Low-surface-brightness science and machine-learning for morphological analysis in LSST

S. Kaviraj, C. Collins, J. Geach, L. Kelvin

LSST:UK All Hands meeting
13 May 2019

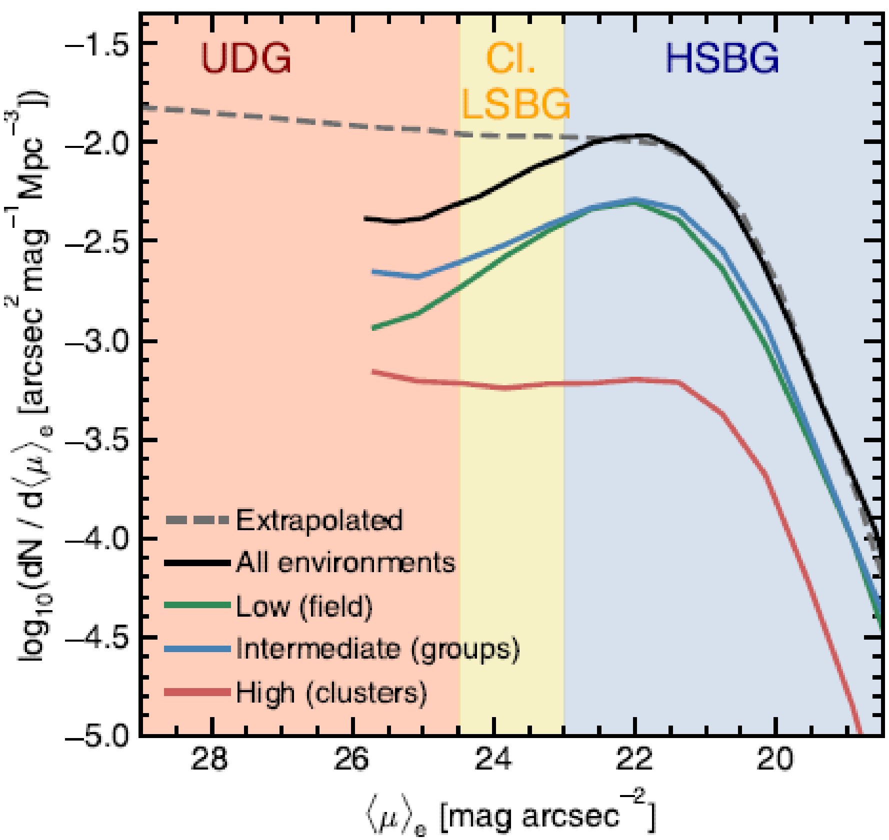
Large Synoptic Survey Telescope

Galaxies, Dark Matter, and Black Holes: Extragalactic Roadmap

Robertson, Brant[1], Banerji, M.[2], Cooper, Michael[3], Davies, R.[4], Ferguson, Henry C.[5], Kaviraj, S.[6], Lintott, C.[4], Lotz, J.[5], Newman, J.[7], Norman, D.[8], Padilla, N.[9], Schmidt, S.[10], Smith, G. P.[11], Verma, A.[4], Working Group Participants, Collaboration Members

