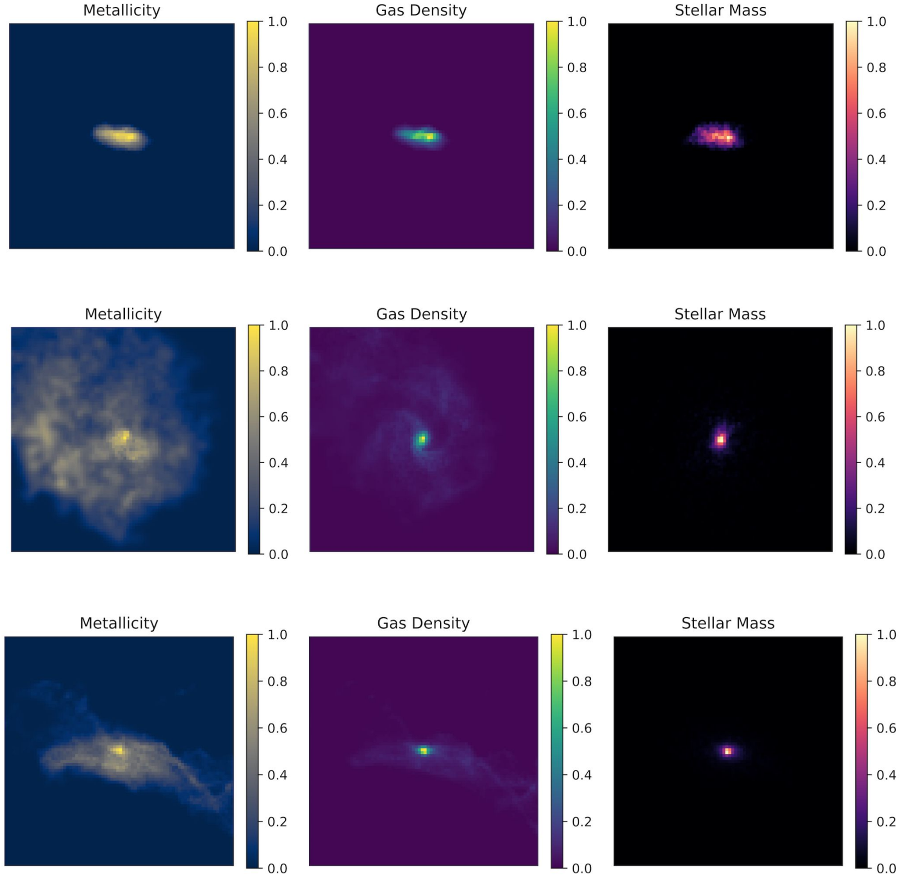


But, radiative transfer takes too long, so we use *yt* to create particle images

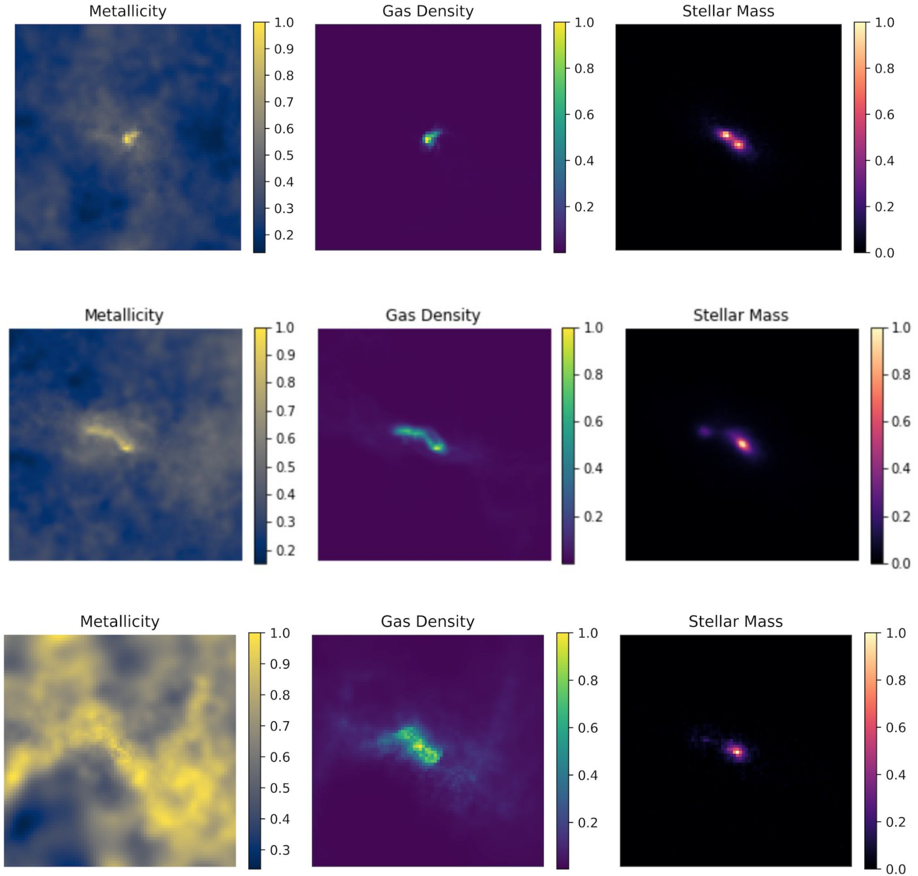


20 kpc width

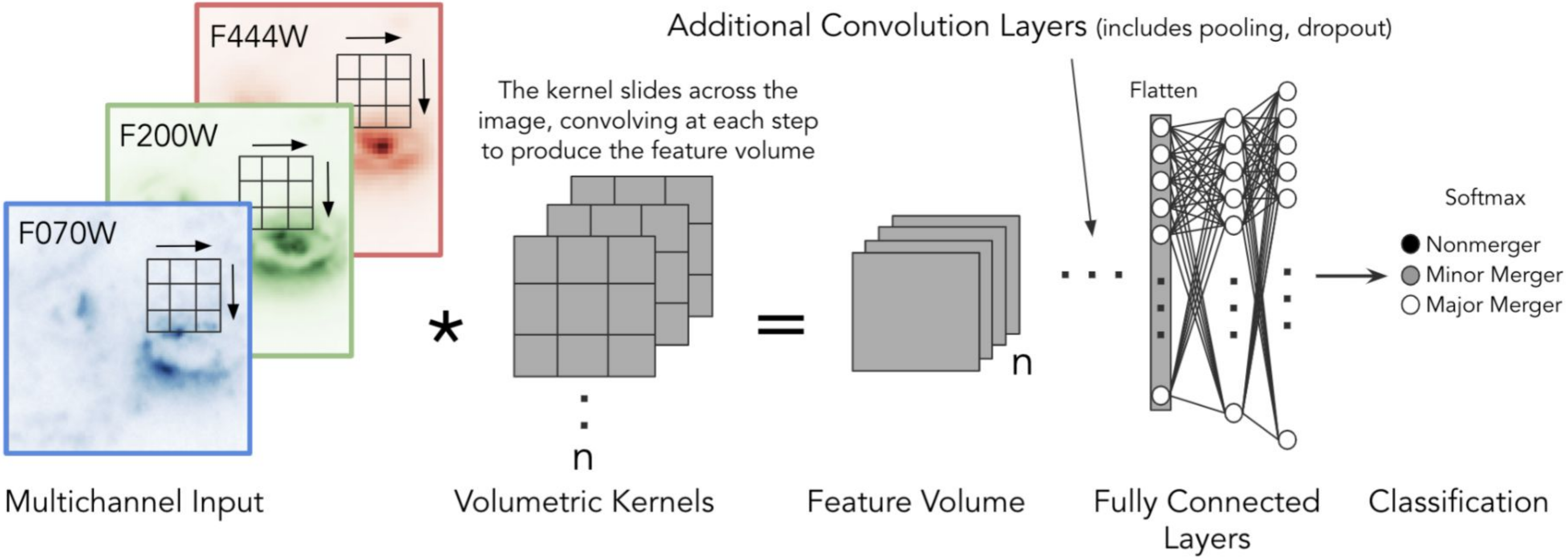
