



HAL
open science

Scientific Synergy Between LSST and *Euclid*

Jason Rhodes, Robert C. Nichol, Éric Aubourg, Rachel Bean, Dominique Boutigny, Malcolm N. Bremer, Peter Capak, Vincenzo Cardone, Benôit Carry, Christopher J. Conselice, et al.

► To cite this version:

Jason Rhodes, Robert C. Nichol, Éric Aubourg, Rachel Bean, Dominique Boutigny, et al.. Scientific Synergy Between LSST and *Euclid*. *The Astrophysical Journal Supplement*, 2017, 233 (2), pp.21. <10.3847/1538-4365/aa96b0>. <hal-01669484>

HAL Id: hal-01669484

<https://hal.science/hal-01669484v1>

Submitted on 2 Sep 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.













L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



Scientific Synergy between LSST and *Euclid*

Jason Rhodes^{1,2} , Robert C. Nichol³, Éric Aubourg⁴, Rachel Bean⁵, Dominique Boutigny⁶, Malcolm N. Bremer⁷, Peter Capak⁸, Vincenzo Cardone⁹, Benoît Carry¹⁰, Christopher J. Conselice¹¹ , Andrew J. Connolly¹² , Jean-Charles Cuillandre^{13,14}, N. A. Hatch¹¹, George Helou⁸, Shoubaneh Hemmati⁸ , Hendrik Hildebrandt¹⁵, Renée Hložek¹⁶, Lynne Jones¹² , Steven Kahn¹⁷, Alina Kiessling¹, Thomas Kitching¹⁸, Robert Lupton¹⁹ , Rachel Mandelbaum²⁰ , Katarina Markovic³, Phil Marshall¹⁷, Richard Massey²¹, Ben J. Maughan⁷, Peter Melchior¹⁹ , Yannick Mellier^{22,23}, Jeffrey A. Newman²⁴ , Brant Robertson²⁵ , Marc Sauvage¹³, Tim Schrabback¹⁵, Graham P. Smith²⁶, Michael A. Strauss¹⁹ , Andy Taylor²⁷, and Anja Von Der Linden²⁸ 

¹ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA; jason.d.rhodes@jpl.nasa.gov

² California Institute of Technology, 1201 East California Boulevard, Pasadena, CA 91125, USA

³ Institute of Cosmology & Gravitation (ICG), University of Portsmouth, Dennis Sciama Building Portsmouth, PO1 3FX, UK

⁴ APC, Astroparticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon & Léonie Duquet, F-75205 Paris Cedex 13, France

⁵ Department of Astronomy, Cornell University, Ithaca, NY 14853, USA

⁶ Laboratoire d'Annecy de Physique des Particules (LAPP), Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy, France

⁷ HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK

⁸ IPAC, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

⁹ I.N.A.F.—Osservatorio Astronomico di Roma Via Frascati 33—I-00040 Monte Porzio Catone (Roma), Italy

¹⁰ Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France

¹¹ School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK

¹² Department of Astronomy, University of Washington, Box 351580, U.W., Seattle WA 98195, USA

¹³ Laboratoire AIM, CEA/IRFU/Service d'Astrophysique, CNRS, Université Paris Diderot, Bât. 709, F-91191 Gif-sur-Yvette, France

¹⁴ Observatoire de Paris, PSL Research University, 61, avenue de l'Observatoire, F-75014 Paris, France

¹⁵ Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121, Bonn, Germany

¹⁶ Dunlap Institute for Astronomy and Astrophysics & the Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, Ontario, M5S 3H4, Canada

¹⁷ SLAC National Accelerator Laboratory, 2575 Sand Hill Road, MS29, Menlo Park, CA 94025, USA

¹⁸ Mullard Space Science Laboratory, University College London, Surrey, RH5 6NT, UK

¹⁹ Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

²⁰ McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

²¹ Durham University, Centre for Extragalactic Astrophysics, South Road, Durham DH1 3LE, UK

²² Institut d'Astrophysique de Paris, UMR7095 CNRS, Université Pierre & Marie Curie, 98 bis boulevard Arago, F-75014 Paris, France

²³ IRFU, Service d'Astrophysique, CEA Saclay, F-91191 Gif-sur-Yvette Cedex, France

²⁴ Department of Physics and Astronomy and PITT PACC, University of Pittsburgh, 3941 O'Hara St., Pittsburgh, PA 15260, USA

²⁵ Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

²⁶ School of Physics and Astronomy, University of Birmingham, Birmingham, B15 2TT, UK

²⁷ Institute for Astronomy, School of Physics and Astronomy, Royal Observatory Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK

²⁸ Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA

Received 2017 July 31; revised 2017 October 17; accepted 2017 October 18; published 2017 December 7

Abstract

Euclid and the Large Synoptic Survey Telescope (LSST) are poised to dramatically change the astronomy landscape early in the next decade. The combination of high-cadence, deep, wide-field optical photometry from LSST with high-resolution, wide-field optical photometry, and near-infrared photometry and spectroscopy from *Euclid* will be powerful for addressing a wide range of astrophysical questions. We explore *Euclid*/LSST synergy, ignoring the political issues associated with data access to focus on the scientific, technical, and financial benefits of coordination. We focus primarily on dark energy cosmology, but also discuss galaxy evolution, transient objects, solar system science, and galaxy cluster studies. We concentrate on synergies that require coordination in cadence or survey overlap, or would benefit from pixel-level co-processing that is beyond the scope of what is currently planned, rather than scientific programs that could be accomplished only at the catalog level without coordination in data processing or survey strategies. We provide two quantitative examples of scientific synergies: the decrease in photo- z errors (benefiting many science cases) when high-resolution *Euclid* data are used for LSST photo- z determination, and the resulting increase in weak-lensing signal-to-noise ratio from smaller photo- z errors. We briefly discuss other areas of coordination, including high-performance computing resources and calibration data. Finally, we address concerns about the loss of independence and potential cross-checks between the two missions and the potential consequences of not collaborating.