The significance of the LSB Universe

Most galaxies are LSB i.e. undetected by today's surveys

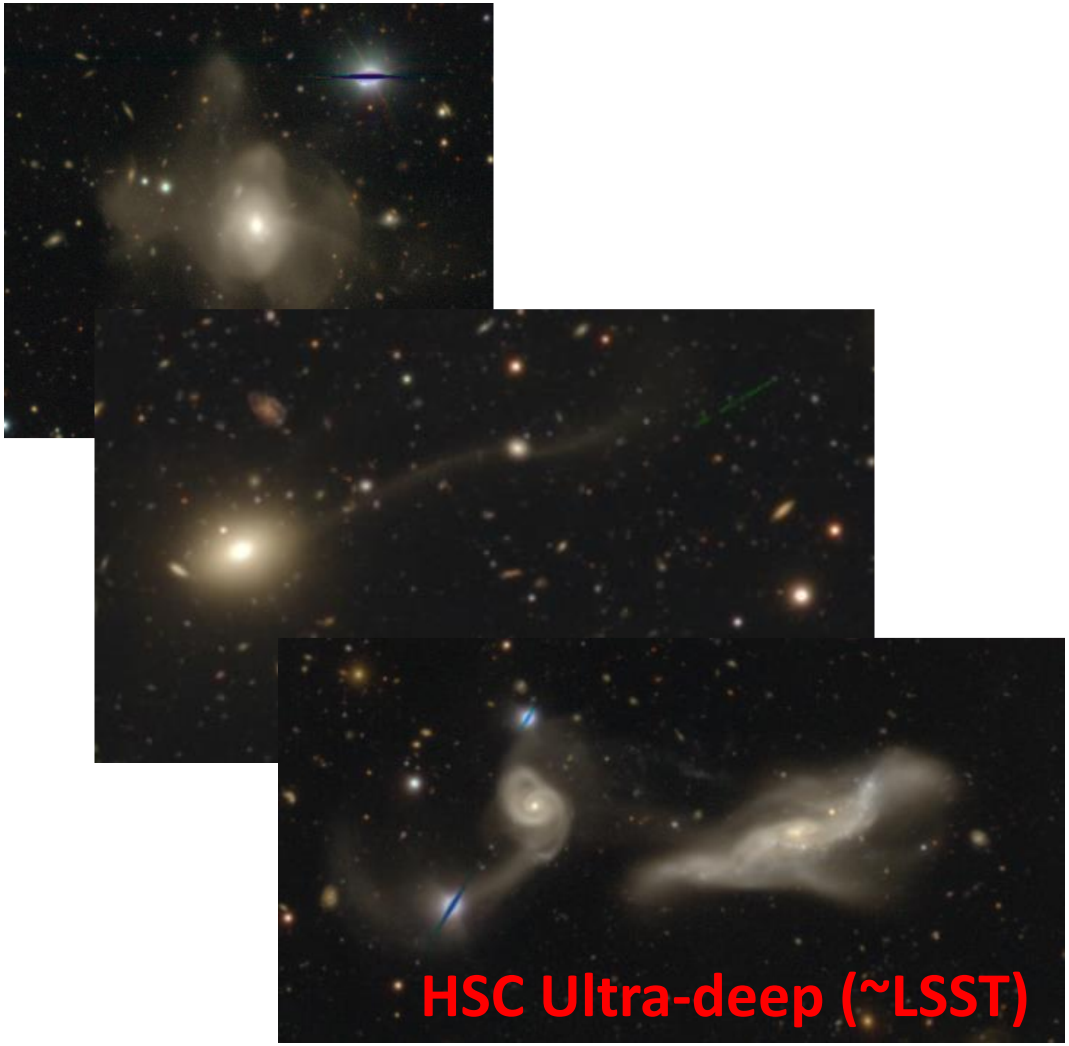


Martin, Kaviraj +19

- At 10^{10} MSun ~20% of galaxies are LSB
- At 10^{8} MSun ~90% of galaxies are LSB
- LSB galaxies are the norm rather than the exception
- Without a good understanding of LSB galaxies our understanding of galaxy evolution remains incomplete

The significance of the LSB Universe

LSB tidal features are key tracers of our theoretical paradigm



- LSB tidal features encode galaxy assembly histories
- Most mergers have low mass ratios which produce faint features
- Key tracers of our structure formation paradigm