Non-mergers



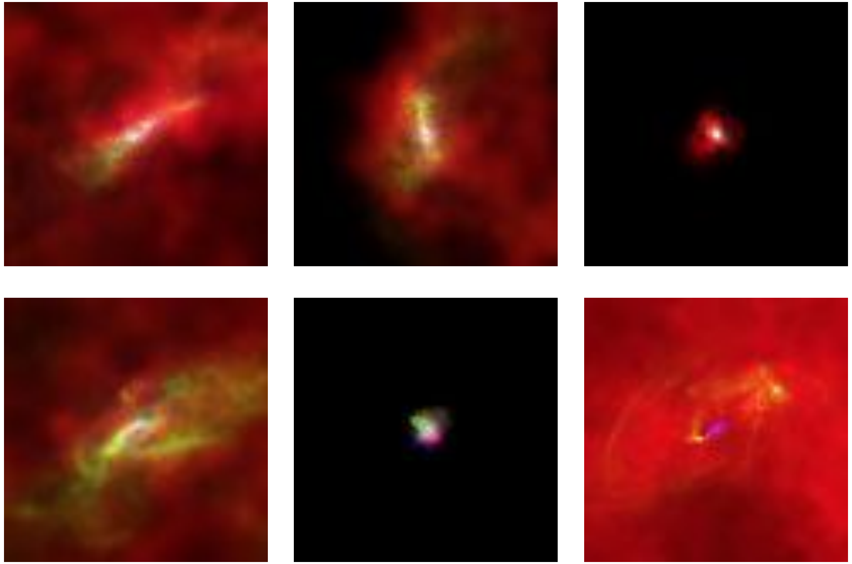
Mergers (pre, current, post)



Convolutional Neural Network design

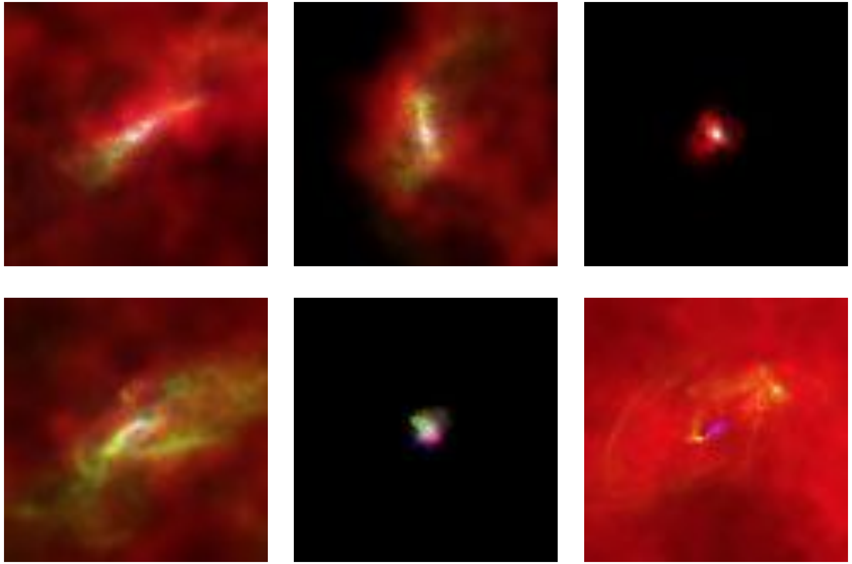


