

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

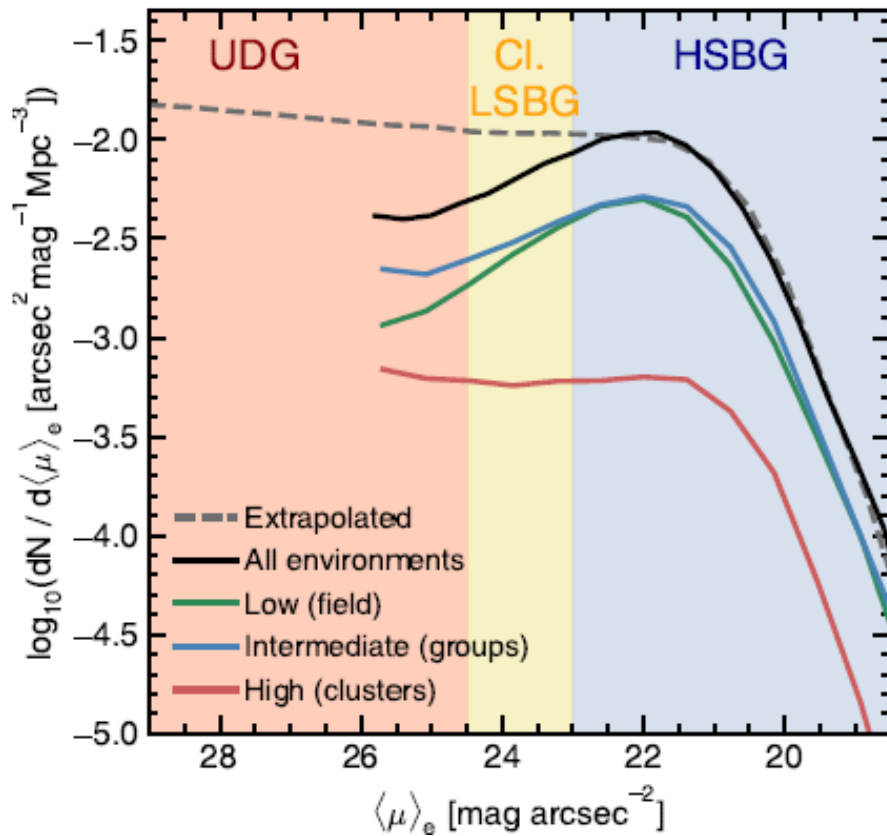
Large Synoptic Survey Telescope Galaxies, Dark Matter, and Black Holes: Extragalactic Roadmap

Robertson, Brant¹, Banerji, M.², Cooper, Michael³, 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

3.2.5	Disentangling Complicated Lines of Sight	22
3.2.6	Forward Modeling LSST Clusters and Groups	22
3.3	Deep Drilling Fields	24
3.3.1	Coordinating Ancillary Observations	24
3.3.2	Observing Strategy Cadence	25
3.3.3	Data Processing	25
3.4	Galaxy Evolution Task Lists	27
3.4.1	Techniques, Algorithms, or Software Development	27
3.4.2	Techniques for identifying and deblending overlapping galaxies	28
3.4.3	Optimizing Galaxy Morphology Measurements	29
3.4.4	Optimizing Galaxy Photometry	30
3.4.5	Optimizing Measurements of Stellar Population Parameters	31
3.6	Low-Surface Brightness Science	35
3.6.1	Techniques for Finding Low-Surface Brightness Tidal Features	35
3.6.2	Low-Surface Brightness Galaxies	37
3.6.3	Probing the Faint Outskirts of Galaxies with LSST	40
3.6.4	Low-Surface Brightness Intracluster Light	41
3.7	Photometric Redshifts	43
3.7.1	Impact of Filter Variations on Galaxy photo-z Precision	43
3.7.2	Photometric Reshifts in the LSST Deep Drilling Fields	43
3.7.3	Multivariate Physical Properties of Galaxies from Photometric Redshifts	44
3.7.4	Identifying Spectroscopic Redshift Training Sets for LSST	45
3.7.5	Develop Techniques to Identify Specific Sub-Populations of Galaxies	45
3.7.6	Simulations with Realistic Galaxy Colors and Physical Properties	46
3.7.7	Using Galaxy Size and Surface Brightness distributions as Photo-z Priors	47
3.8	Theory and Mock Catalogs	48
3.8.1	Image Simulations of Galaxies with Complex Morphologies	48
3.8.2	New Theoretical Models for the Galaxy Distribution	48
3.8.3	Design of New Empirical Models for the Galaxy Distribution.	49
3.8.4	Estimating Uncertainty for Large Scale Structure Statistics	50