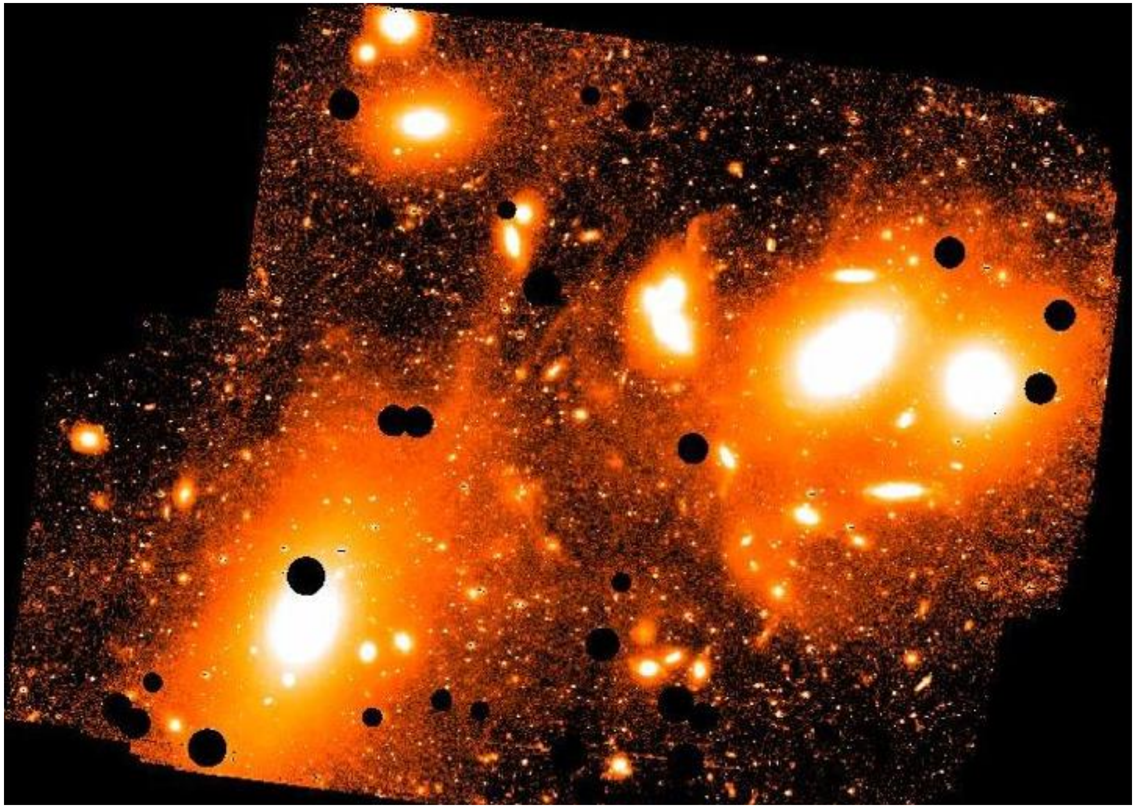
The significance of the LSB Universe

LSB tidal features are key tracers of our theoretical paradigm

Duc +11

The significance of the LSB Universe

Intra-cluster light (ICL)



Mihos +05

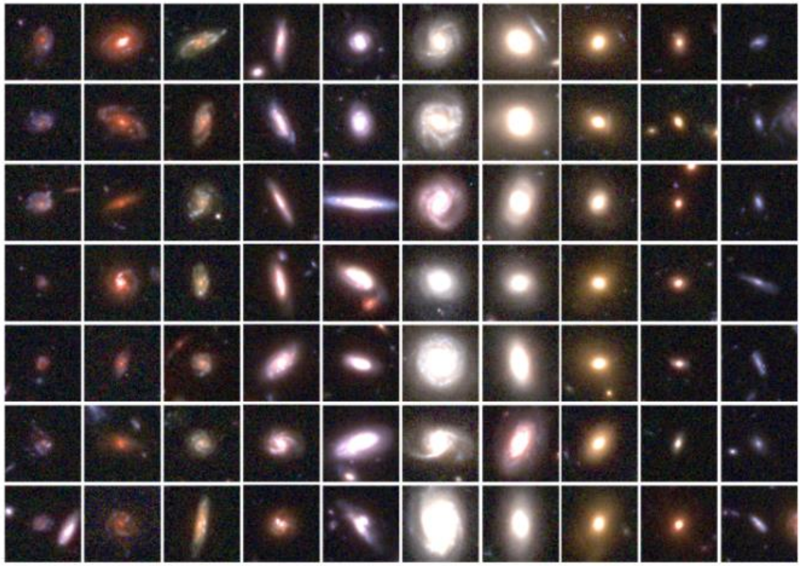
- Galaxy clusters test our cosmological model
- But a significant fraction of baryons in the low-surface-brightness ICL
- Utility of clusters closely linked to how well we understand the evolution of the ICL

The significance of the LSB Universe

- Statistical LSB studies using deep-wide surveys are LSST's niche
- Huge unexplored discovery space – a new frontier in galaxy evolution studies
- But…impossible without proper preparatory work
- Project pipelines are not optimised for LSB studies and no project effort planned in this area

