JADES • JWST ADVANCED DEEP EXTRAGALACTIC SURVEY

UNIVERSITY OF CAMBRIDGE
Cavendish Laboratory

William M. Baker

Using machine learning and Bayesian inference in galaxy evolution

Kavli Institute for Cosmology Cambridge
KICC

Outline

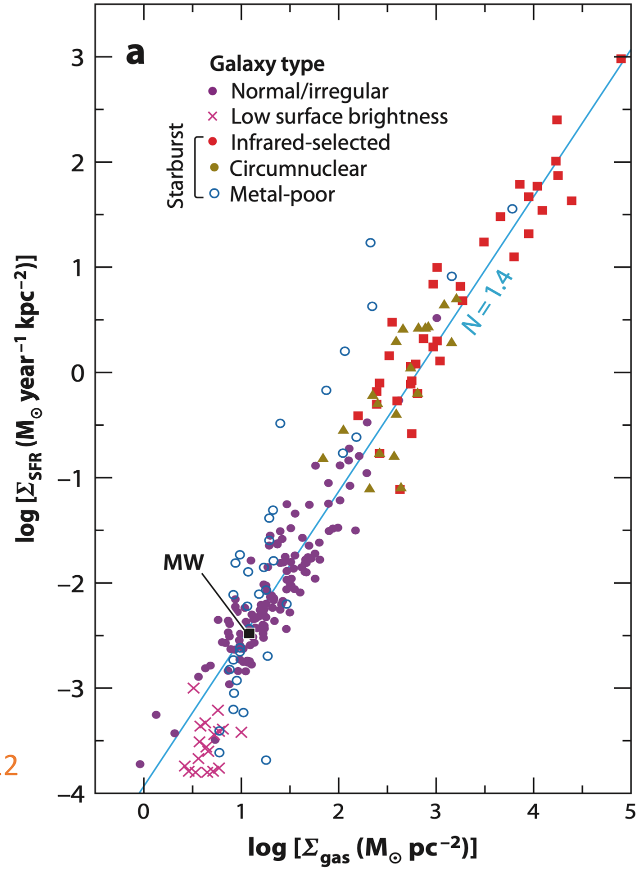
1. Disentangling indirect byproducts from intrinsic relations with partial correlation coefficients and random forest regression

2. Understanding the origin of the mass-metallicity relation

3. Forward modelling light distributions with ForcePho to determine fluxes

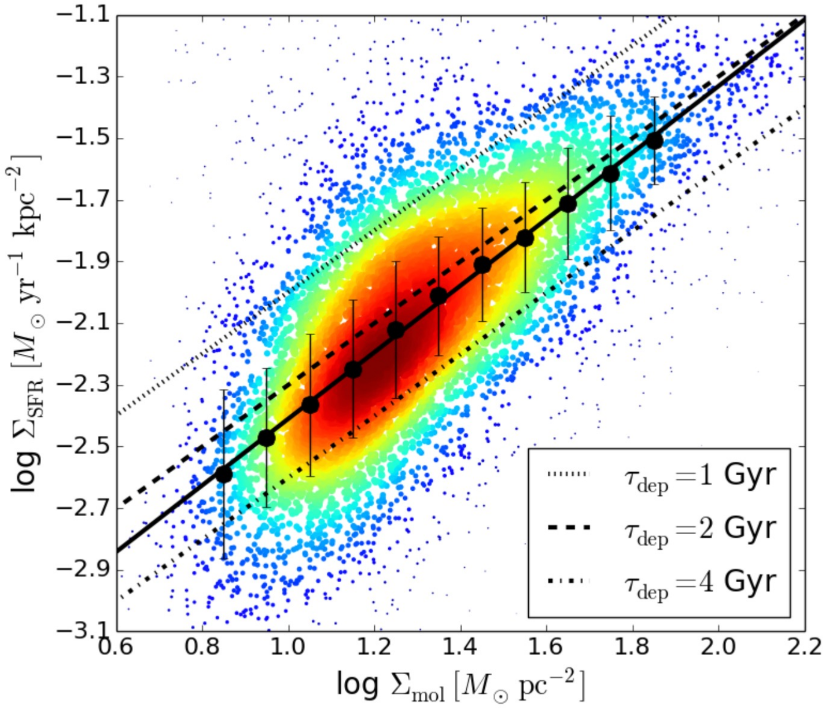
4. Fitting SEDs with Prospector to infer physical quantities

The Star Forming Schmidt-Kennicutt (SK) relation



Kennicutt & Evans 2012

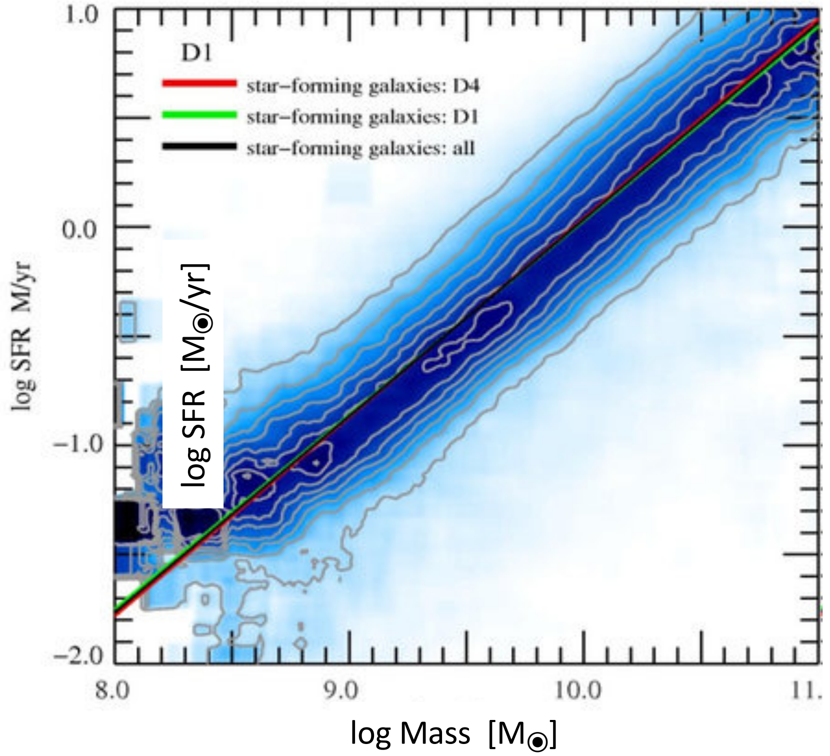
Integrated version

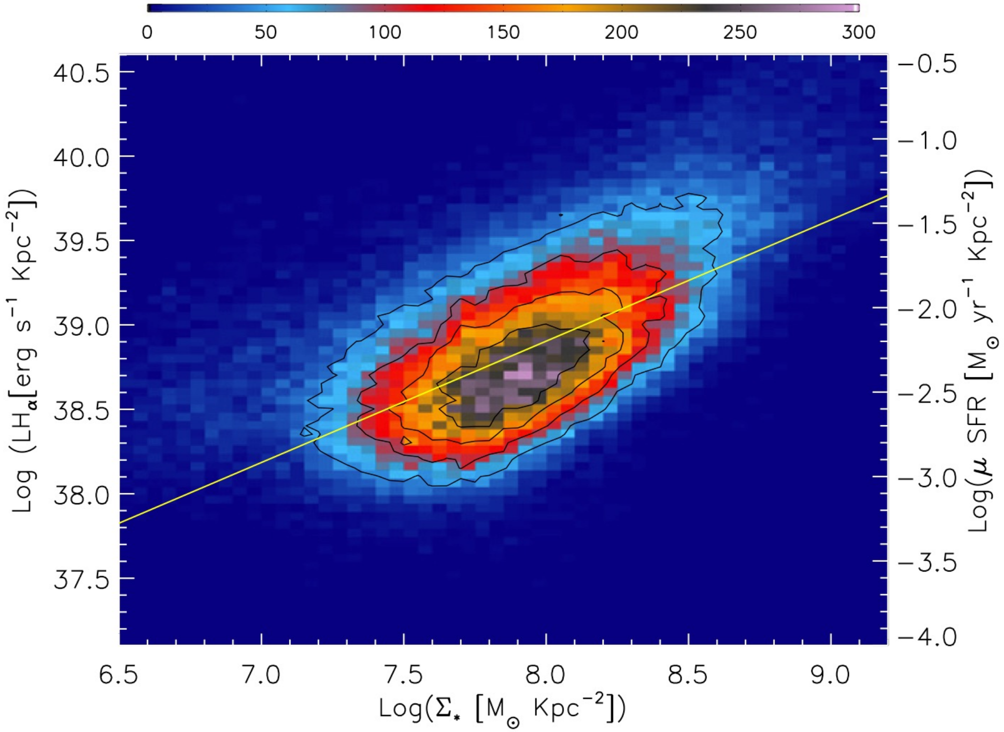


Utomo 2017

Resolved version

The Star Forming 'Main Sequence' (SFMS)





Cano Diaz+16

Spatially resolved version

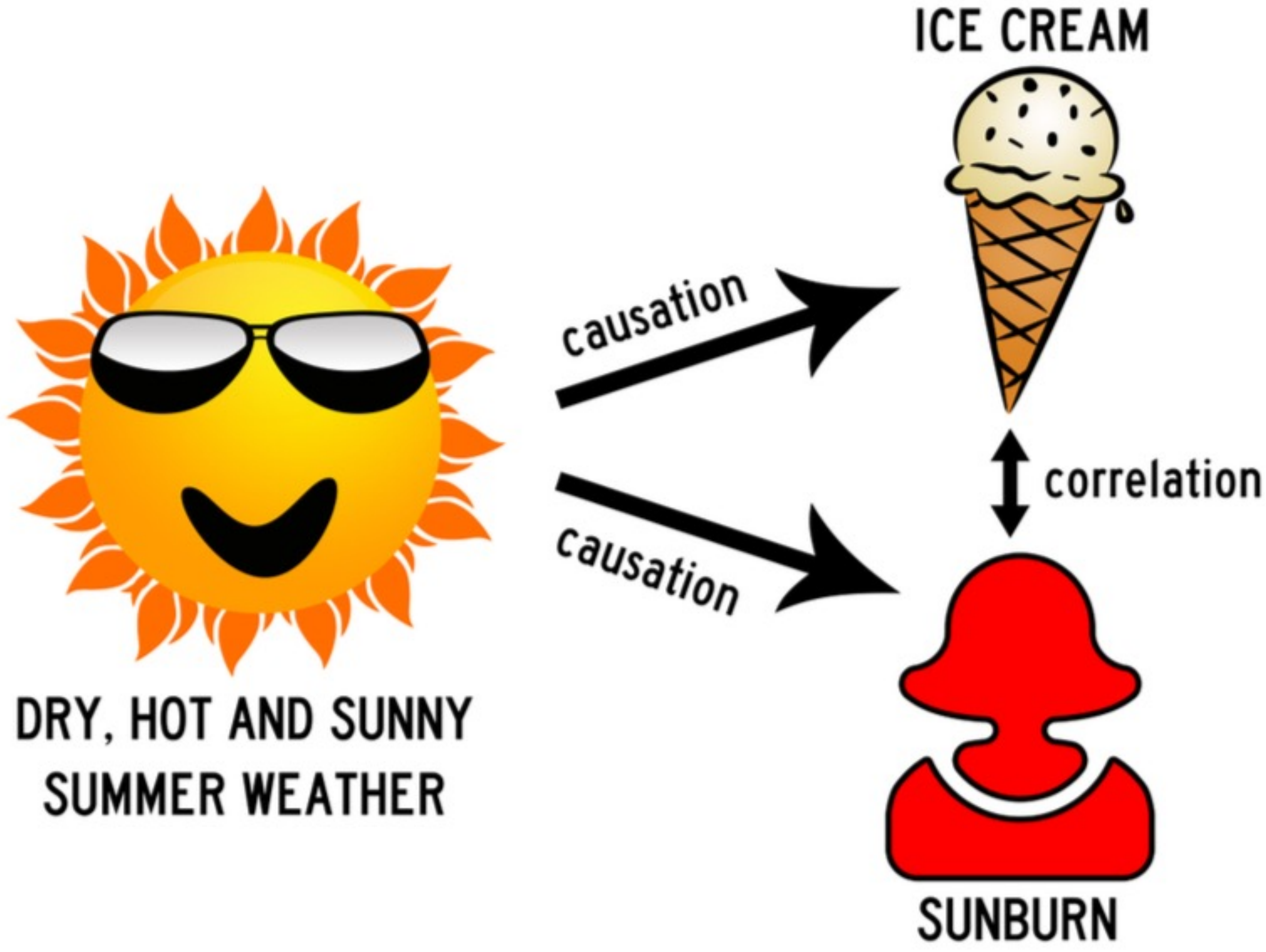
Integrated, global version

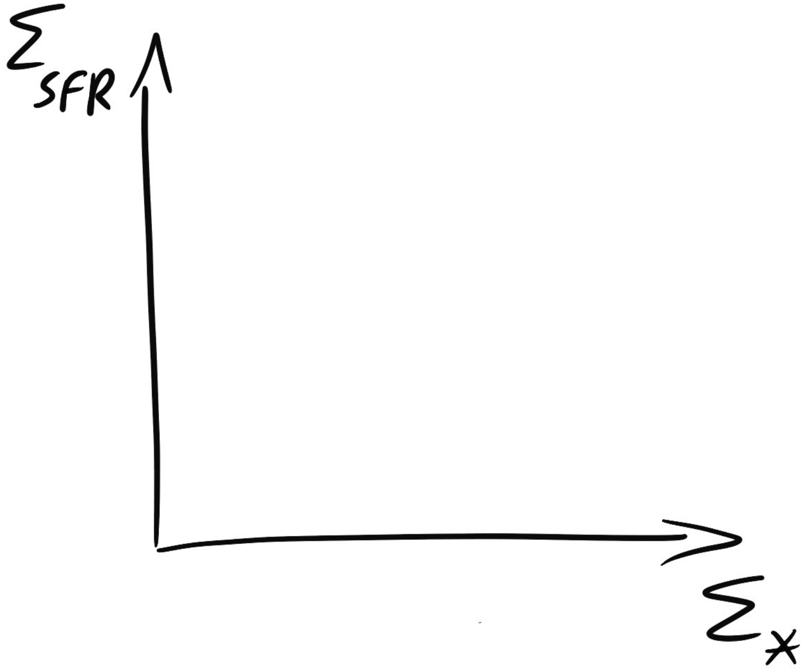
Peng+ 2010

e.g Brinchmann+01, Peng+10, Renzini & Peng'15,....