The significance of the LSB Universe

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

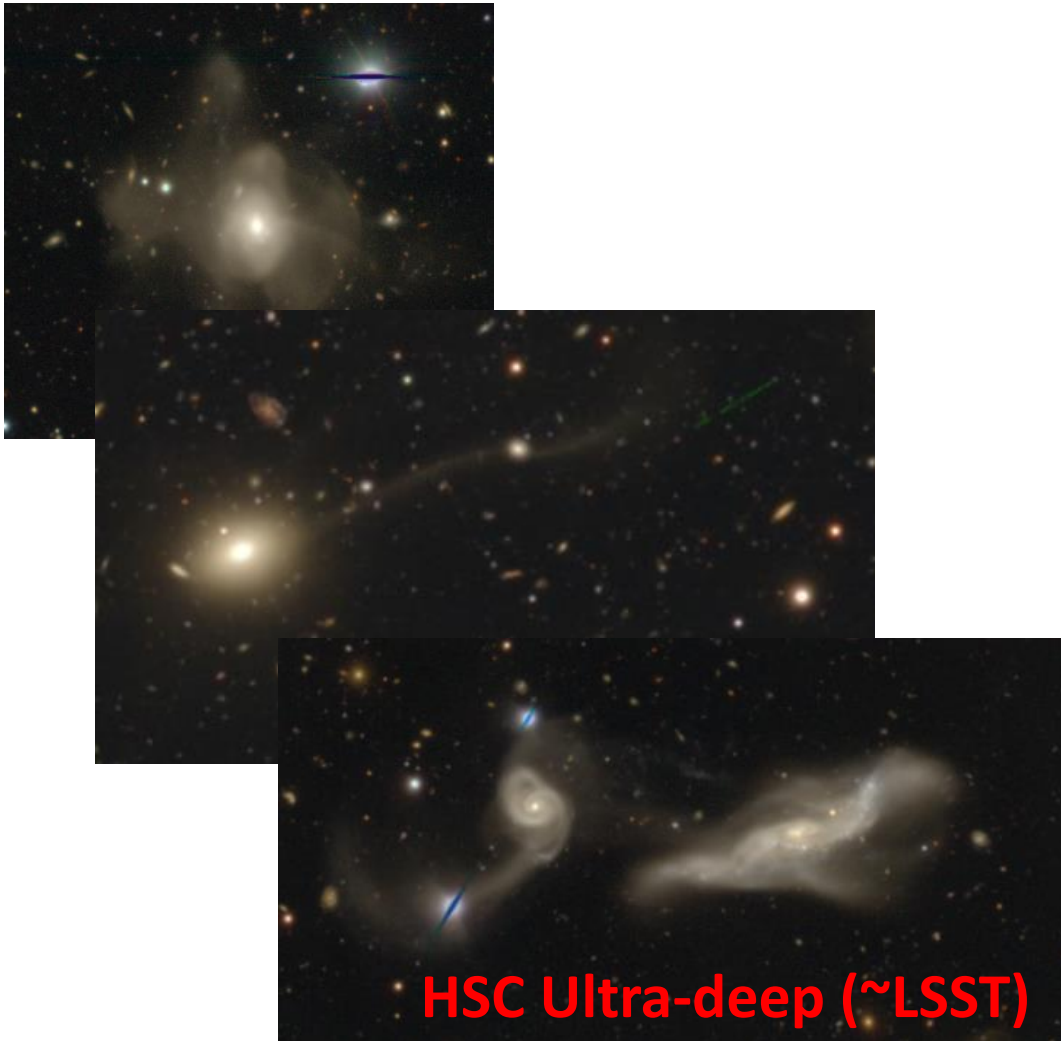


- At 10^{10} MSun $\sim 20\%$ of galaxies are LSB
- At 10^8 MSun $\sim 90\%$ of galaxies are LSB
- LSB galaxies are the norm rather than the exception
- Without a good understanding of LSB galaxies our understanding of galaxy evolution remains incomplete