Galaxy morphology

A fundamental quantity in observational cosmology



- Galaxy morphology is a fundamental parameter for all galaxy evolution studies
- Also key for a plethora of science in observational astrophysics
- …e.g. an important prior in photo-z pipelines which underpin weak lensing studies and contextual data for transient lightcurve classifications
- Important for the science goals of several science collaborations e.g. Galaxies, AGN, strong lensing, transients etc.
- But…significant challenges due to LSST's data volume and cadence

Phase B project

Part 1: Enabling low-surface-brightness science using LSST

The problem:

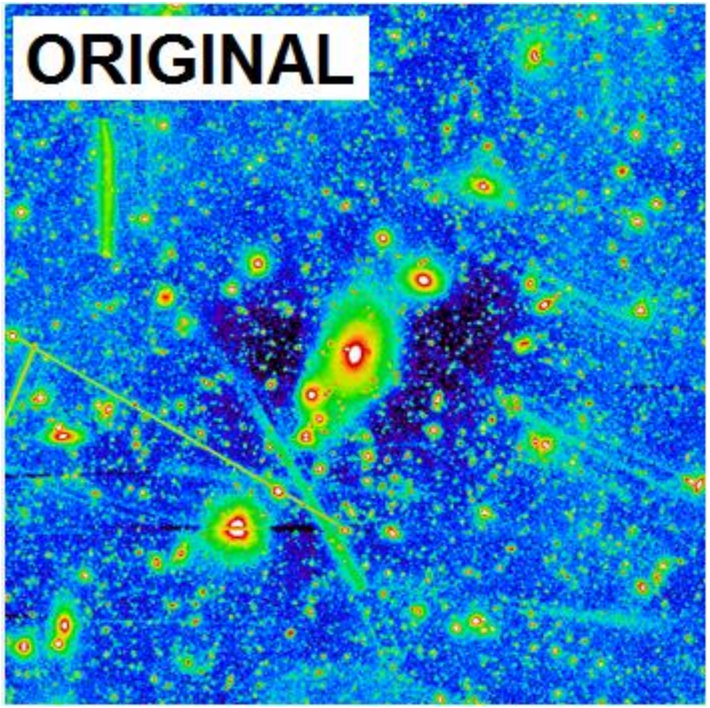
- LSB structures acutely susceptible to sky over-subtraction and shredding by de-blenders
- Means that galaxy population is incomplete and LSB structures are removed/truncated/shredded -> LSB science impossible.
- Project pipeline known to suffer these effects (optimised for photometry i.e. smaller spatial scales than LSB structures) and no LSB-enabling work is planned by Project

The solution (provided by this WP):

- Optimised sky subtraction to preserve LSB structures at any spatial scale
- Machine-learning algorithms for mitigating shredding