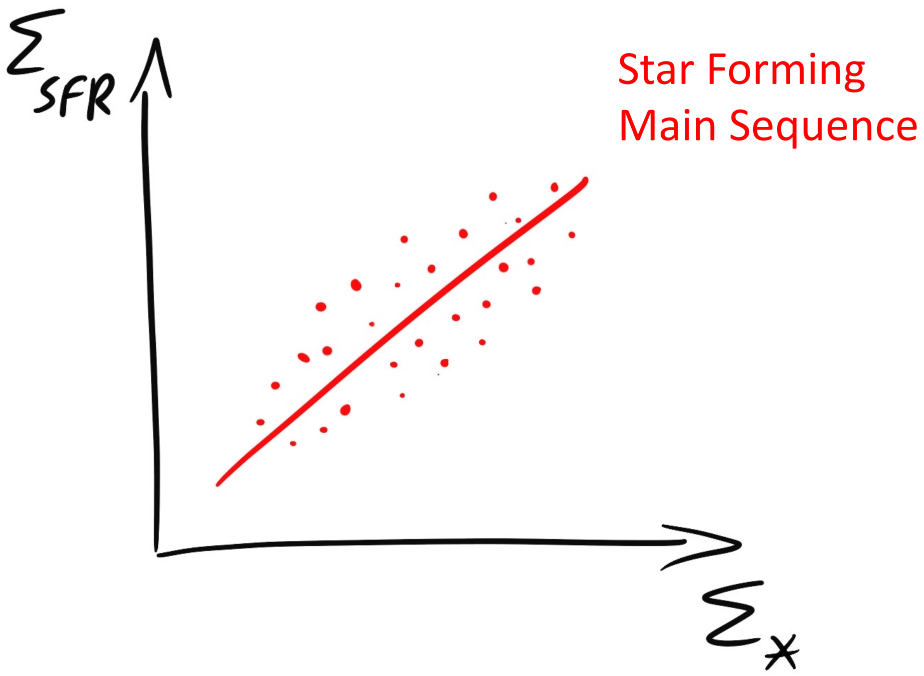
e.g. Cano Diaz+16, Wuyts+13, Akiyama+18, Hsieh+17, Baker+2022

Correlation does not imply causation

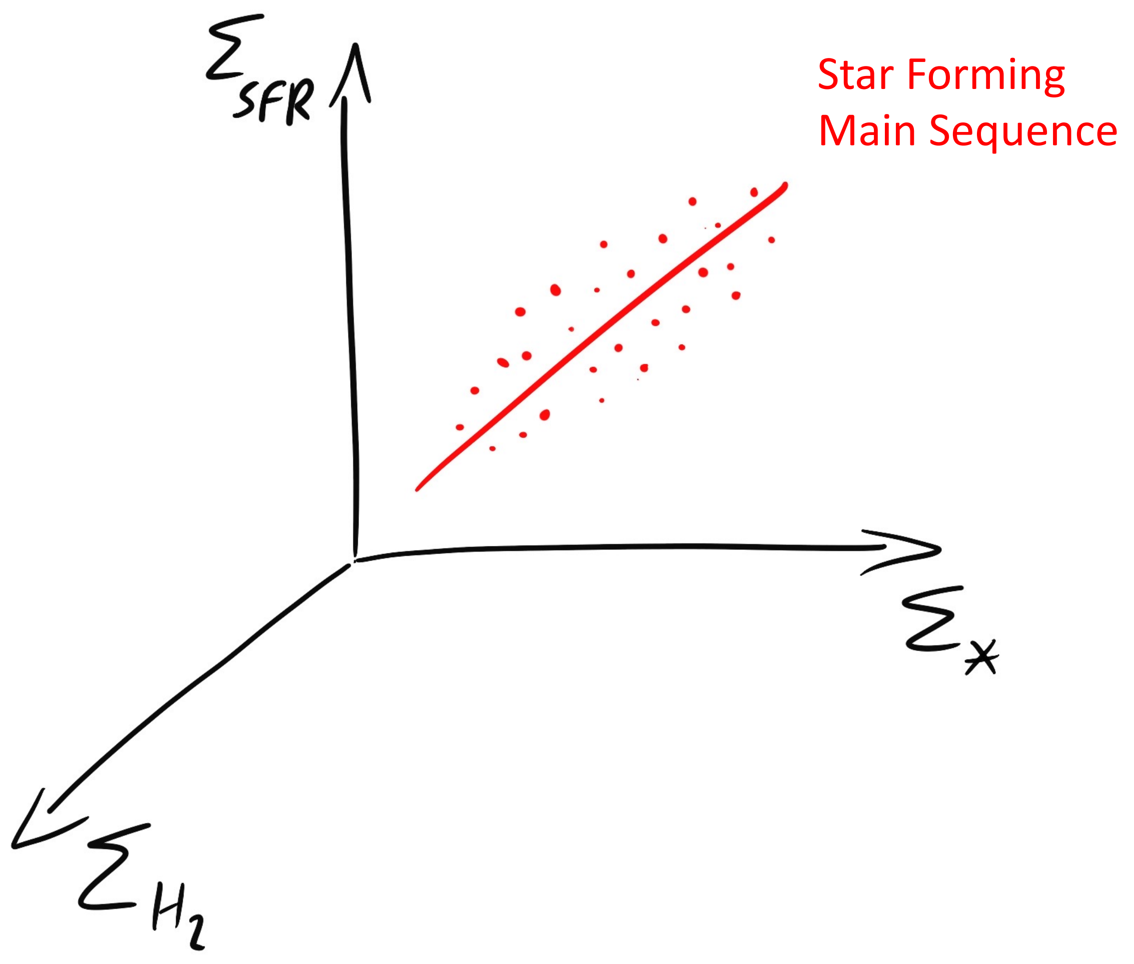




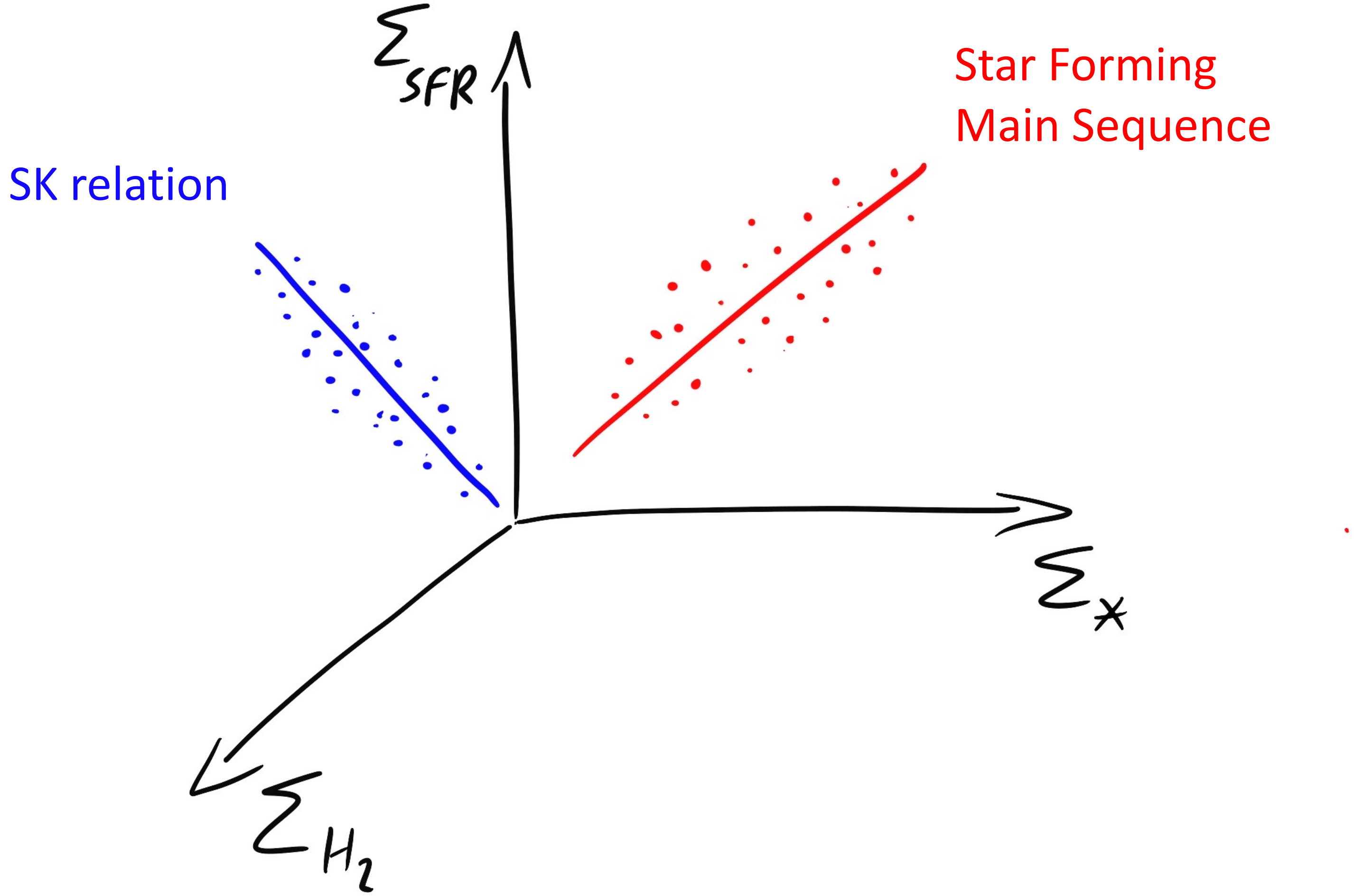
Σ_{SFR}
Σ_{*}



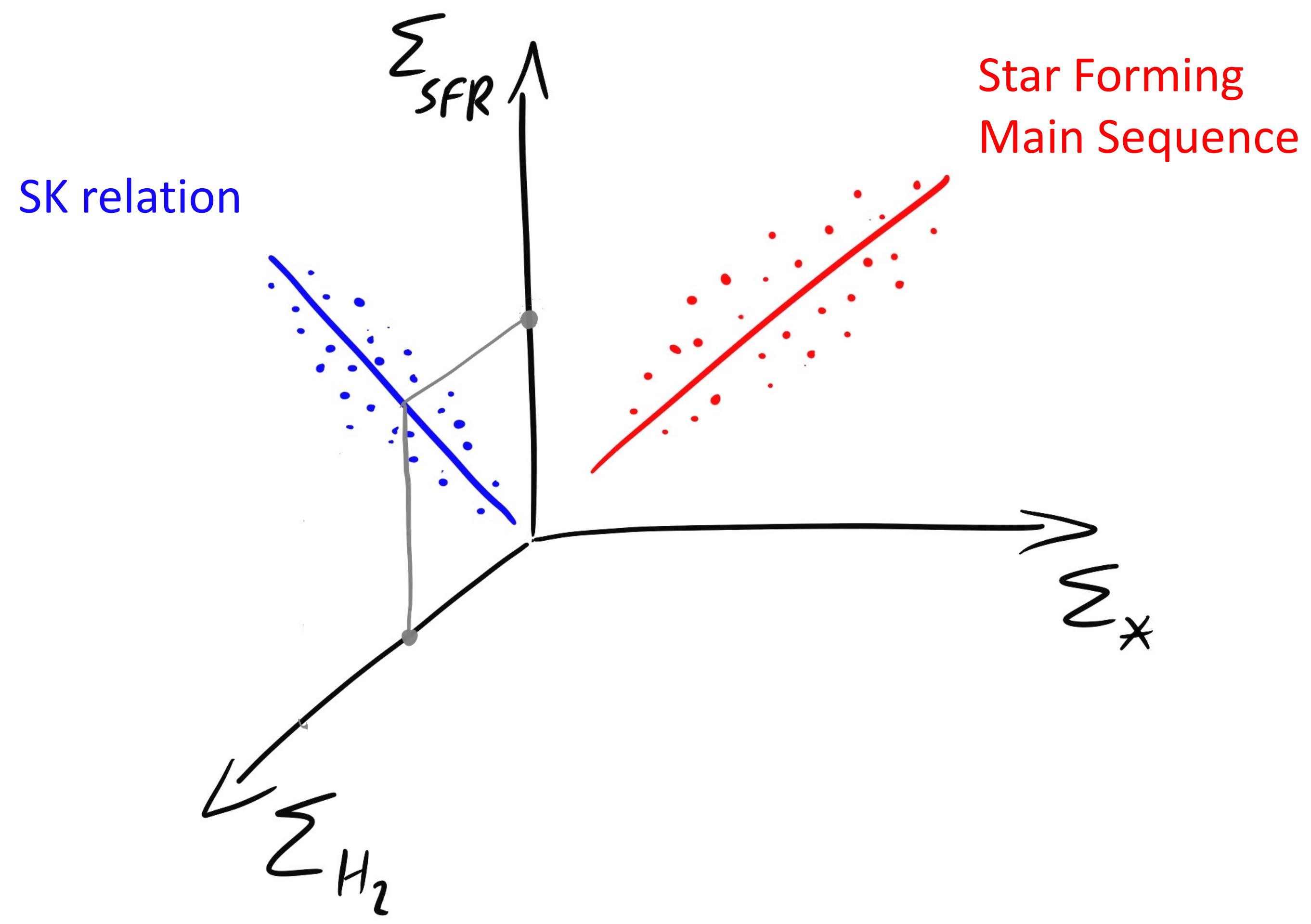
Σ_{SFR}
Star Forming
Main Sequence
Σ_{*}