Martin, Kaviraj +19

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

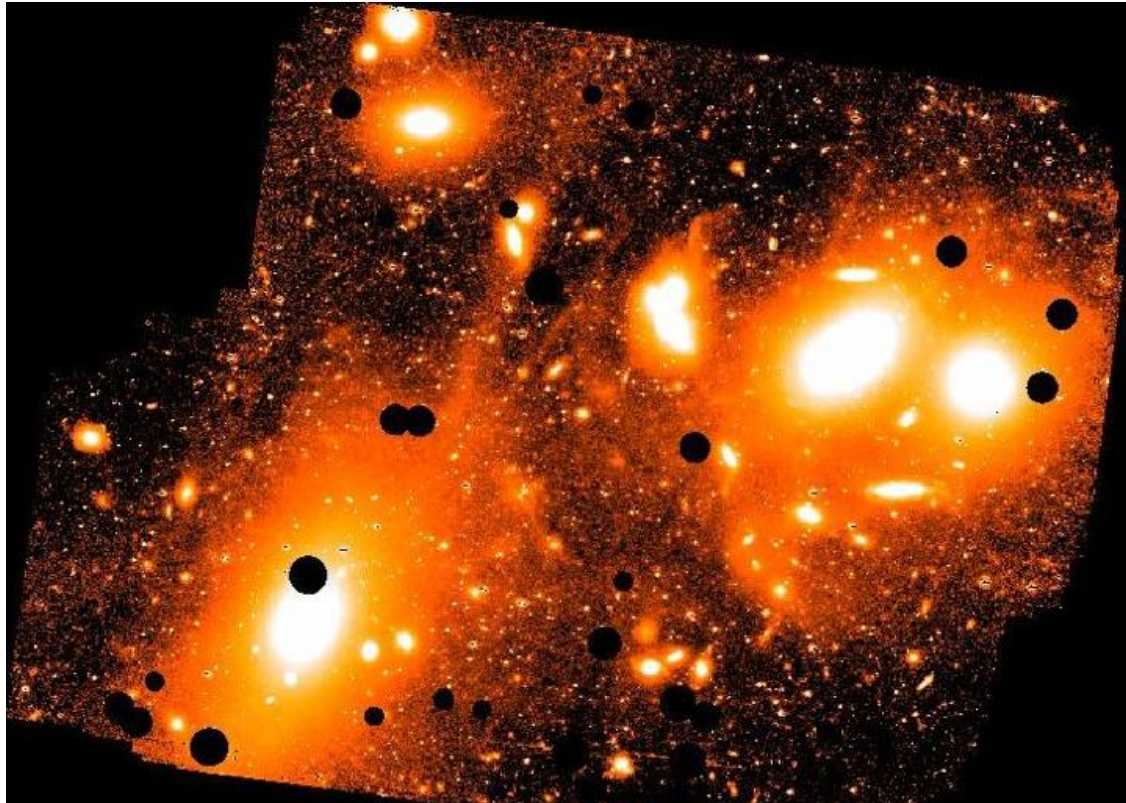
LSB tidal features are key tracers of our theoretical paradigm



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

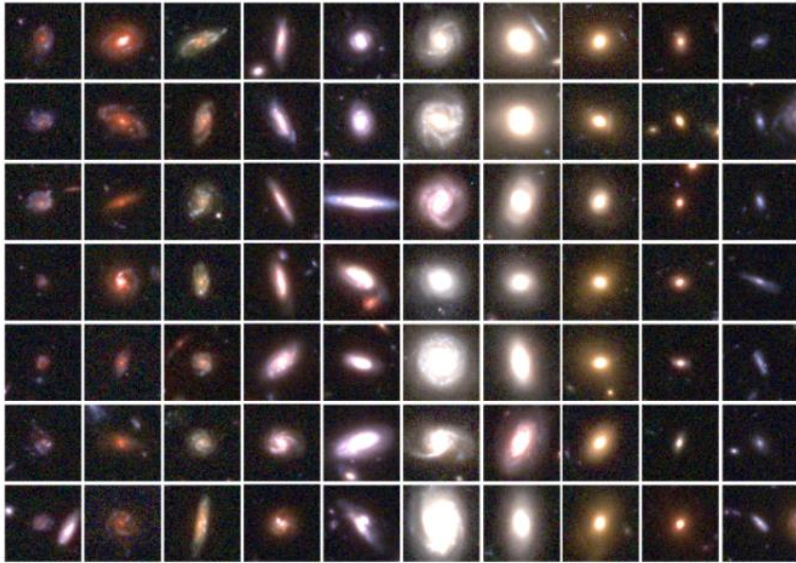


The significance of the LSB Universe

- 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:

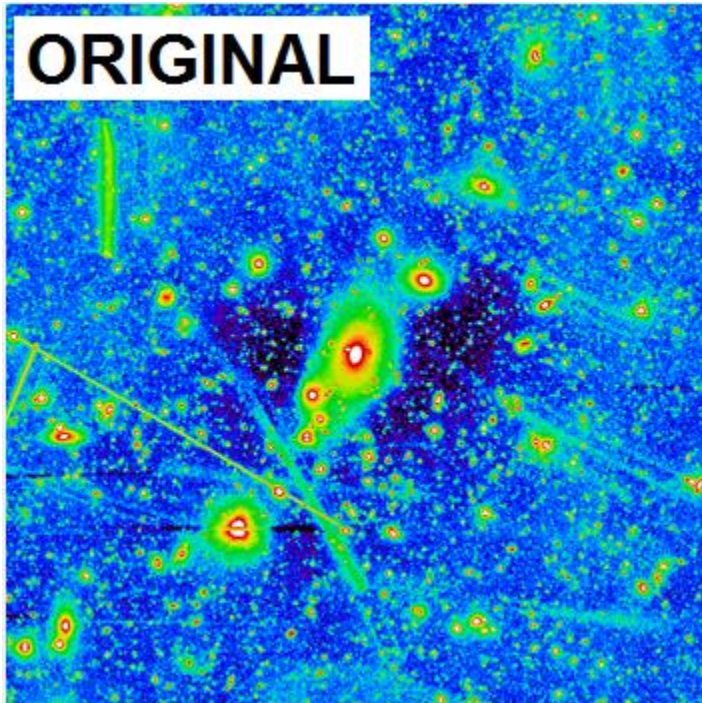
- LSB structures acutely susceptible to sky over-subtraction and shredding by de-blenders
- Means that galaxy population is incomplete and LSB structures are removed/truncated/shredded -> LSB science impossible.
- Project pipeline known to suffer these effects (optimised for photometry i.e. smaller spatial 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

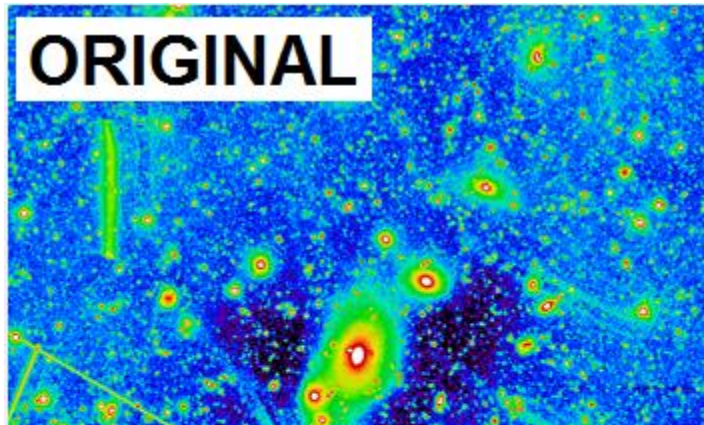
Part 1: Enabling low-surface-brightness science using LSST



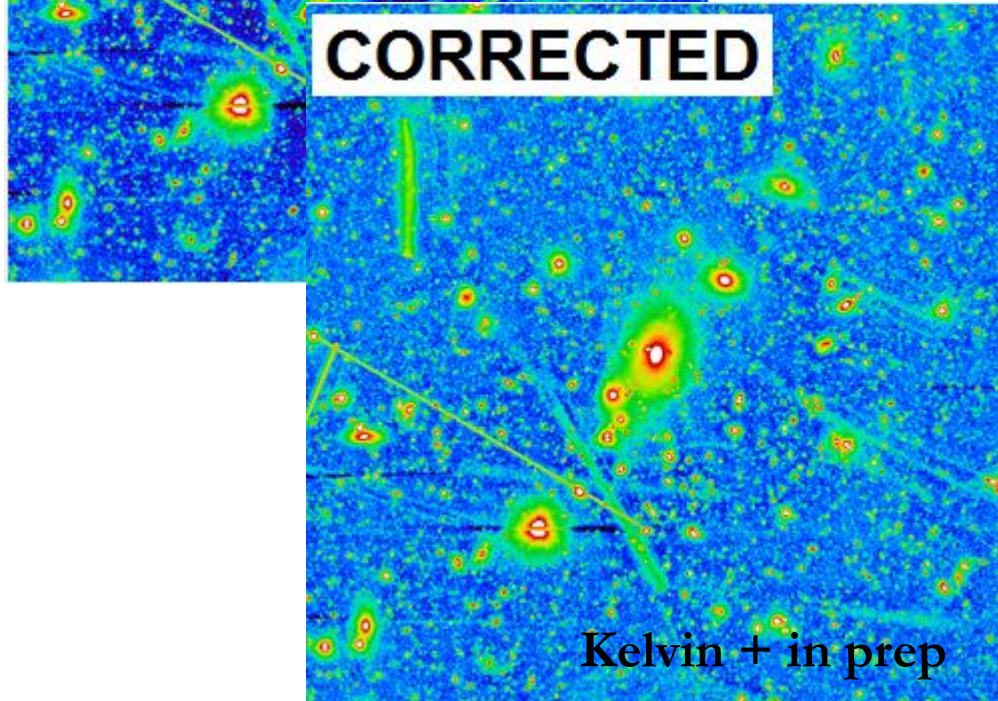
← XCS 35 in HSC Deep DR1 i-band reduced through the current LSST pipeline. Sky over-subtraction is visible around bright and extended sources.