Phase B project

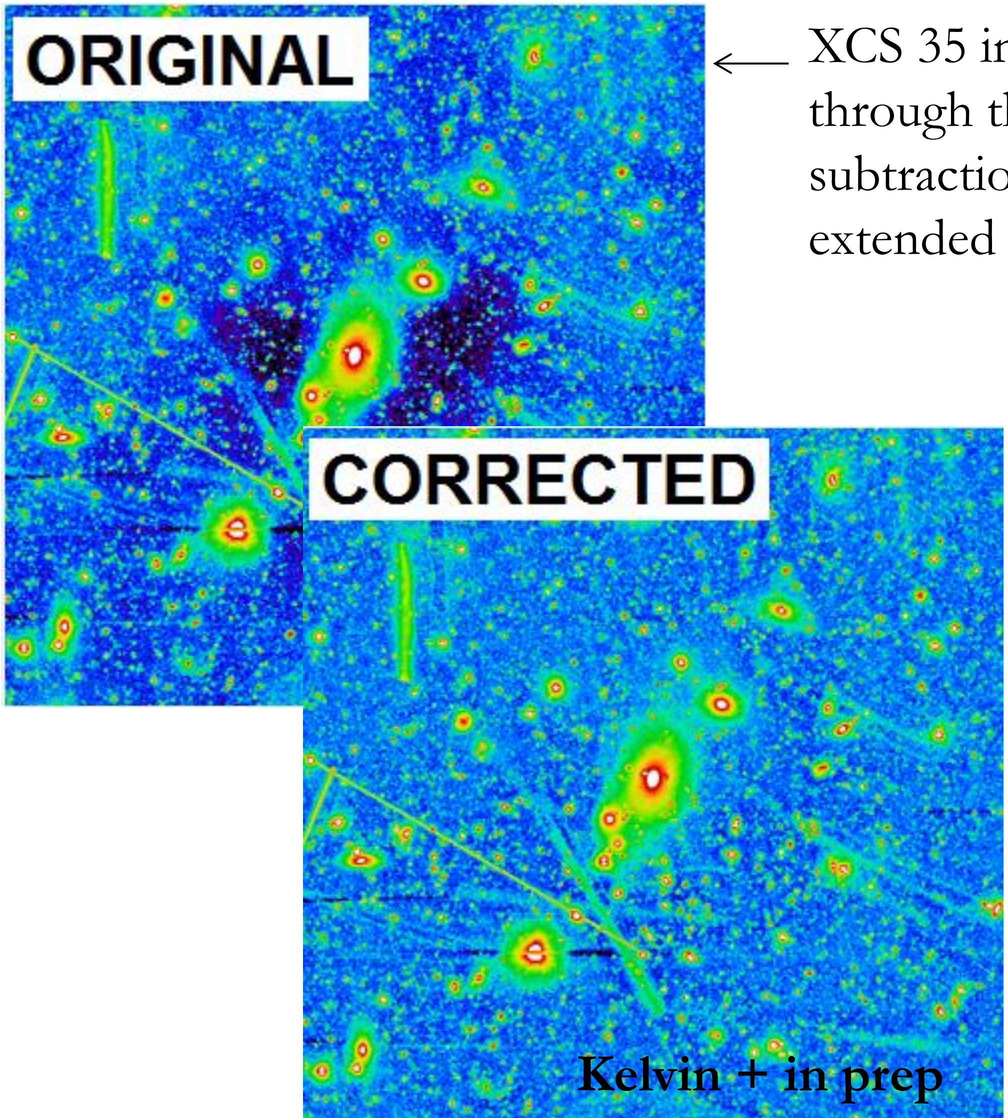
Part 1: Enabling low-surface-brightness science using LSST



XCS 35 in HSC Deep DR1 i-band reduced through the current LSST pipeline. Sky over-subtraction is visible around bright and extended sources.

Phase B project

Part 1: Enabling low-surface-brightness science using LSST



XCS 35 in HSC Deep DR1 i-band reduced through the current LSST pipeline. Sky over-subtraction is visible around bright and extended sources.

- Fit 2D Sersic models to sources
- Characterise the expected flux in the wings
- Residual between model and science map used to define over-subtraction threshold
- Values below threshold added back

Phase B project

Part 1: Enabling low-surface-brightness science using LSST

- Explore optimal background subtraction strategies for multiple spatial scales

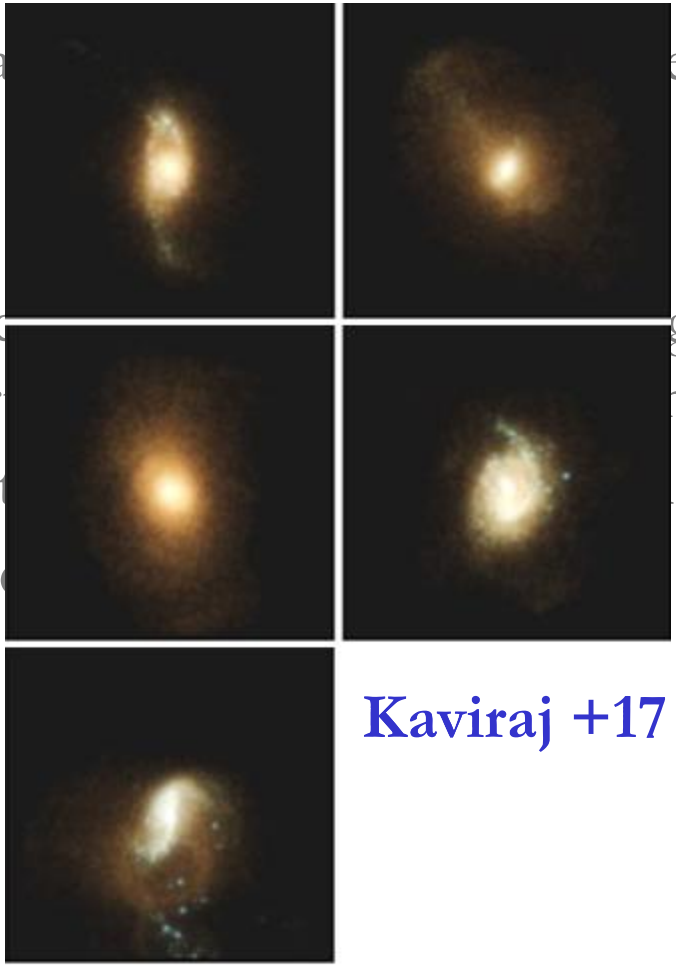
- Benchmark on mock images from cosmological simulations e.g. Horizon-AGN **(Kaviraj +17)** inserted into HSC and ComCam frames to ensure proper representation of noise/background sources/camera effects in the data