Σ_{SFR}
Star Forming
Main Sequence
Σ_*
Σ_{H_2}

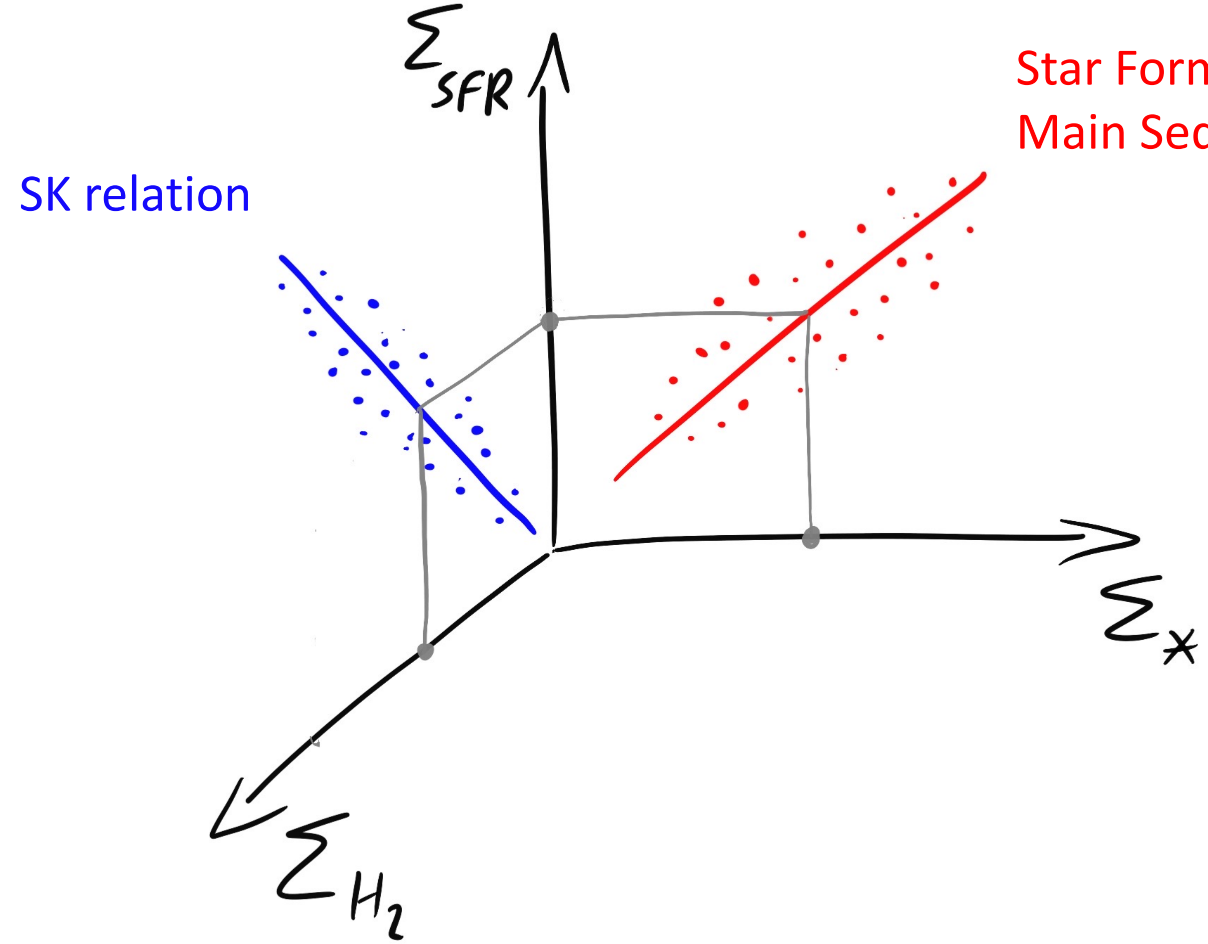


Σ_{SFR}
SK relation
Star Forming
Main Sequence
Σ_*
Σ_{H_2}

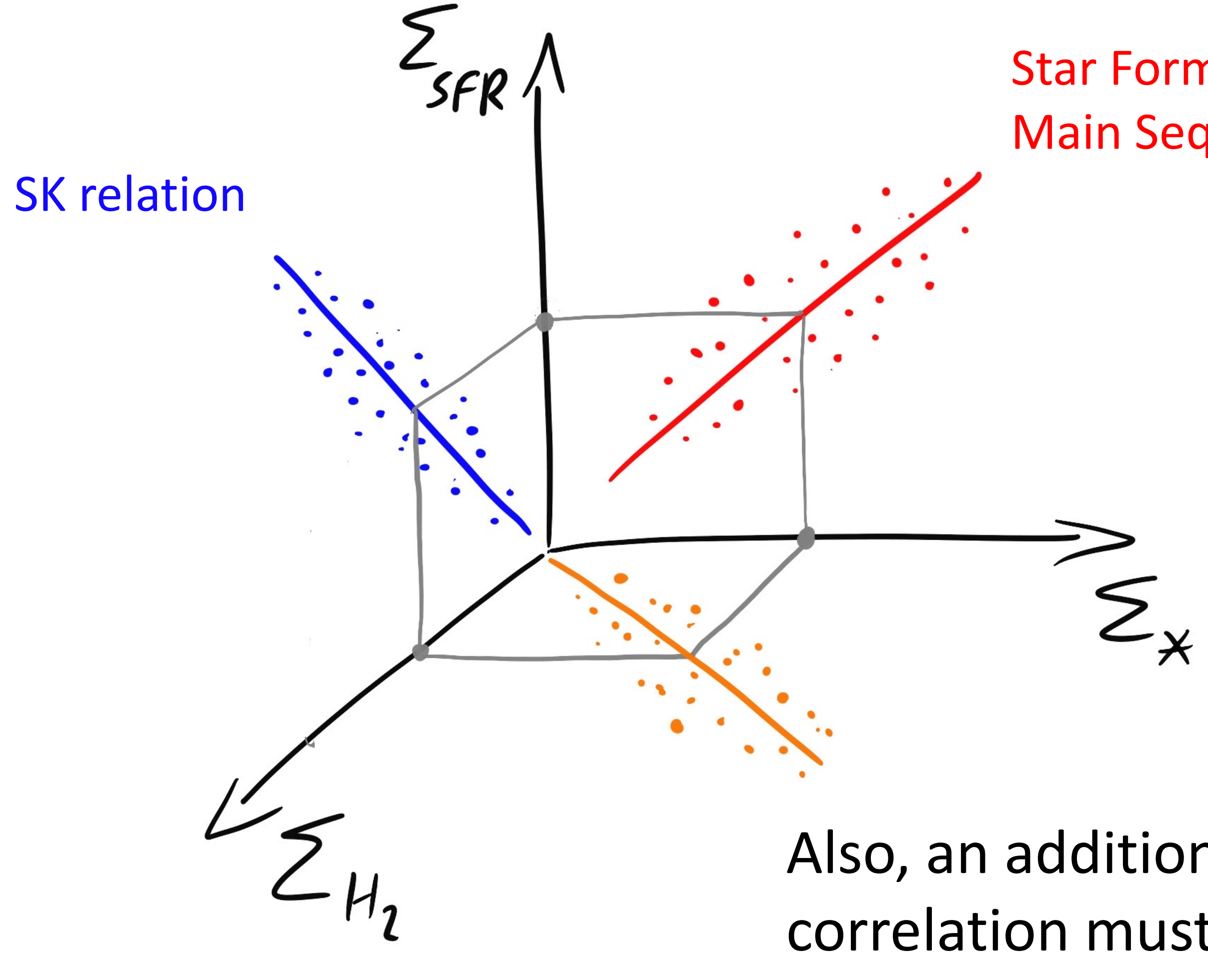


Σ_{SFR}
SK relation
Star Forming
Main Sequence
$\Sigma_{\star}$
Σ_{H_2}

Is it Σ_* or Σ_{H2} driving Σ_{SFR}?



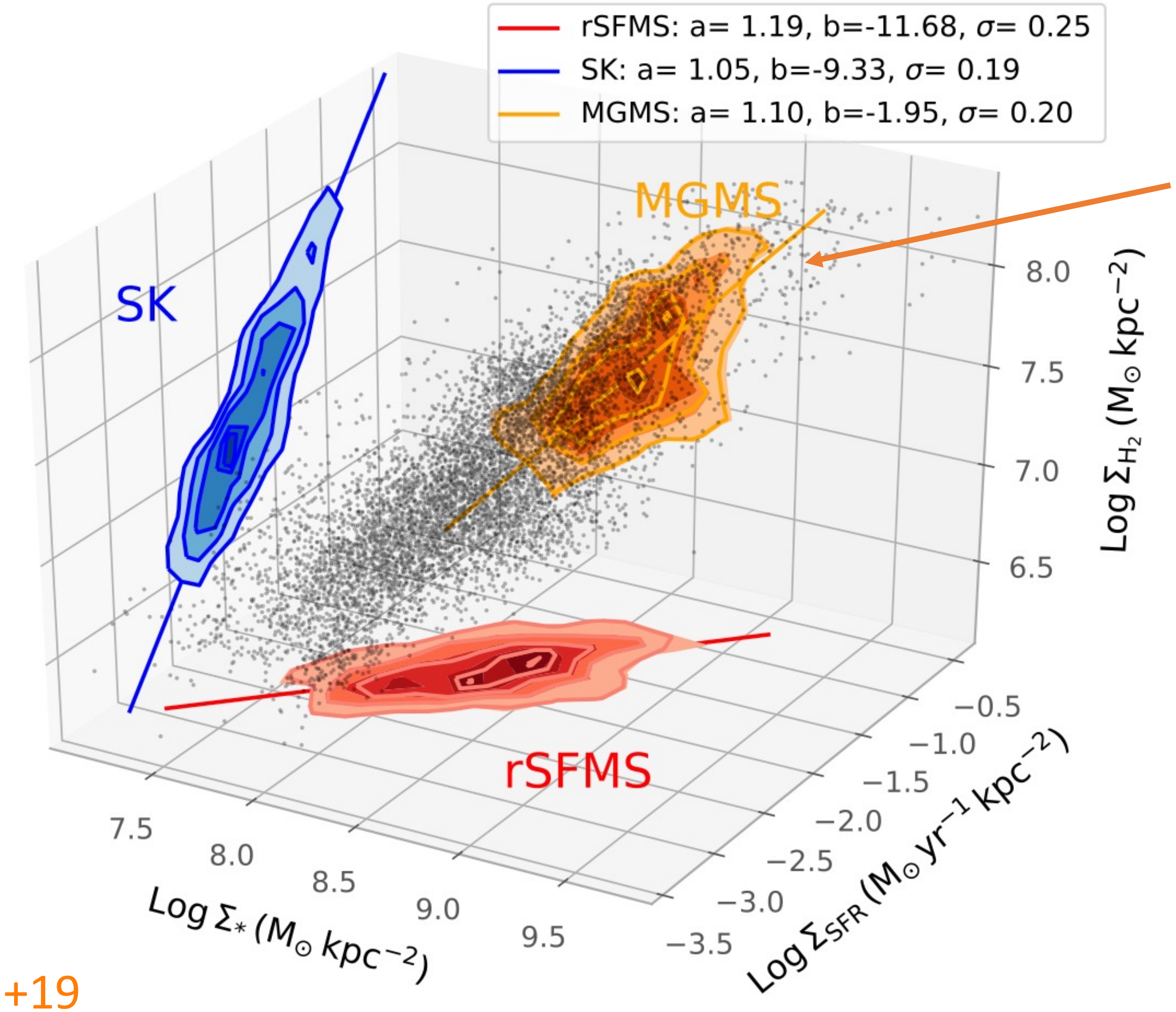
Is it Σ_* or Σ_{H2} driving Σ_{SFR}?



Also, an additional correlation must be present between Σ_* or Σ_{H2}

The ALMaQUEST survey and the 'Main Sequences'

46 galaxies with resolved CO (ALMA) and optical spectroscopy (MaNGA)



Molecular Gas Main Sequence

Lin+19

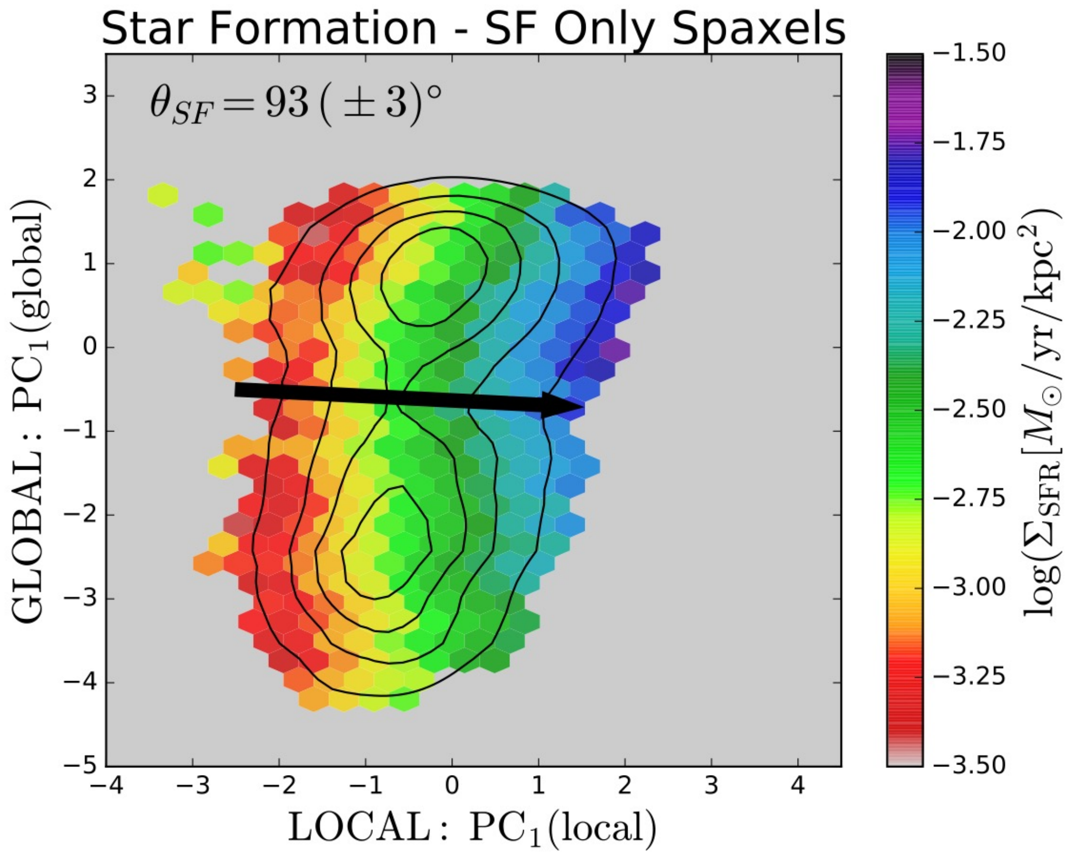
Which of these three relations are primary/fundamental and which are instead a by-product of the others?

Partial Correlation Coefficients

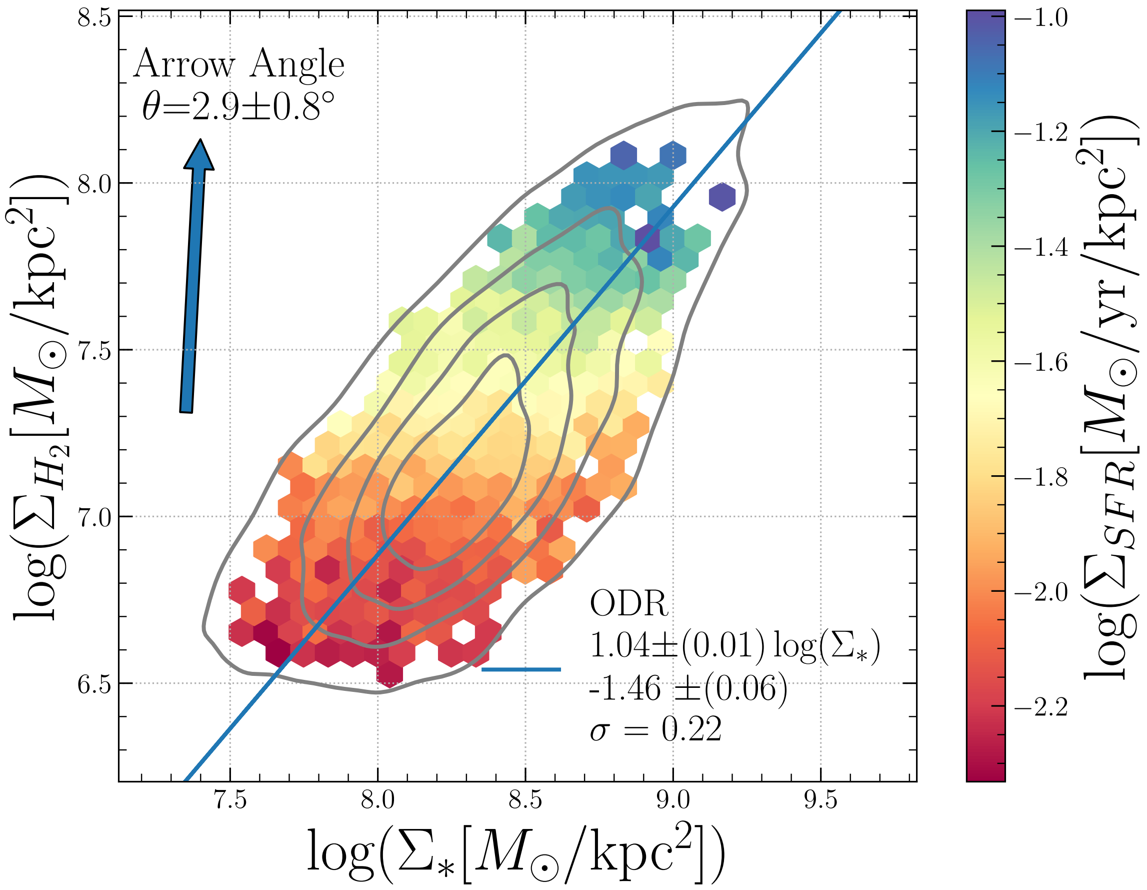
- Partial correlation coefficients give the correlation between two quantities whilst holding further quantities constant
- -> powerful tool to disentangle intrinsic (direct) correlations from indirect correlations which are a by-product of other correlations

$$\tan(\theta) = \frac{\rho_{AC|B}}{\rho_{BC|A}}$$

Bluck+2020



Disentangling direct, intrinsic correlations from those that are induced/by-product requires using/controlling all data simultaneously