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

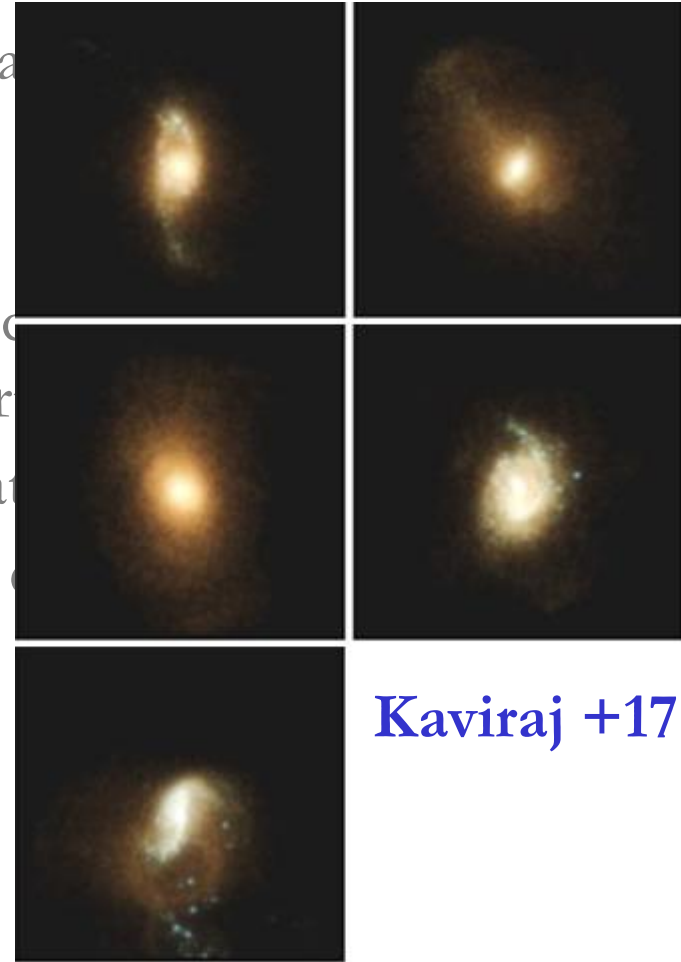
Part 1: Enabling low-surface-brightness science using LSST

- 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 subtraction at various spatial scales
- Benchmark on mock images from Horizon-AGN (Kaviraj +17) inserted into real frames to ensure proper representation of low-surface-brightness sources/camera effects in the data



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


$$\begin{pmatrix} 1.2 & 0.3 & 3.01 & \dots & 2.5 \\ 3.1 & 0.03 & 2.3 & \dots & 0.04 \\ 0.01 & 0.3 & 6.2 & \dots & 3.1 \\ 0.5 & 0.6 & 1.1 & \dots & 0.05 \\ 0.01 & 0.04 & 3.2 & \dots & 5.32 \end{pmatrix}$$


**GNG &
Hierarchical
Clustering
(Cosine dist)**

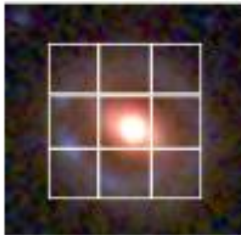
Phase B project

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

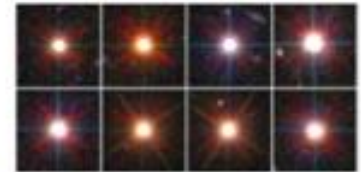
An Unsupervised Approach

Step 3 – Sort Galaxy Vectors into Collections

Galaxy	Galaxy Vectors						
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