Transfer learning is an exciting option



Options: TNG100 (8 times the volume)

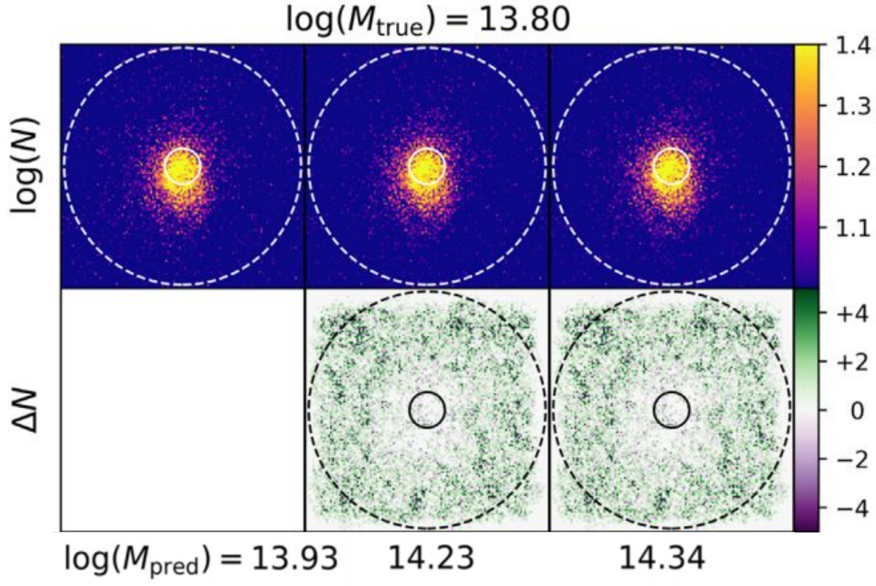
Transfer learning is an exciting option



Options: TNG100 (8 times the volume) or dogs and cats!!