Baker+22

Arrow:
direction of average gradient
from Partial Correlation Coefficients

Σ_{SFR} depends strongly on Σ_{H2}

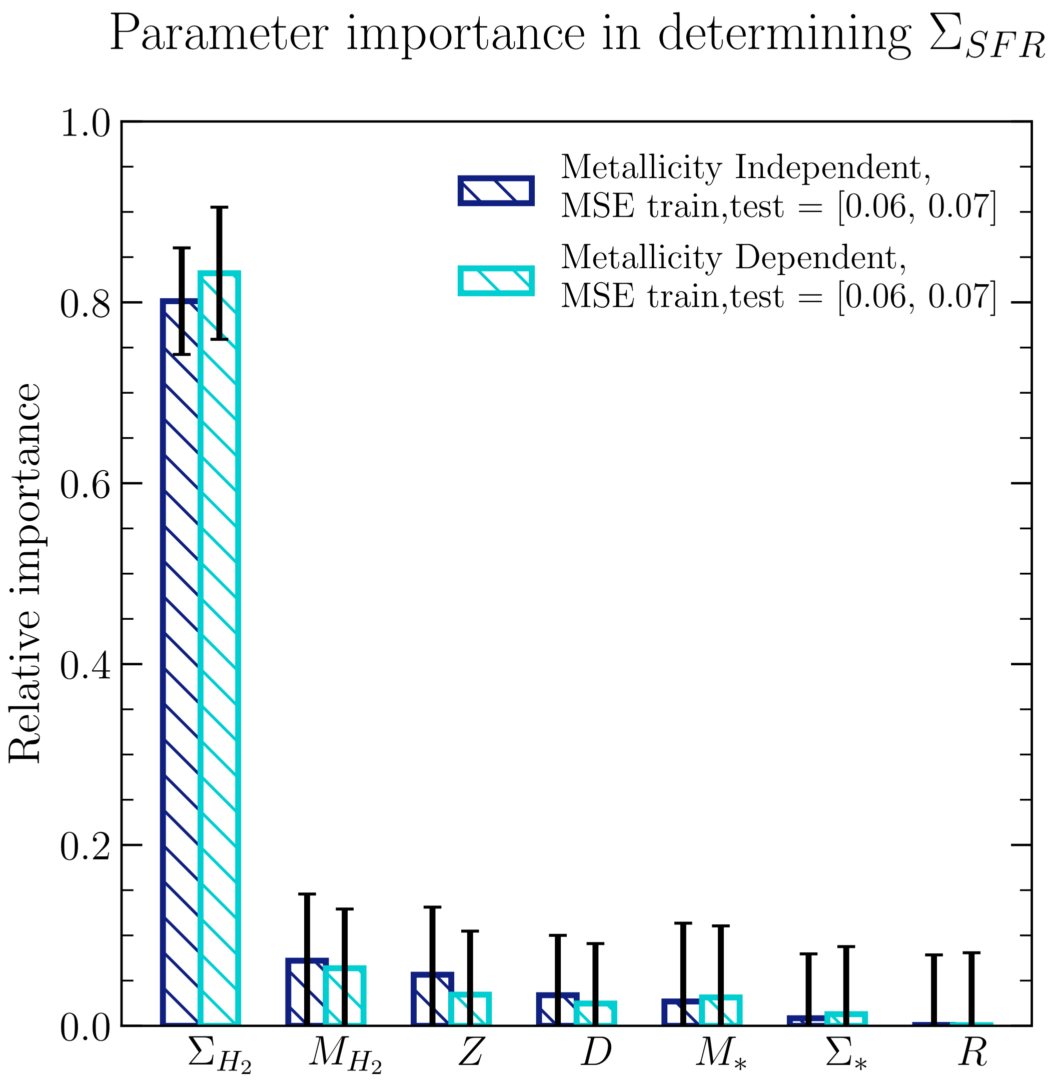
Σ_{SFR} does <u>not</u> depend on Σ_* at a fixed Σ_{H2}

Random Forest Regression

- Form of supervised (i.e. identifiable labels) Machine Learning
- Multiple decision trees – split via Gini Impurity (a measure of the quality of a split)
- Can probe several inter-correlated quantities simultaneously, uncover non-linear relationships, and determine the intrinsic dependence of a quantity (Bluck et al. 2022)
- For further details on Partial correlations or Random Forests see Bluck et al. 2020, 2022, Piotrowska+2022 or Baker+2022, 2023a, 2023b, Baker & Maiolino 2023

Random Forest (Machine Learning) Regression analysis

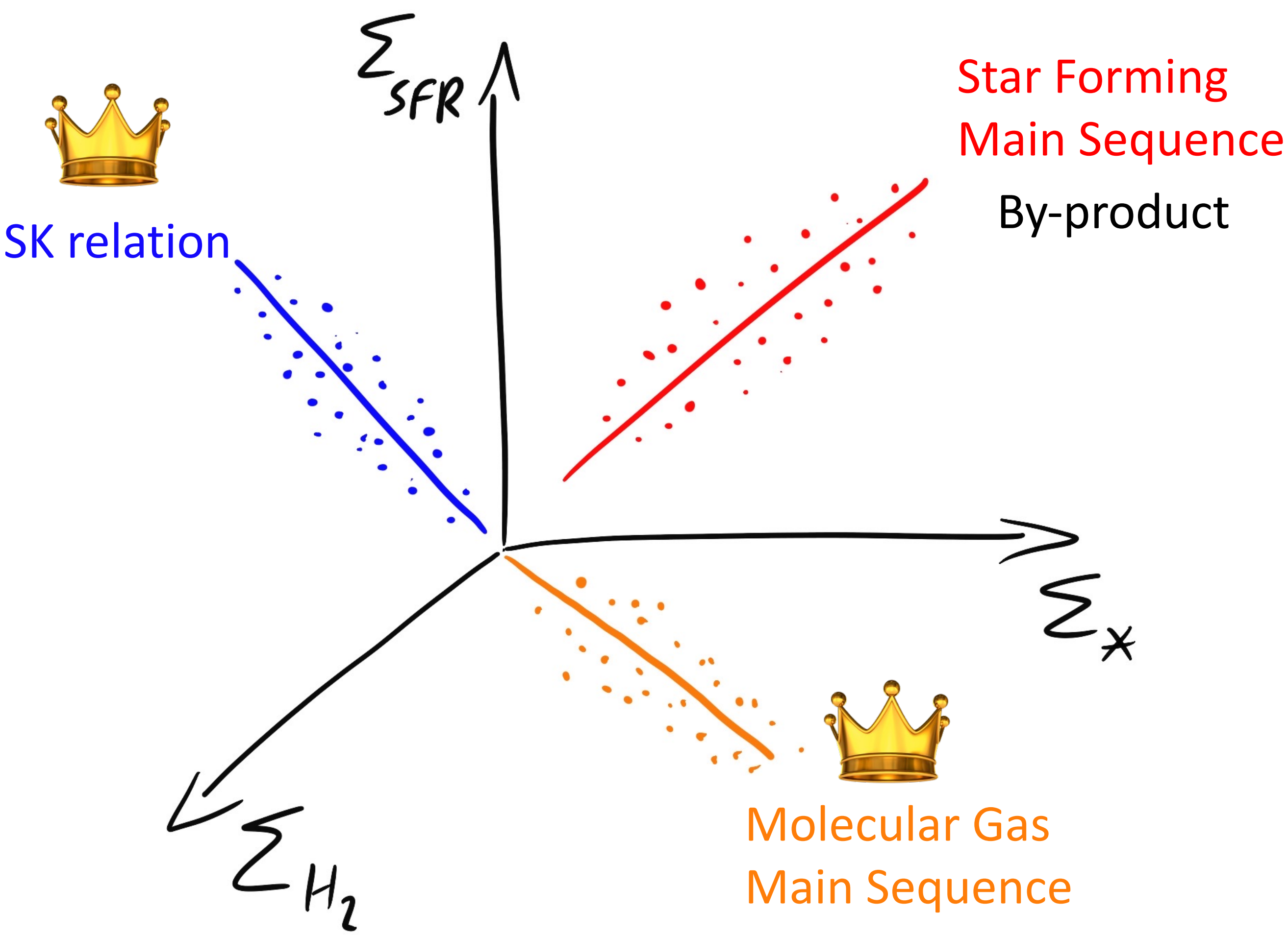
Relative importance of various galactic parameters in **predicting Σ_{SFR}**



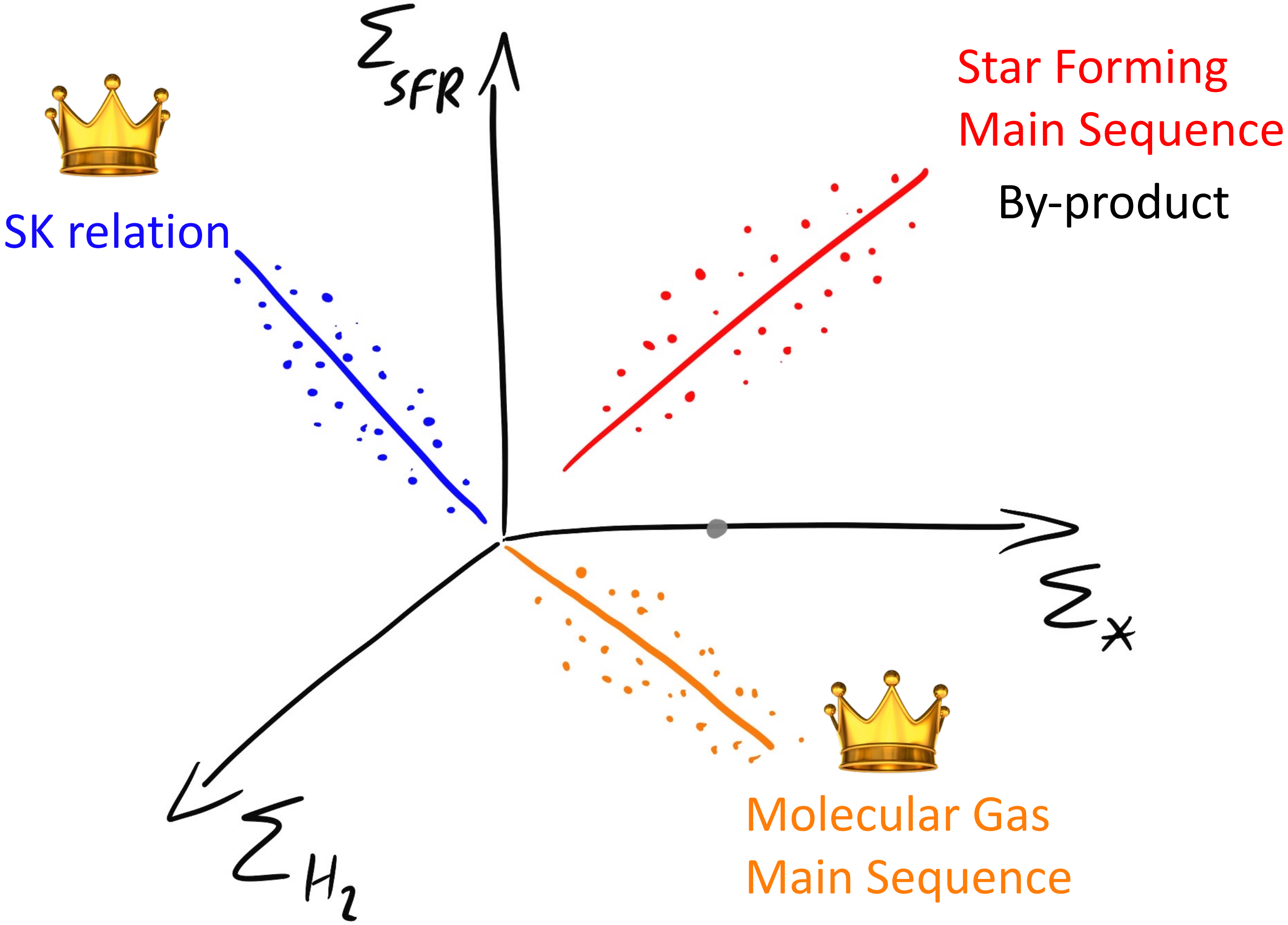
Σ_{H2} (i.e. Schmidt-Kennicutt) unambiguously, by far, the most important

Σ_* totally unimportant once the dependence on Σ_{H2} is taken into account

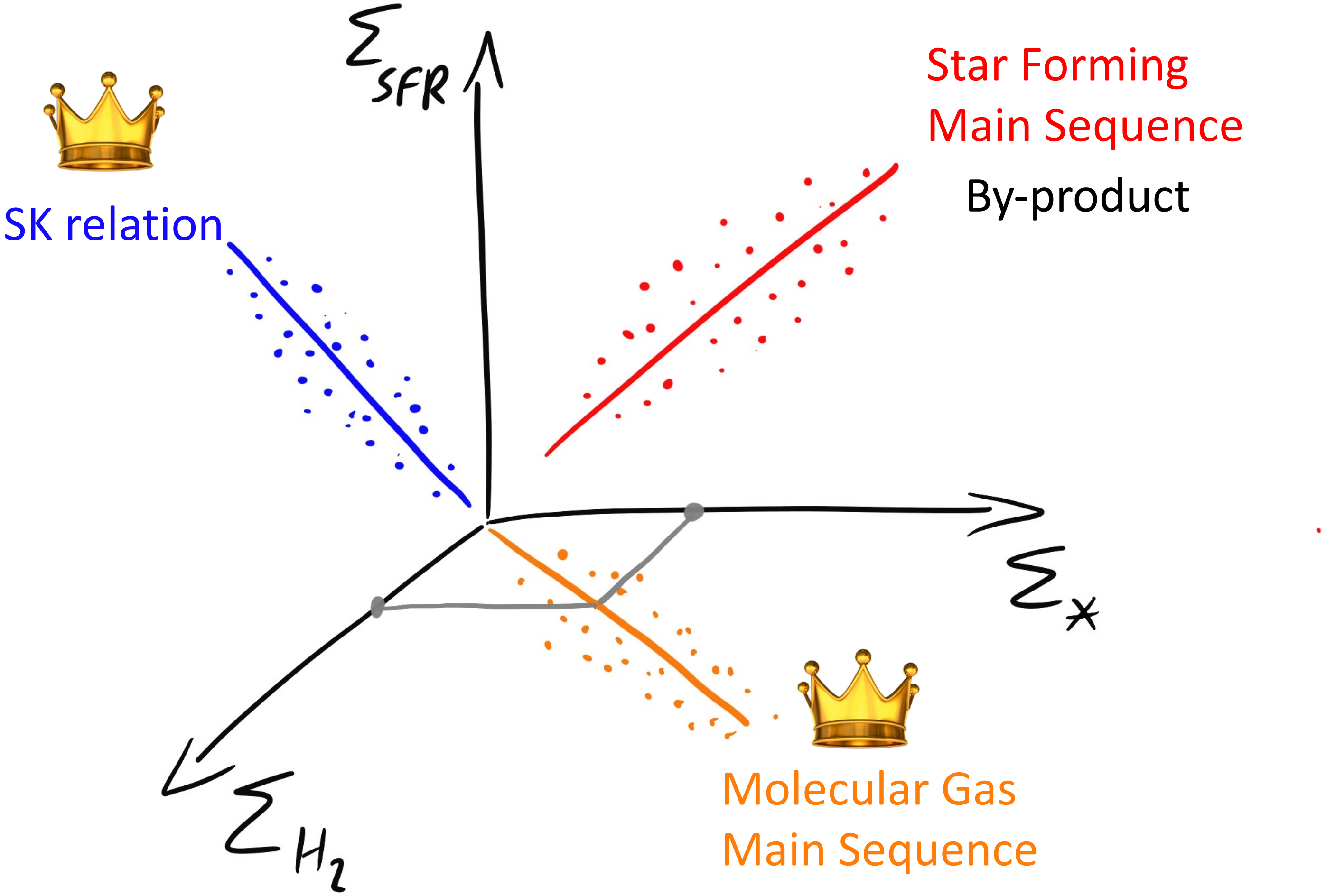
Baker+22



Σ_{SFR}
Star Forming
Main Sequence
By-product
SK relation
Σ_{*}
Σ_{H_2}
Molecular Gas
Main Sequence



Σ_{SFR}
Star Forming
Main Sequence
By-product
SK relation
$\Sigma_{\star}$
Σ_{H_2}
Molecular Gas
Main Sequence