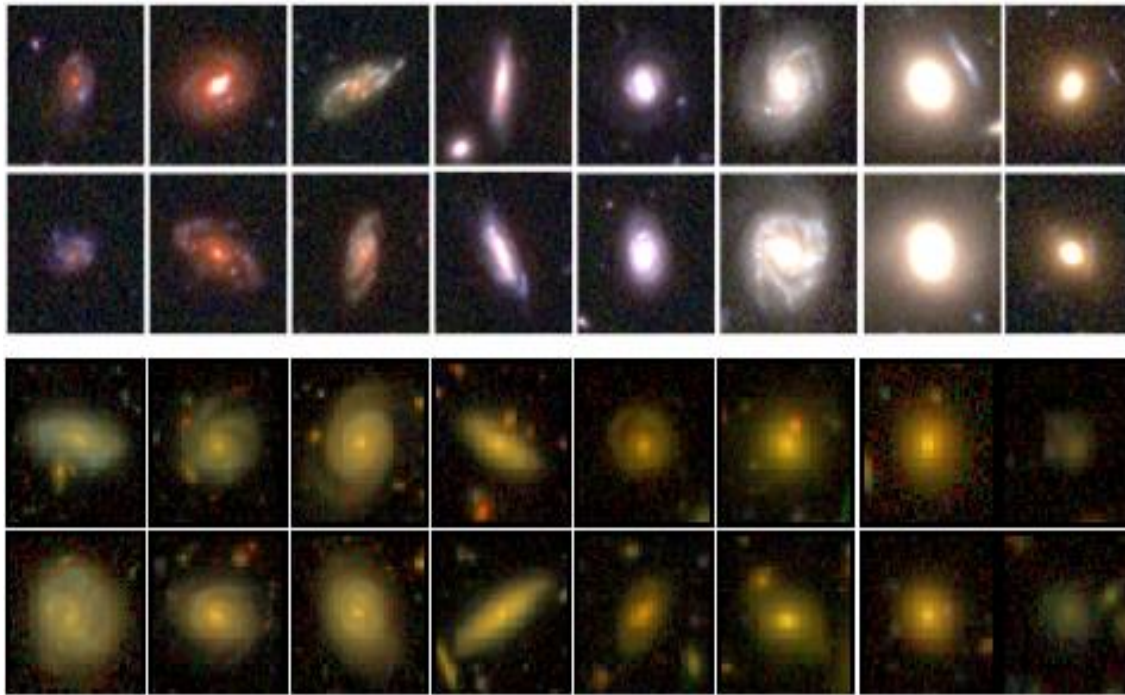
Sort galaxies into groups and create a dictionary of galaxies

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

ng,

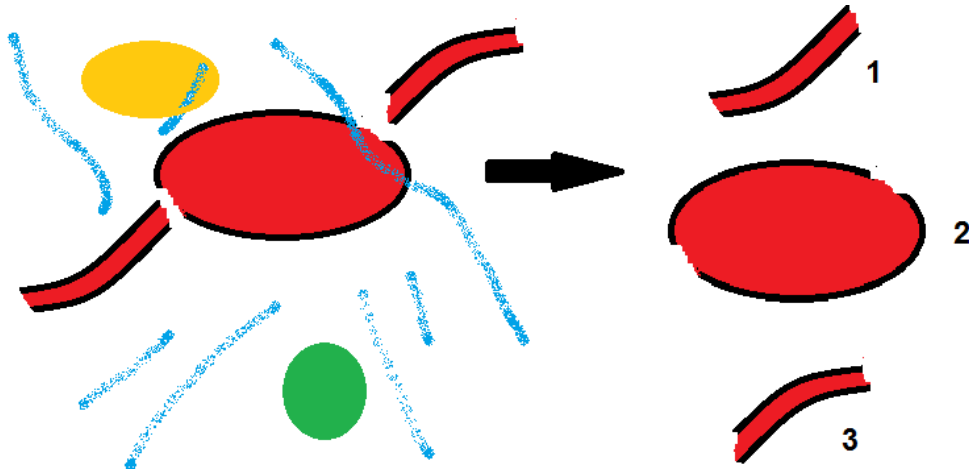
deep

yr)

raj +)

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)
 - Mitigating shredding (funded in Phase B)