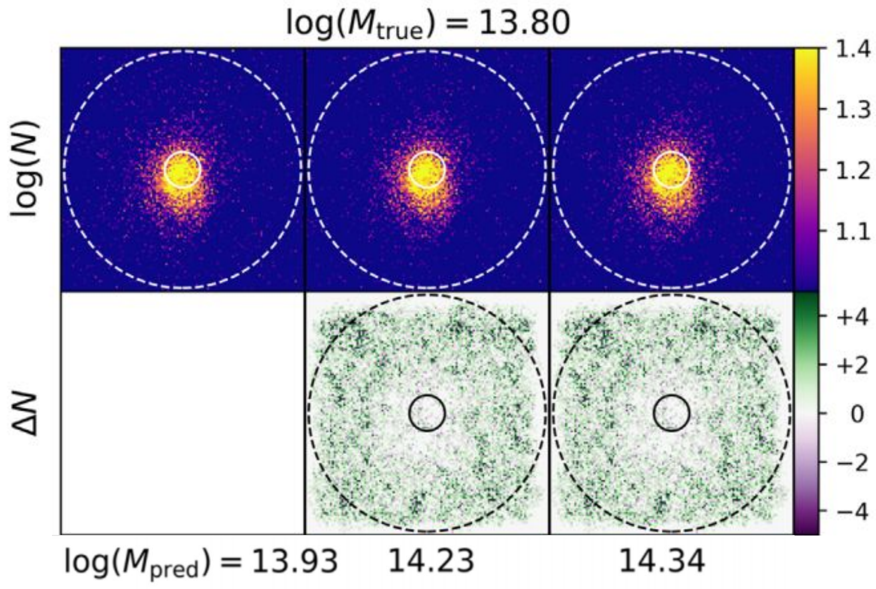
How do we untangle the CNN's decisions?

Saliency methods - e.g., Ntampaka+2018 use Google DeepDream to compute the gradient of the output



How do we untangle the CNN's decisions?

Saliency methods - e.g., Ntampaka+2018 use Google DeepDream to compute the gradient of the output



However, saliency maps can be misleading (Adebayo+2018)

TCAVs: Testing with concept activation vectors

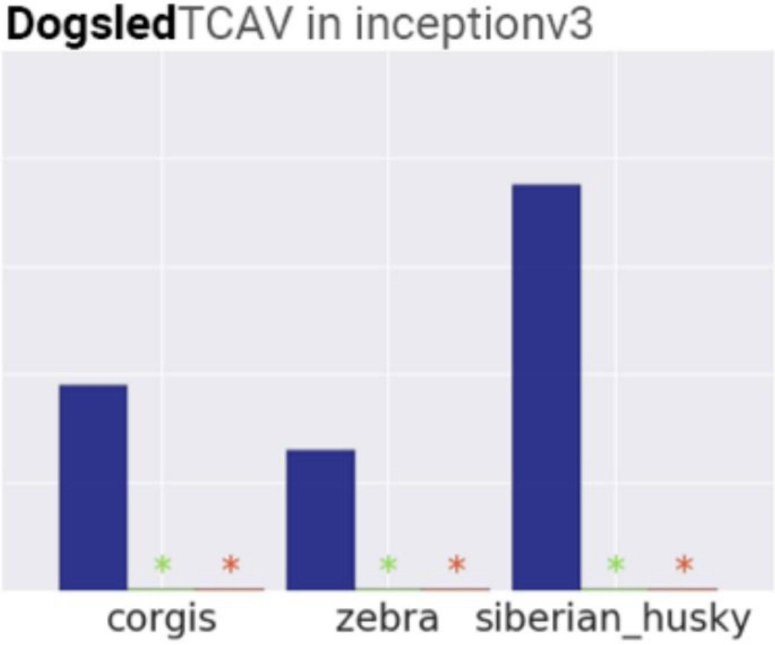
"[After the fact,] CAVs are learned by training a linear classifier to distinguish between the activations produced by a concept's examples and examples in any layer"