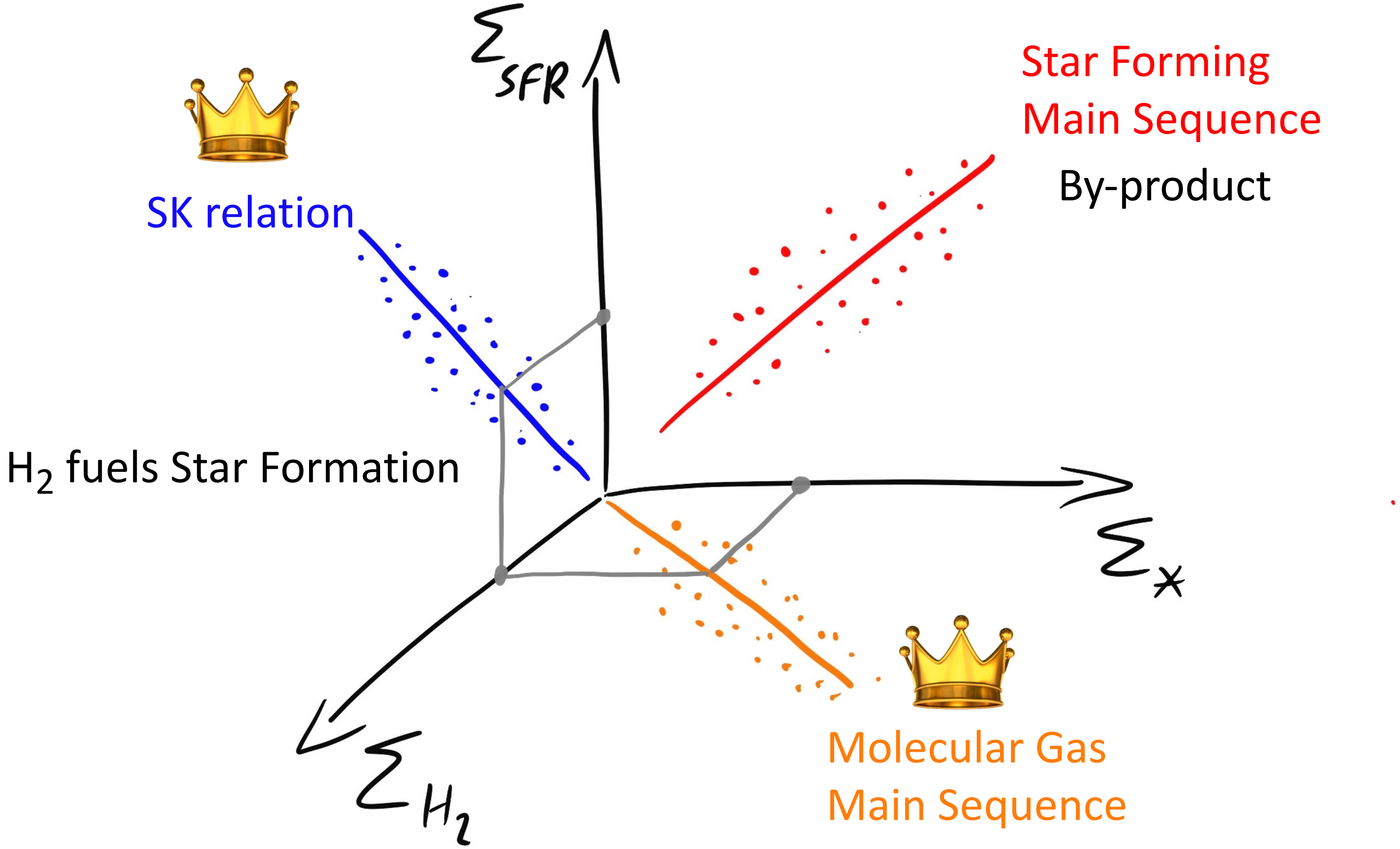


Possible explanations:

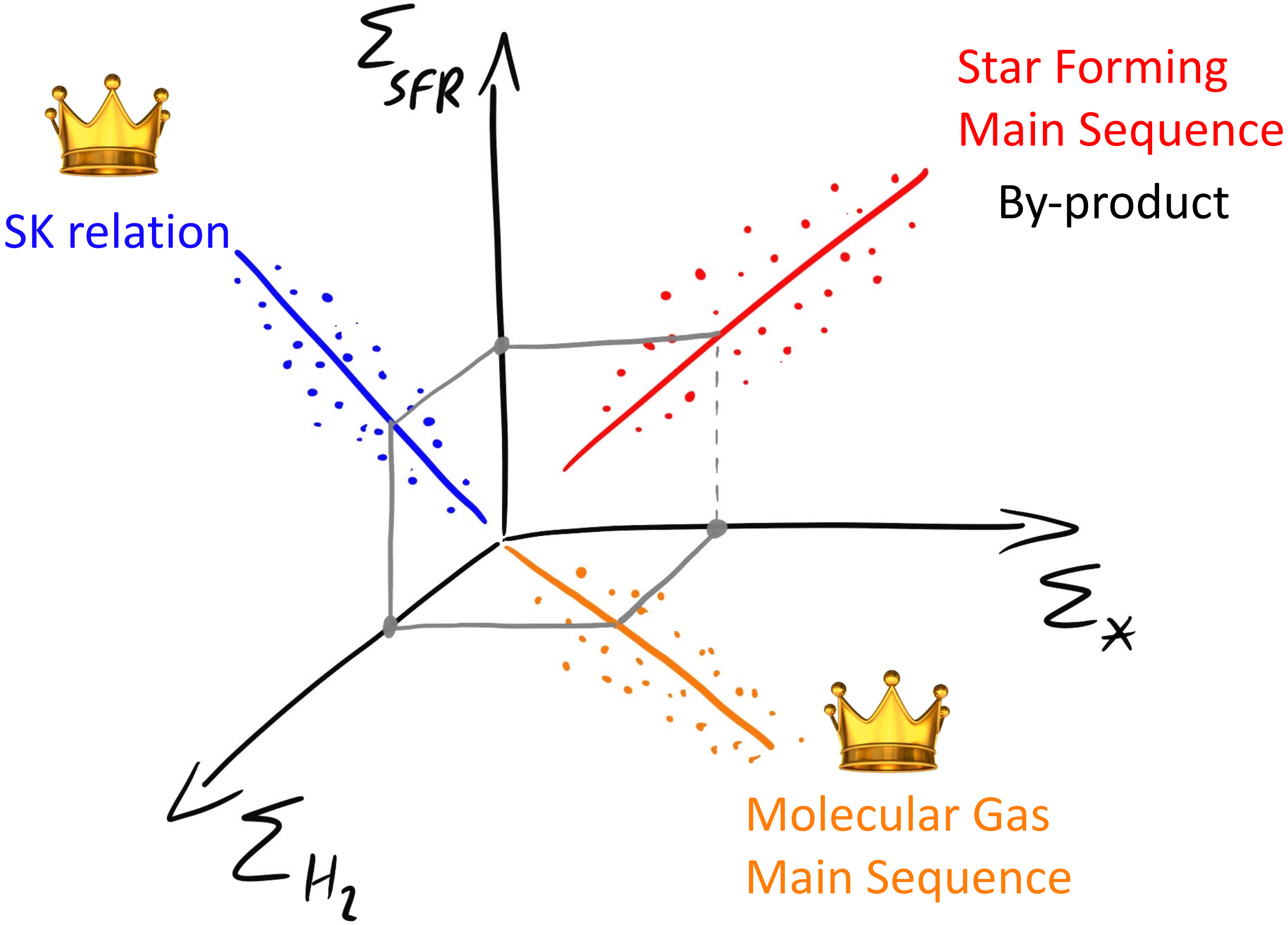
Gas accretion via gravity

HI->H2 conversion

.....

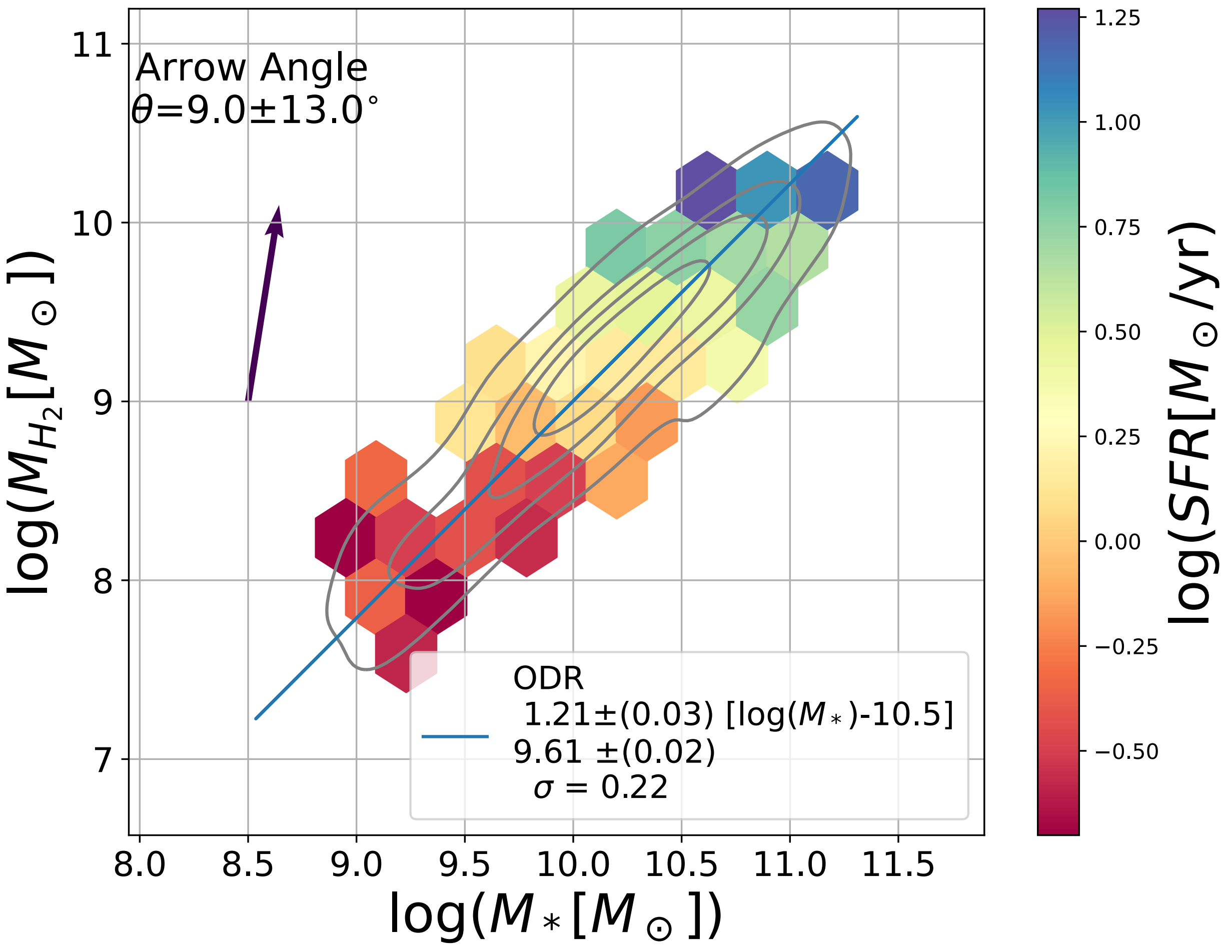


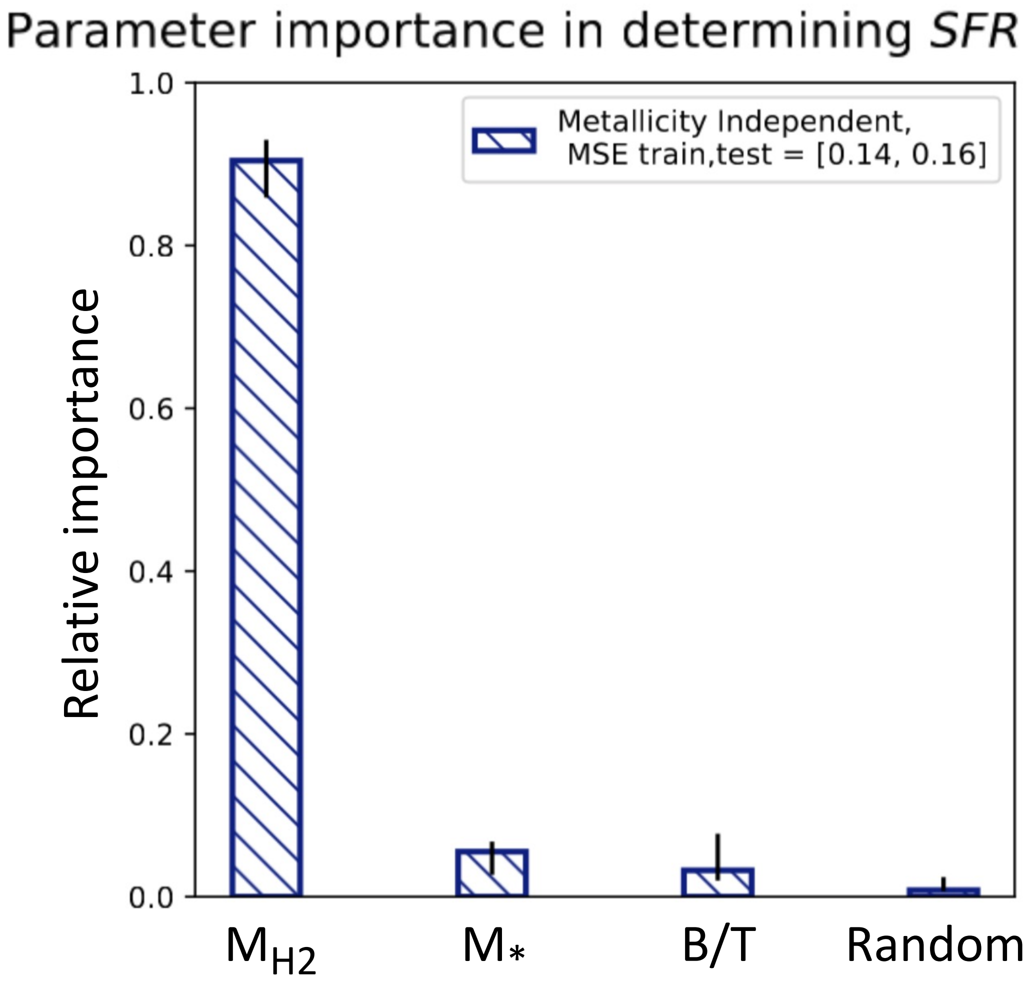
Σ_{SFR}
Star Forming
Main Sequence
By-product
SK relation
H$_2$ fuels Star Formation
Σ_*
Σ_{H_2}
Molecular Gas
Main Sequence



Σ_{SFR}
Star Forming
Main Sequence
By-product
SK relation
Σ_*
Σ_{H_2}
Molecular Gas
Main Sequence

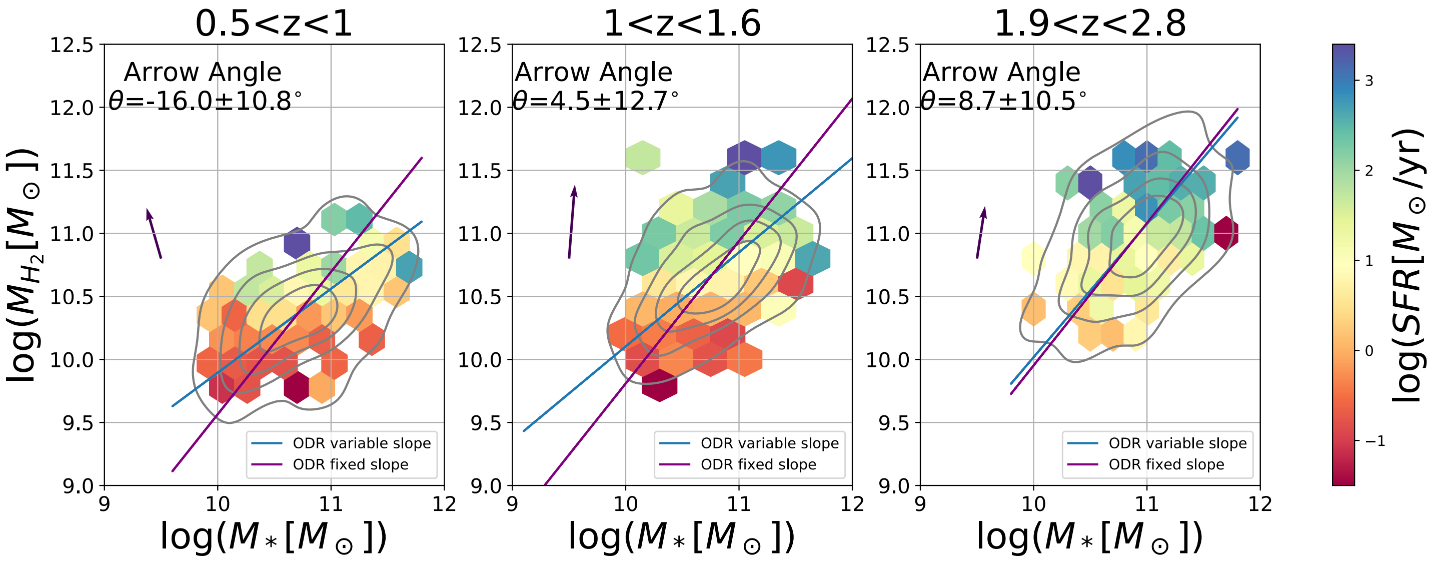
Same finding for the integrated, global quantities – Locally