Phase B project

Part 1: Enabling low-surface-brightness science using LSST

- Explore optimal background subtra e spatial scales

- Benchmark on mock images from g. Horizon-AGN **(Kaviraj +17)** inser n frames to ensure proper representat d sources/camera effects in the



Kaviraj +17

Phase B project

Part 1: Enabling low-surface-brightness science using LSST

- Explore optimal background subtraction strategies for multiple spatial scales

- Benchmark on mock images from cosmological simulations e.g. Horizon-AGN **(Kaviraj+17)** inserted into HSC and ComCam frames to ensure proper representation of noise/background sources/camera effects in the data

- The results of this WP will be critical for the global LSST community

Phase B project

Part 2: Measuring galaxy morphologies in LSST

The problem:

- LSST data volumes are unprecedented
- Makes visual classification (even using systems like Galaxy Zoo) intractable
- Requirement: **fast and accurate** automated techniques benchmarked via visual classification
- LSST's short cadence is an extra hurdle – repeatedly producing training sets for supervised ML impractical

The solution (provided by this WP):

- *Unsupervised* ML + benchmarking via visual classification
- UML compresses arbitrarily large galaxy pop. into small number of \`morphological clusters'
- …and these **clusters** are then benchmarked against visual classification

Phase B project

Part 2: Measuring galaxy morphologies in LSST

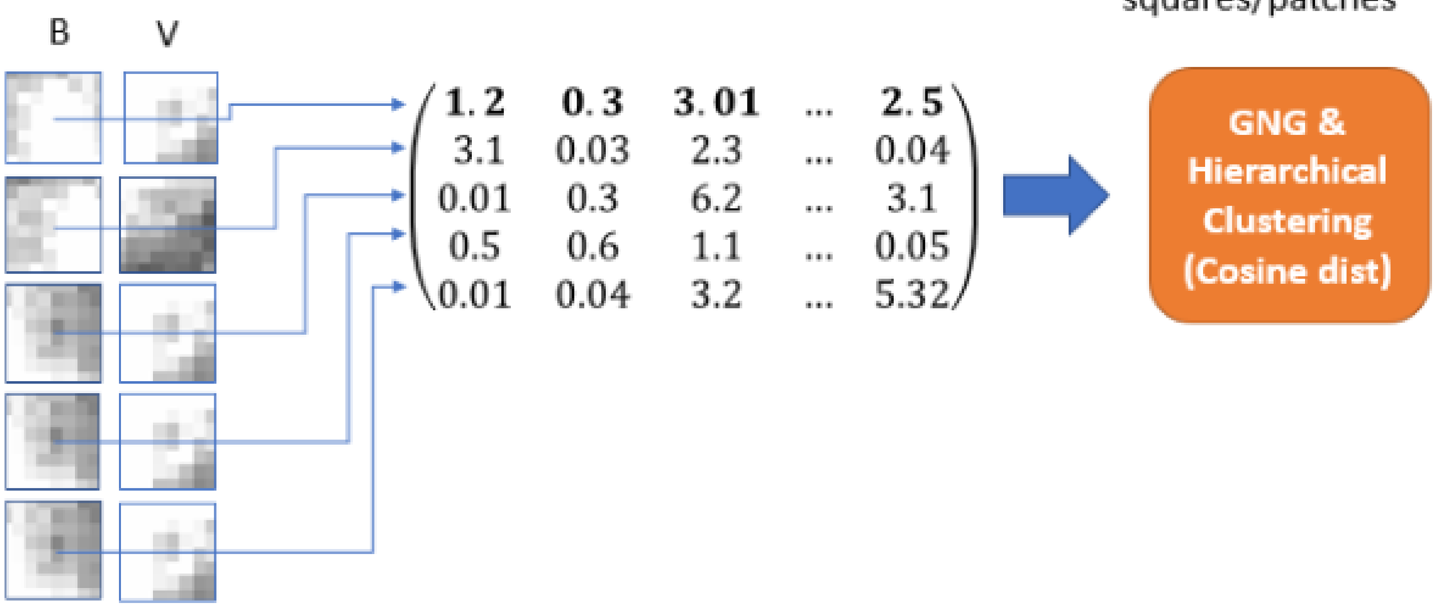
An Unsupervised Approach

Step 1 – Build a dictionary of patches

N px by N px squares
(multiple filters)

Convert pixel values into vectors
(multiple filters)

Sort vectors into groups -
a dictionary of 1000 average
squares/patches



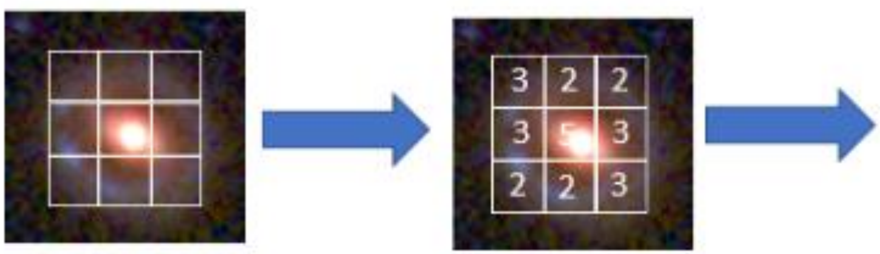
Phase B project

Part 2: Measuring galaxy morphologies in LSST

An Unsupervised Approach

Step 2 – Create Galaxy Vector Representations

Map dictionary vectors to the galaxy



Each vector is a histogram of the dictionary patches that form the galaxy

	Patch Dictionary Mapping -> Galaxy Vectors						
Galaxy Id	1	2	3	4	5	...	1000
1	0	4	4	0	1	...	0
2	10	0	254	0	0	...	0
3	0	33	0	34	0	...	0
4	0	67	2	3	45	...	10
5	45	5	21	89	0	...	0
6	3	23	0	0	0	...	67
...	...	...	...	...	...	...	...
n	0	0	3	0	54	...	3

Phase B project

Part 2: Measuring galaxy morphologies in LSST

An Unsupervised Approach

Step 3 – Sort Galaxy Vectors into Collections

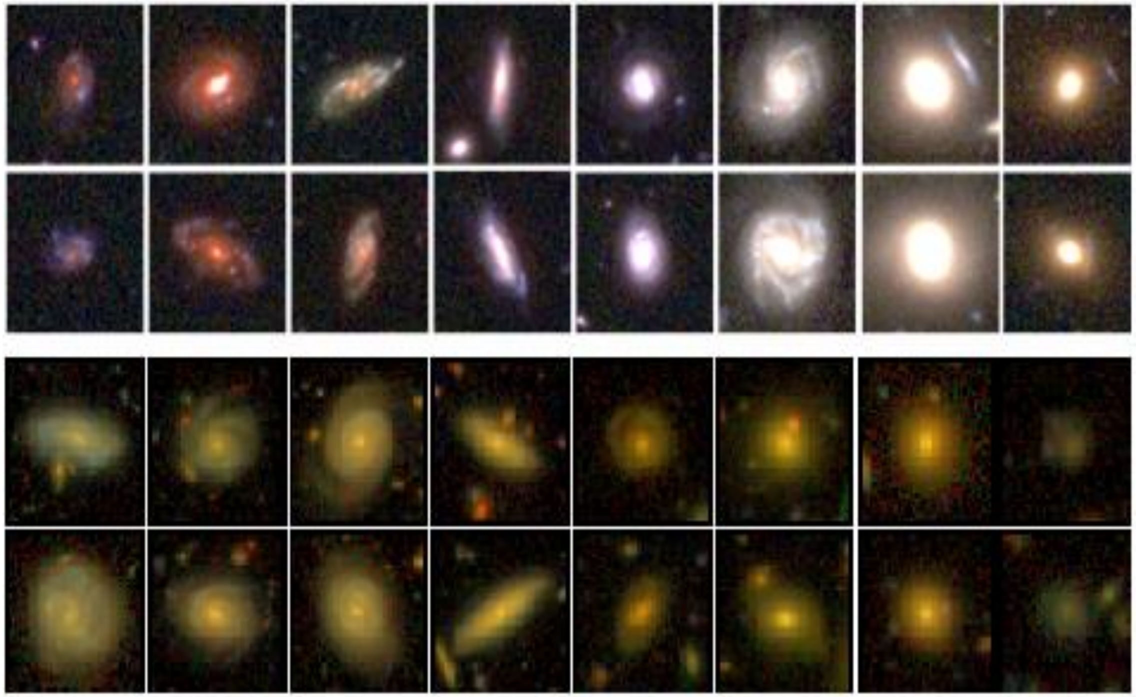
Galaxy Id	Galaxy Vectors 1	2	3	4	5	...	1000
1	0	4	4	0	1	...	0
2	10	0	254	0	0	...	0
3	0	33	0	34	0	...	0
4	0	67	2	3	45	...	10
5	45	5	21	89	0	...	0
6	3	23	0	0	0	...	67
...	...	...	...	...	...	...	...
n	0	0	3	0	54	...	3