Interpretability beyond feature attribution: Kim+2018 https://arxiv.org/pdf/1711.11279.pdf, also https://www.youtube.com/watch?v=Ff-Dx79QEEY&ab_channel=MLconf

TCAVs: Testing with concept activation vectors

Interpretability beyond feature attribution: Kim+2018 https://arxiv.org/pdf/1711.11279.pdf,
also https://www.youtube.com/watch?v=Ff-Dx79QEEY&ab_channel=MLconf

TCAVs: Testing with concept activation vectors



Ideas for galaxy-based CNNs:

- 'Gas-rich' concept
- 'Disky' concept
- 'Busy field' concept

Team 'Fake it till you make it'
A smorgasbord of mocks from Illustris TNG50

HTST NIRCam

Becky Nevin

HST CANDELs

Aimee Schechter

SKIRT9 + AGN

Jacob Shen

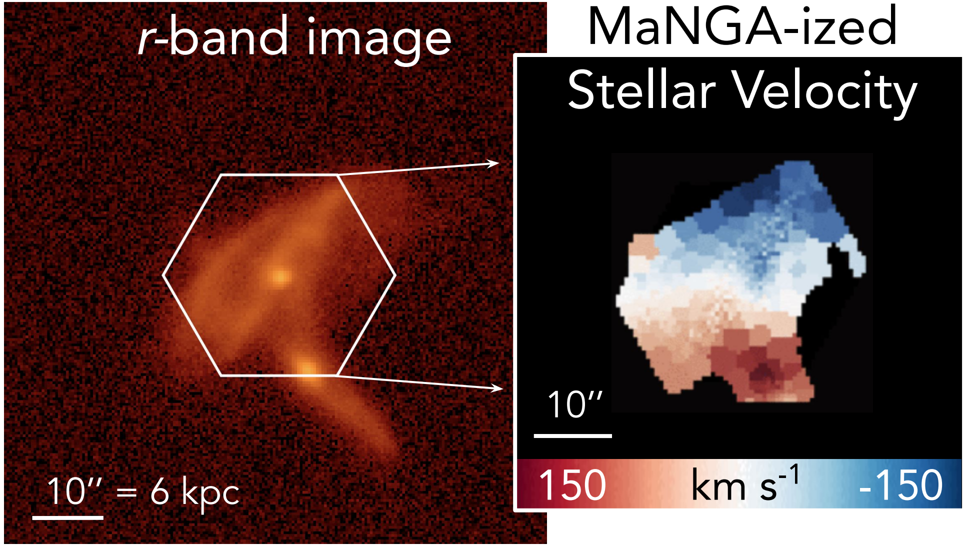
HSC-Joint,
MaNGA, SAMI, HECTOR

Connor Bottrell

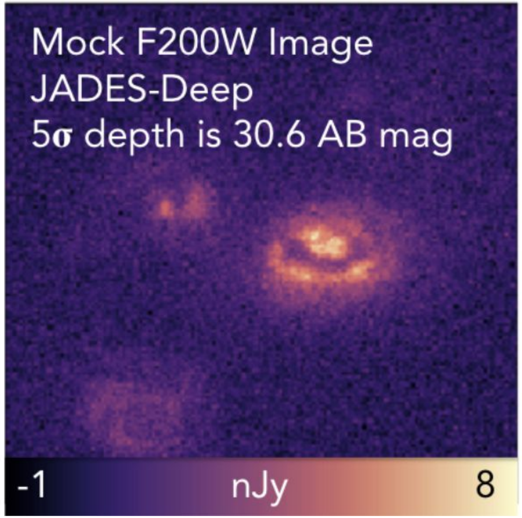
How do we best identify merging galaxies?:
Expanding the toolkit to include stellar kinematics and *HTST* NIRCam imaging

SDSS-ized

MaNGA-ized



HTST-ized



Becky Nevin | beckynevin.github.io

CENTER FOR
ASTROPHYSICS
HARVARD & SMITHSONIAN