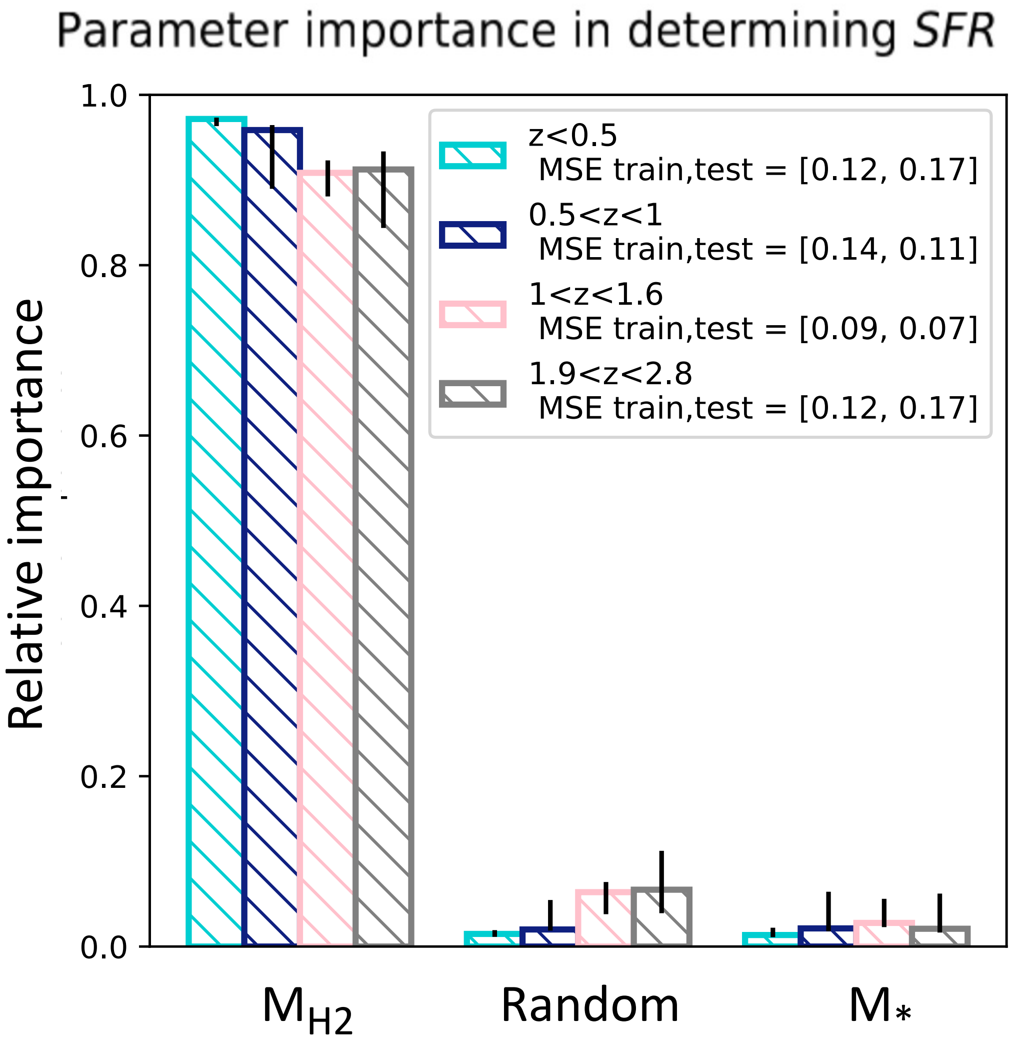




Baker+2023a

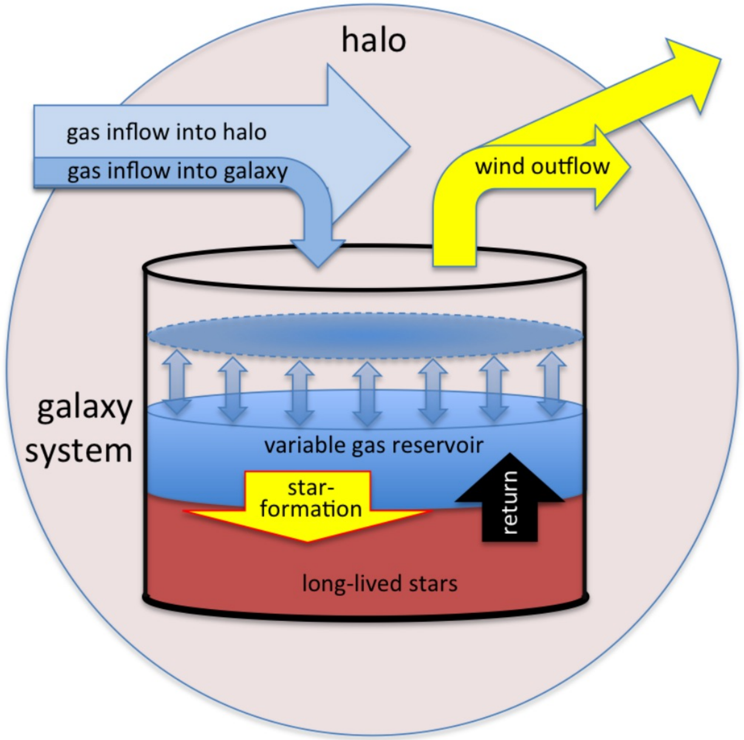
Same finding for the integrated, global quantities – At high-z





Baker+2023a

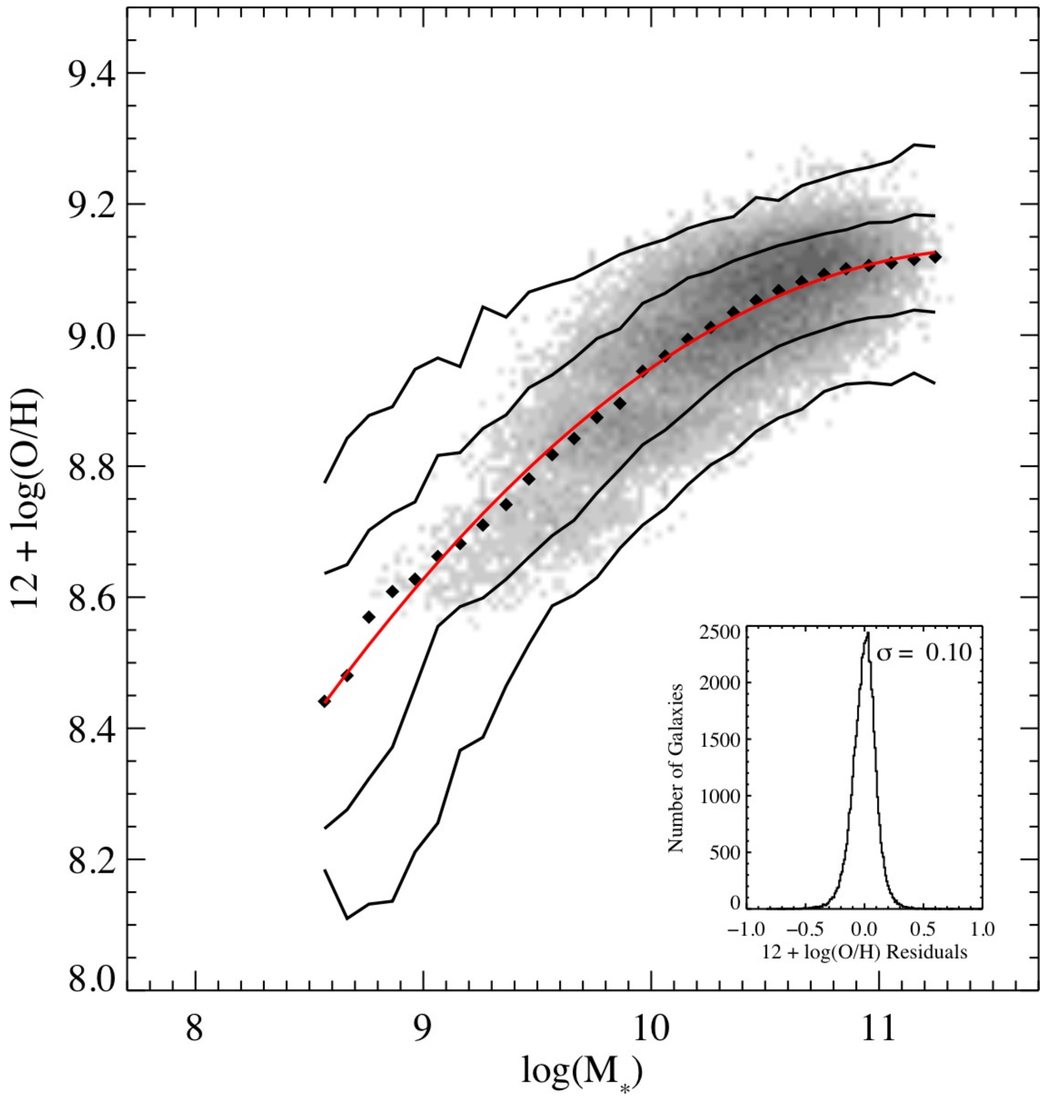
What about other quantities? Gas phase metallicity



Lilly+2013

- Gas phase metallicity, 12+log(O/H), is the metallicity of the ISM as traced by emission lines
- Important tracer of many baryonic processes taking place in galaxies
- Can trace gas inflows, outflows, star-formation, etc.
- Can be reasonably well-modelled by simple gas-regulator models e.g. Lilly+2013

Mass-Metallicity Relation (MZR)



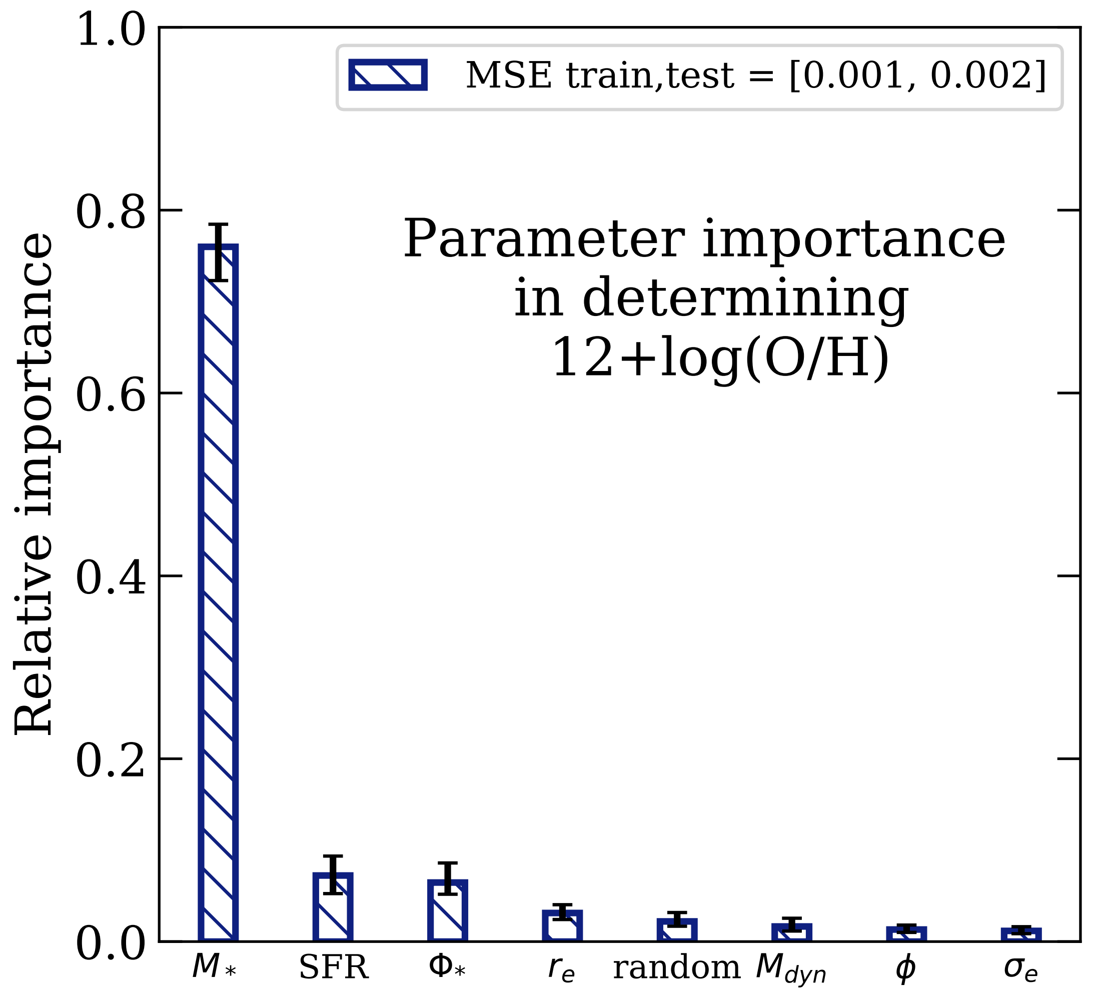
- Metallicity increases with stellar mass up to $M_*\sim10^{10.5-11}M_\odot$
- more massive galaxies have larger gravitational potentials → better able to hold onto metals?
- Or larger stellar mass, hence greater star-formation over its history, hence greater metal production?

Tremonti+ 2004

Does the mass-metallicity relation (MZR) truly trace stellar mass?

Or is it tracing dynamical mass or the underlying gravitational potential?

Can investigate many more quantities simultaneously with random forests



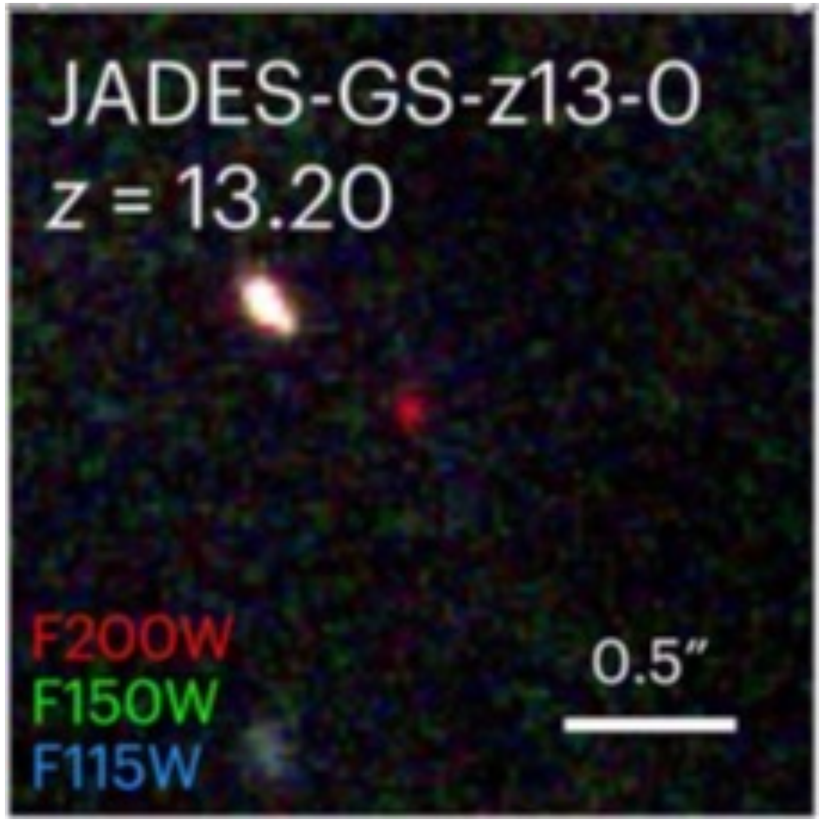
Baker & Maiolino, 2023

- E.g. what does the gas-phase metallicity of SF galaxies depend on?
- Here we show metallicity depends on stellar mass not the underlying gravitational potential

Bayseian Inference with ForcePho

- NIRCam on JWST gives us some of the deepest images of the universe

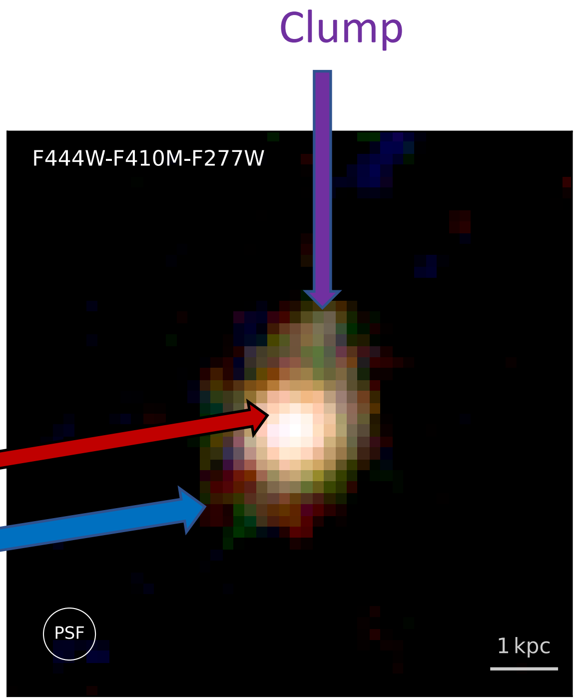
- But how do we accurately extract the light of the source? (whilst accounting for multiple components, PSF effects etc.)