Sort galaxies into groups and create a dictionary of galaxies

GNG or Kmeans or Hierarchical Clustering

Phase B project

Part 2: Measuring galaxy morphologies in LSST



← **HST** **(Hocking, Geach +18)**

← **HSC Ultra-deep ~ LSST 10 yr** **(Martin, Kaviraj +)**

↑

Tested on bright galaxies only

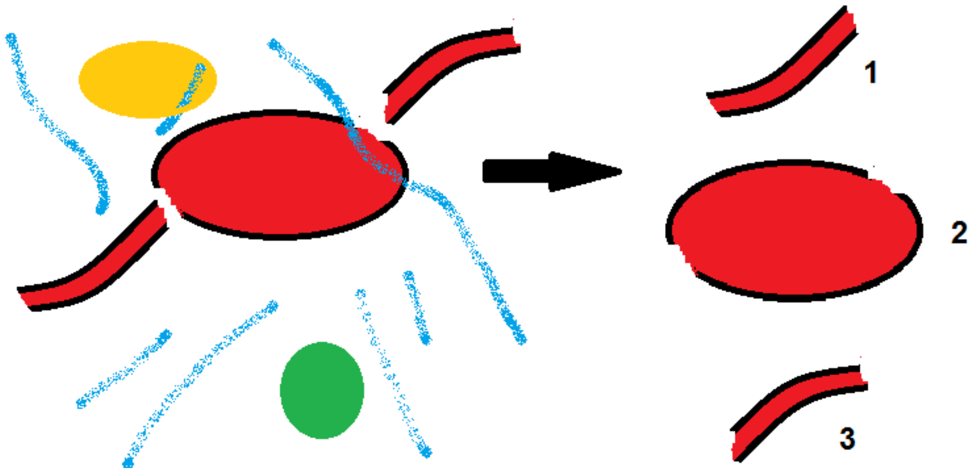
Phase B project

Part 2: Measuring galaxy morphologies in LSST

- Successfully tested on relatively bright galaxies in HST-CANDELS and HSC Udeep
- Adapt for LSB galaxies (which dominate the number density)
- Develop star-galaxy separation capability (works reasonably well on HST data but untested on HSC/LSST type data)
- Develop capability for strong lens detection (works on HST data but untested on HSC/LSST type imaging)
- Develop capability for mitigating shredding

Phase B project

Part 2: Measuring galaxy morphologies in LSST



- If sky is over subtracted then galaxies and tidal features can be shredded
- Since it works at the **pixel level,** the UML algorithm can mitigate this
- …because tidal features inherit properties from their parent galaxies

LSB science and UML for galaxy morphology

Summary

Principal aims:

(1) Enable LSB science using LSST by developing optimised background modelling and sky subtraction at all spatial scales (funded in Phase B)

(2) Develop unsupervised machine-learning infrastructure for morphological analysis

- Morphological classification (not funded)
- Star-galaxy separation (not funded)
- Lens identification (not funded)
- Mitiating shredding (funded in Phase B)