Robertson+2023

ForcePho

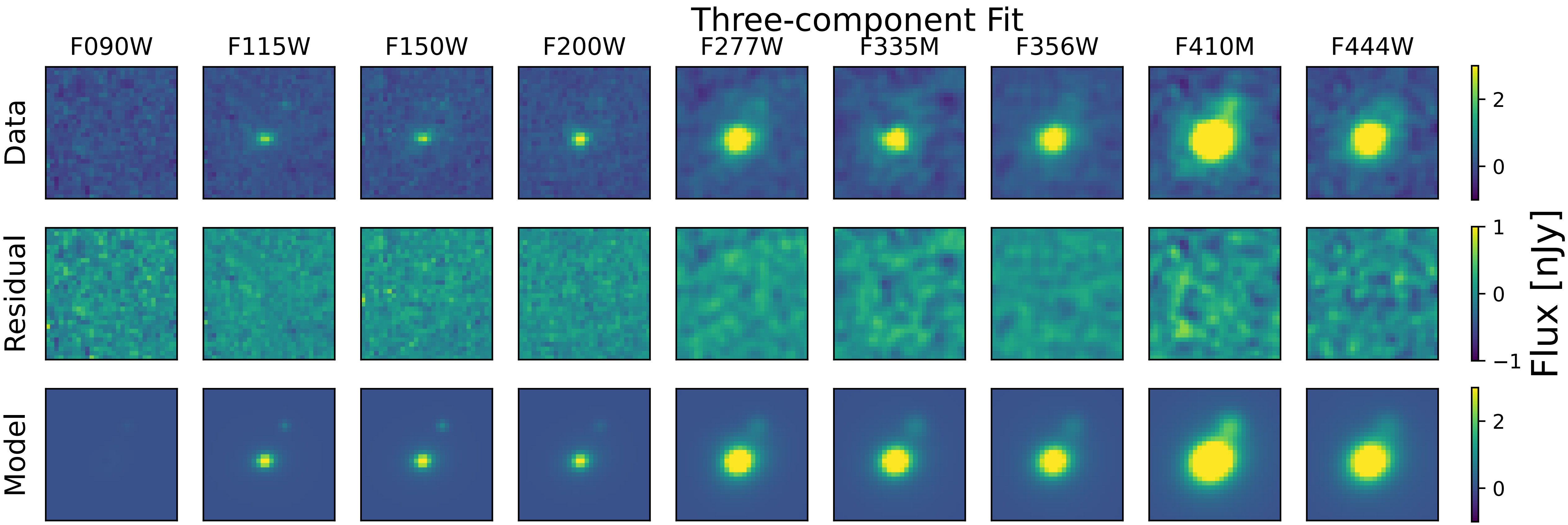
- ForcePho (Johnson+ in prep) is a forced photometry tool that fits multiple PSF convolved Sersic profiles simultaneously to each filter https://github.com/bd-j/forcepho
- Uses Dynamic Nested sampling through Dynesty (Speagle+2020)
- We can use it to fit a three-component model to this z=7.43 galaxy!
- Central Core
- + Disc
- + Clump



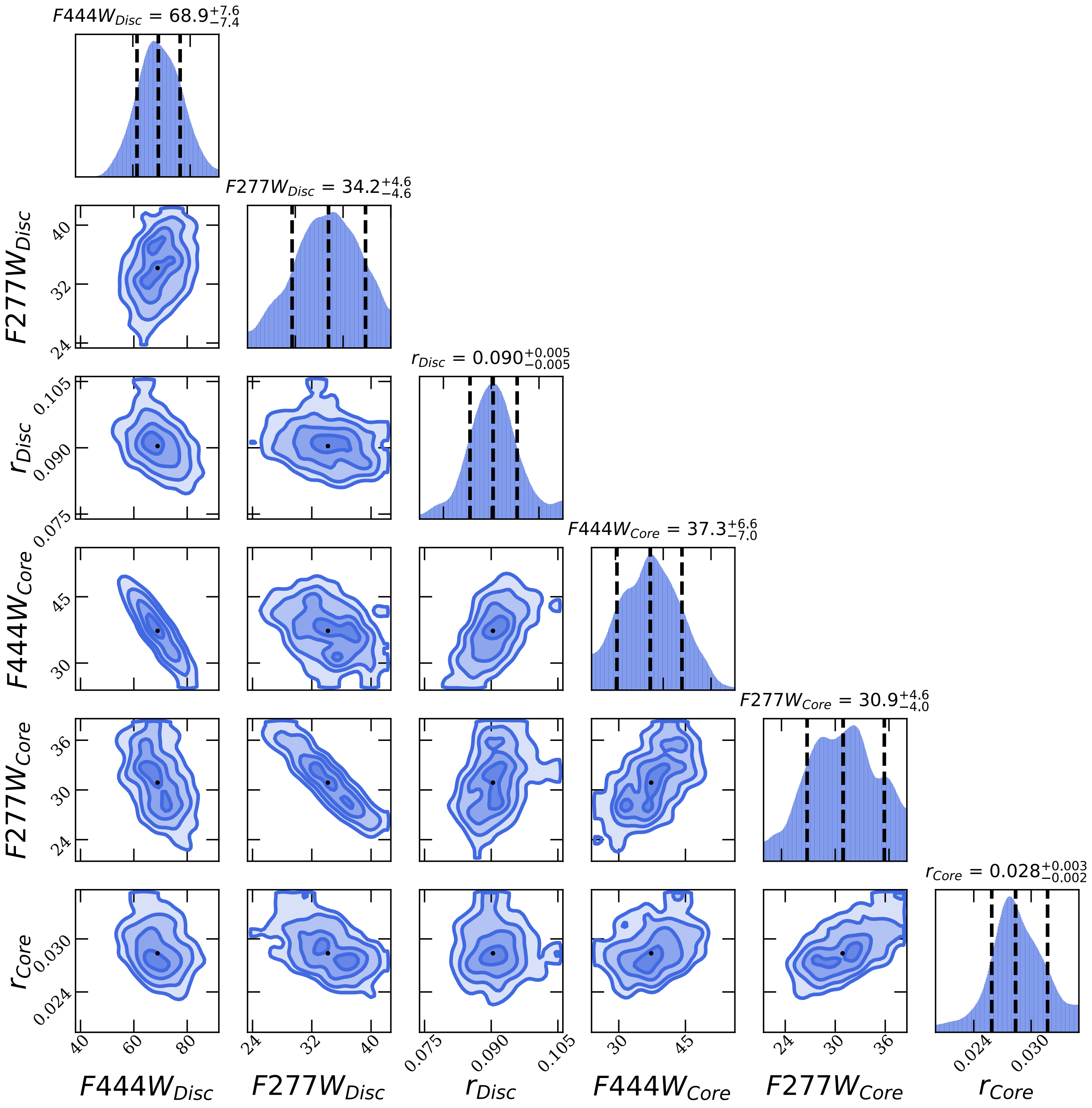
Baker et al., submitted

ForcePho fits

- Can see how ForcePho models the components well in each band



Baker et al., submitted



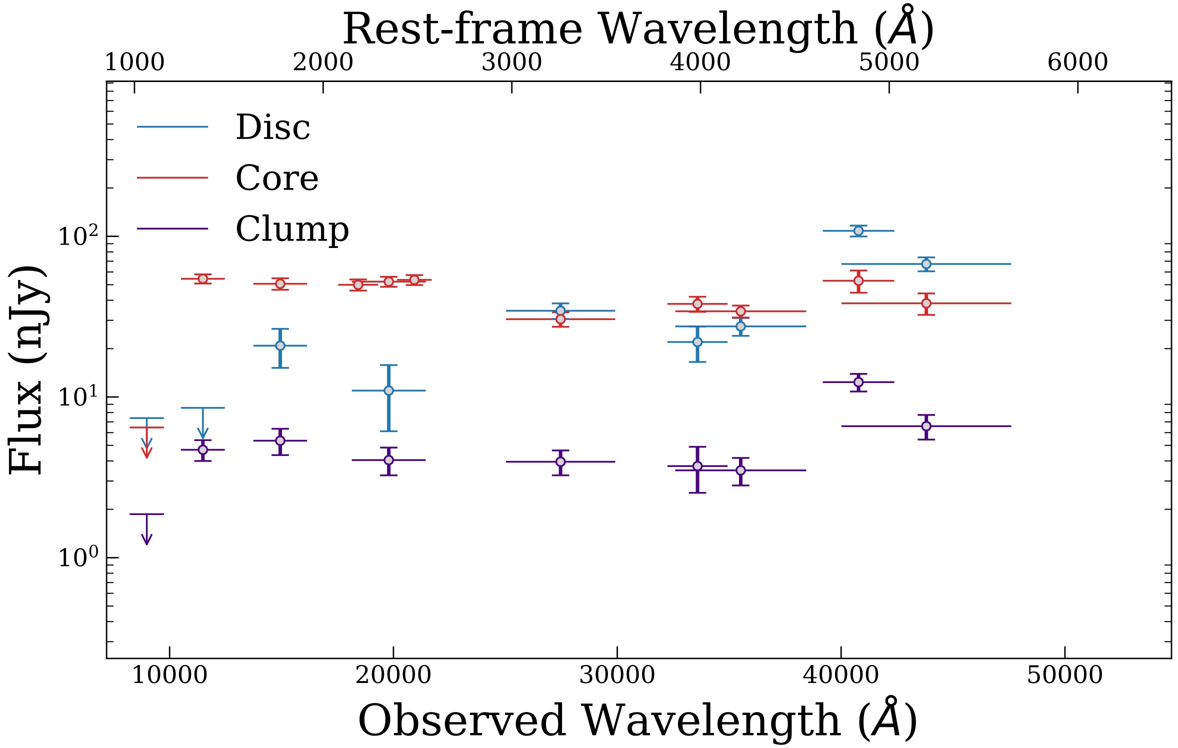
ForcePho Fits

- Corner fit showing fluxes and sizes for the Core and Disc components

Baker et al., submitted

Now we have the SED – Bayesian Inference with Prospector

- SED fitting with Prospector (Johnson+ 2021)
- Nested Sampling through Dynesty
- Non-parametric SFH (Continuity prior, Leja+ 2019)
- Flexible dust attenuation model
- Nebular emission via Cloudy – fit for Ionization parameter and gas phase metallicity

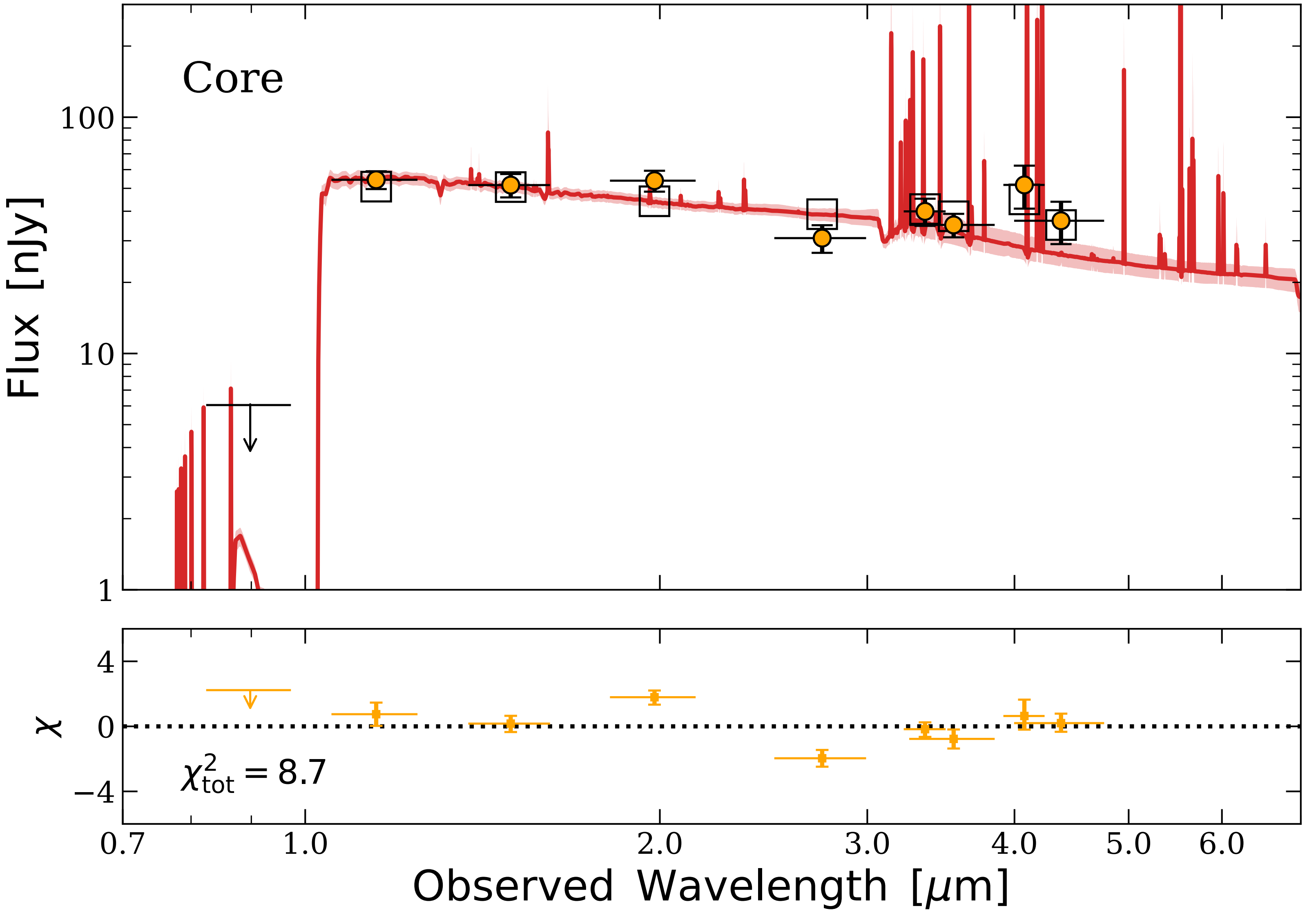


Baker et al., submitted

Core Component

- SED fitting with Prospector (Johnson+ 2021)

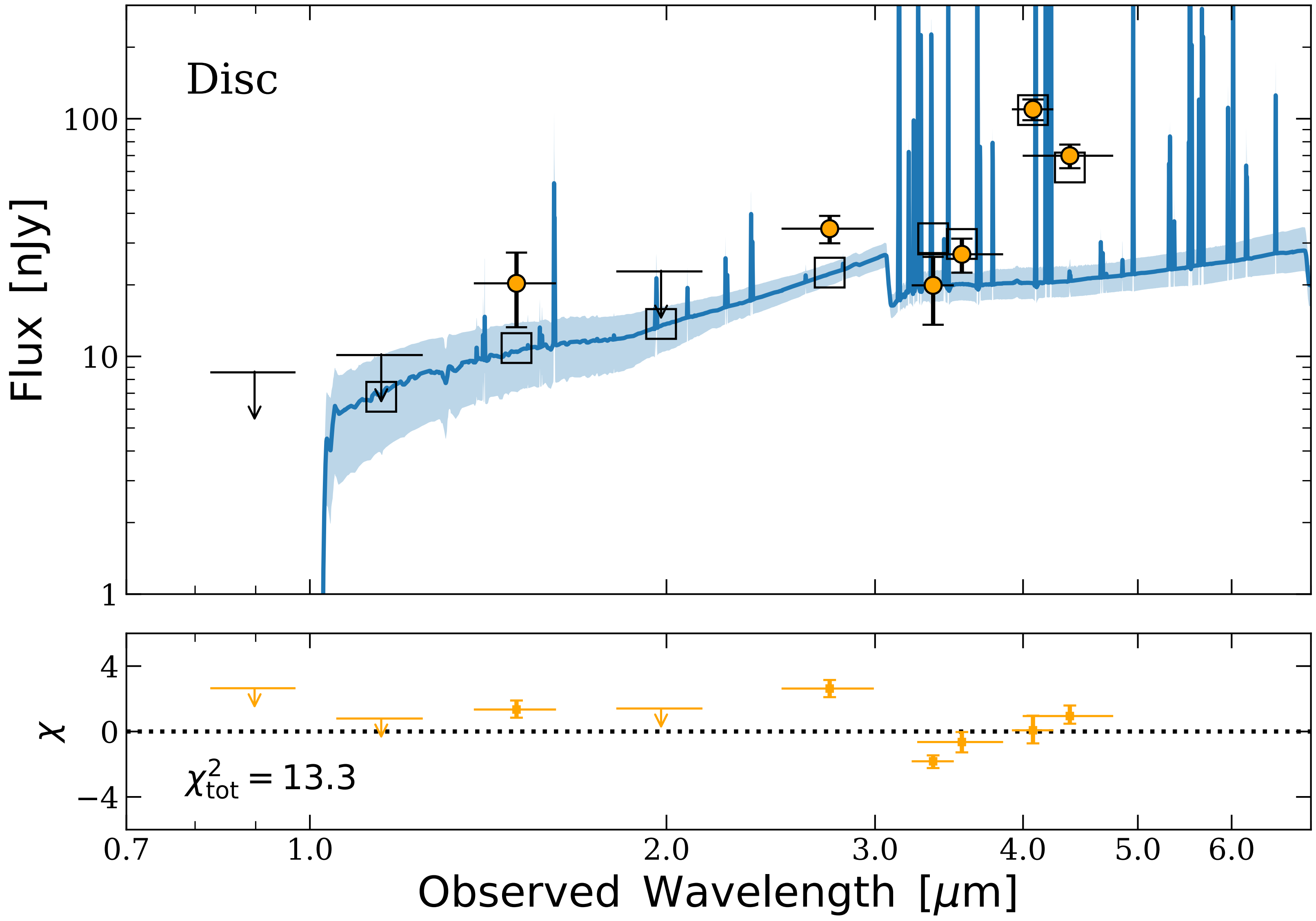
- $\log(M_*/M_\odot)$= 8.39



Baker et al., submitted

Disc Component

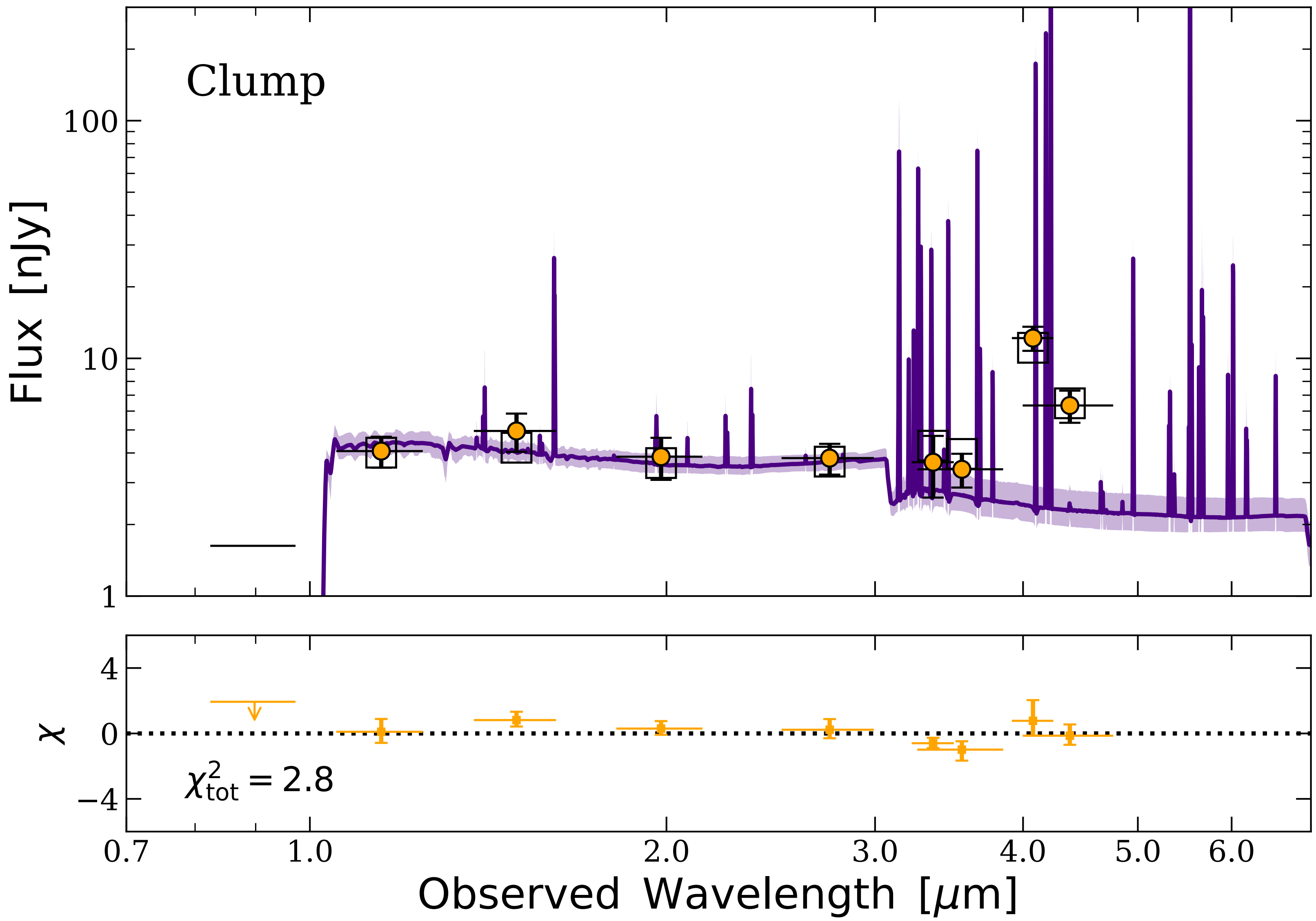
- Fit independently, but with the same fitting routine
- $\log(M_*/M_\odot)$= 8.3
- Also fit combined photometry – find we miss stellar mass → spatially resolved photometry important!



Baker et al., submitted

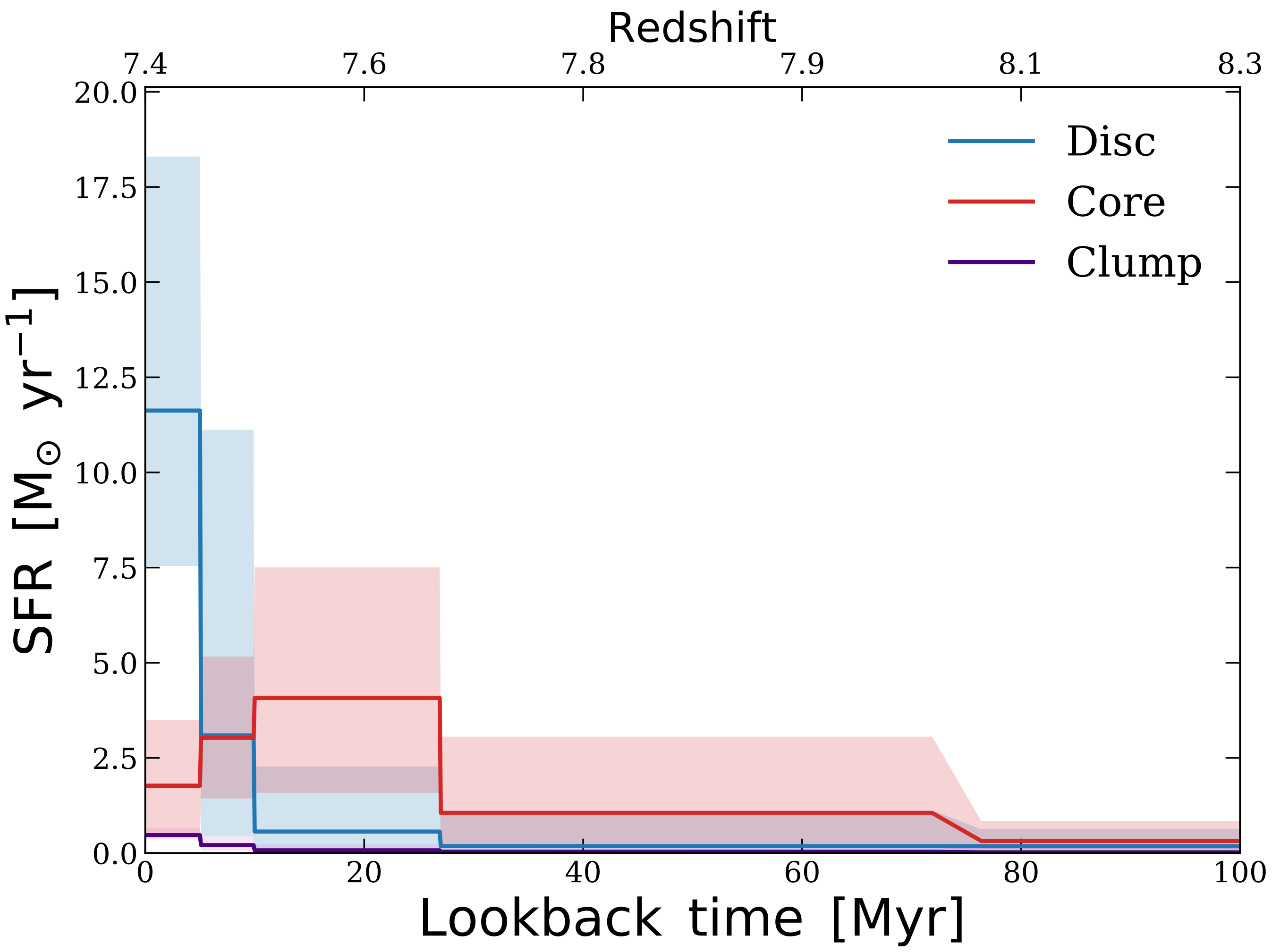
Clump Component

- Fit independently, but with the same fitting routine
- Can see that the clump has a distinct stellar population → might be a small merging galaxy?
- $\log(M_*/M_\odot)$= 7.2



Baker et al., submitted

Star-formation histories



- Disc appears to be undergoing a recent burst
- Core appears to be decreasing in SFR
- Core and Clump appear to be older, Disc appears to be younger

Baker et al., submitted

Summary

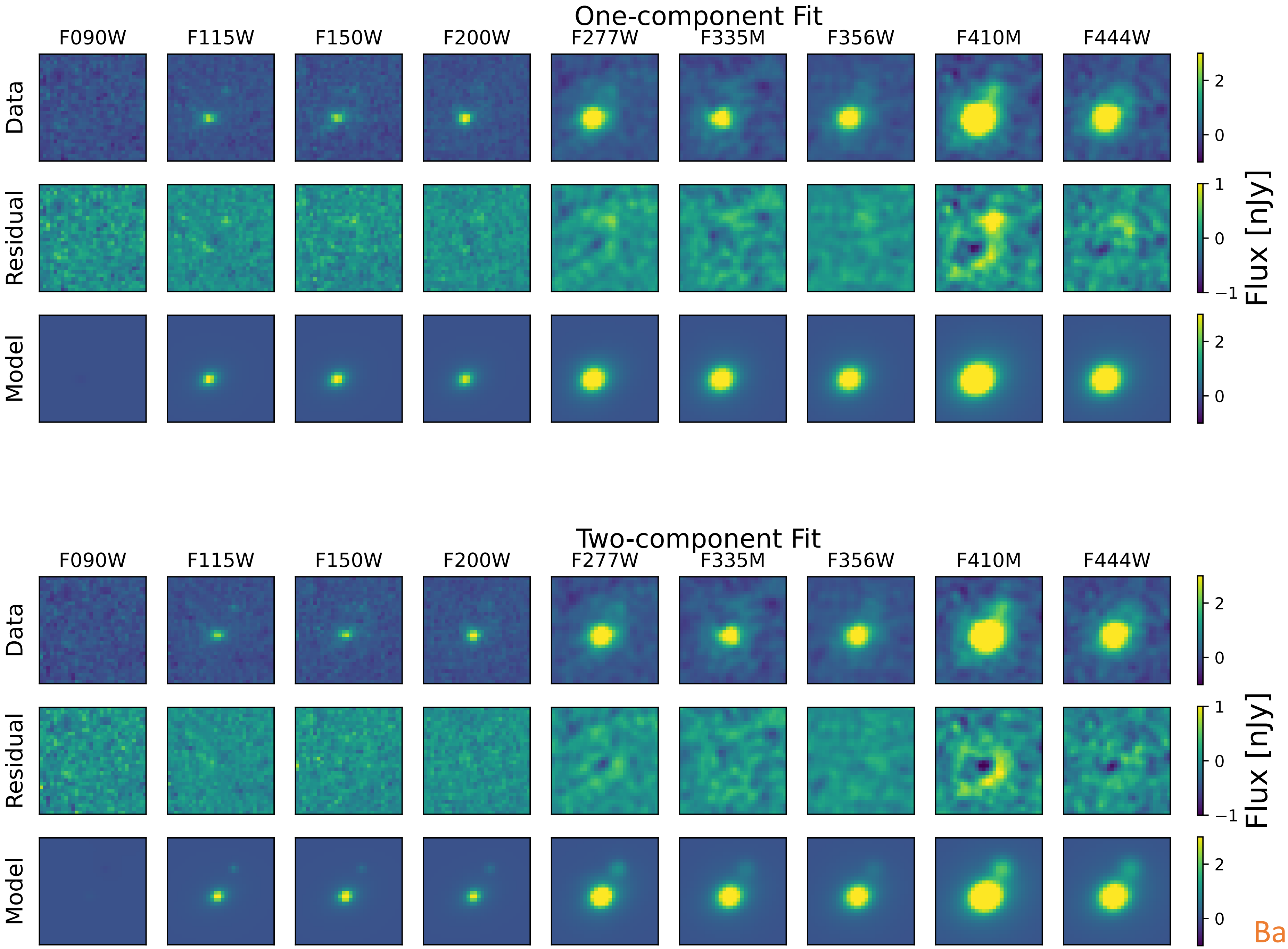
- Partial correlation coefficients and random forest regression can be used to help uncover intrinsic relationships amongst highly inter-correlated quantities

- → SFMS is not an intrinsic scaling relation, rather a byproduct of the MGMS and SK relations

- → MZR does actually trace stellar mass not the gravitational potential

- Nested sampling and Bayesian inference (in this case ForcePho and Prospector) important to accurately fit models for high-z galaxies
- → enables us to model structure in a z=7.43 galaxy

Thank you for listening!



One and two component fits

Baker et al., in prep