Key words: cosmological parameters – cosmology: dark energy – galaxies: clusters: general – gravitational lensing: weak – surveys – telescopes

1. Introduction

We present a broad overview of the potential synergies between the European Space Agency's *Euclid* mission²⁹

(Laureijs et al. 2011) and the Large Synoptic Survey Telescope³⁰ (LSST; Ivezić et al. 2008; LSST Science Collaboration et al. 2009). This work builds on the white paper by Jain et al. (2015), which started the discussion on the

²⁹ <http://www.euclid-ec.org>

³⁰ <http://www.lsst.org>

scientific and technical merits of combining LSST, *Euclid*, and the *Wide Field Infrared Survey Telescope* (*WFIRST*; Spergel et al. 2015). As this paper is a joint LSST and *Euclid* effort, we only focus herein on synergies between these two facilities and continue to emphasize the scientific and technical gains from combining these data. We do not attempt to provide an exhaustive review of all possible gains; we leave that for others to continue to explore as well as to develop more quantitative analyses that are beyond the scope of this paper. We assume that future work will require some “figures of merit” to help optimize these gains, especially to help in the allocation of (limited) resources.

We have not attempted to broaden this paper beyond LSST and *Euclid* to include other international projects including *WFIRST*, the Dark Energy Spectroscopic Instrument (DESI),³¹ or the Square Kilometre Array (SKA).³² Other authors have begun to explore such synergies, and we refer the reader to those other papers (e.g., Bacon et al. 2015; Jain et al. 2015; Kitching et al. 2015).

1.1. Dark Energy Experiments

By the late 1990s, there was definitive evidence that the universe was in a phase of accelerated expansion. The unknown cause of this acceleration was dubbed dark energy. At the start of the millennium, it became clear that more data were required to determine the nature of dark energy. At the same time, new ways to analyze astronomical data from large wide-field surveys were being developed, in particular weak gravitational lensing and Baryon Acoustic Oscillations (BAO) observed via galaxy clustering. Techniques for classifying supernovae also continued to advance. The diversity of this field at this time has been captured in the ESO-ESA Report on dark energy (Peacock et al. 2006) and the Dark Energy Task Force (DETF) reports (Albrecht et al. 2006, 2009).

What evolved from this intense period of discussion of new experiments was an international program that, over the two decades from 2010 to 2030, will observe successively larger areas of the sky at wavelengths from the optical to infrared (and eventually, radio), from both the ground and space. It was realized early on that future dark energy experiments should exploit the combination of multiple cosmological techniques as this would lead to more secure dark energy measurements, tighter control of systematics, and the ability to distinguish between dark energy and possible modifications to general relativity (modified gravity). Thus, both *Euclid* and LSST have evolved through a process of niche separation and complementarity to existing proposals to encompass a unique combination of methodologies and sensitivity to systematic effects.

Each experiment was designed to simultaneously minimize statistical error and systematics for dark energy studies, while finding maximum complementarity with existing, ongoing surveys. The resulting projects approximately follow a “staged” classification proposed by the DETF report, whereby early experiments designed to confirm the existence of dark energy are classified as “Stage II,” mid-cycle experiments that can determine cosmological parameters within our current paradigm to high accuracy and test systematic effects are “Stage III,” and nearly all sky experiments that have the capability to

determine the physical nature of dark energy are “Stage IV.” The majority of Stage II experiments are now complete, proving the existence of dark energy beyond doubt for a majority of cosmologists (Weinberg et al. 2013).

Stage III experiments are now coming to maturity, with surveys including the Kilo Degree Survey (KiDS; Hildebrandt et al. 2017), the Baryon Oscillation Spectroscopic Survey (Alam et al. 2015), the Subaru HyperSuprimeCam Survey (HSC; Aihara et al. 2017), and the Dark Energy Survey (DES; DES Collaboration et al. 2017) delivering new results. These latest surveys still highlight possible tensions within our present cosmological model including ongoing differences between low- and high-redshift measurements of the Hubble constant (Joudaki et al. 2017b and Bonvin et al. 2017) and the growth of structure (Kitching et al. 2016; Efstathiou & Lemos 2017; Hildebrandt et al. 2017). These differences could be due to unresolved systematic uncertainties that would need to be understood to allow us to make further progress, or to the first signs that our incomplete cosmological understanding may require new physics.

There are now several Stage IV experiments under construction, including *Euclid* and LSST. The methods used by these experiments all drive the experiment design toward large wide-field surveys, ideally with high spatial resolution and high cadence of observations. Neither *Euclid* or LSST individually satisfy all of these criteria, but together, they do. The sum of all data from these two facilities will survey the majority of the extragalactic sky, with high spatial resolution, rapid cadence, from ground and space, and in multiple filters from blue optical to near-infrared (NIR). This abundance of new survey data will change astronomy in ways that go far beyond dark energy and follows the rich history of other surveys like the Sloan Digital Sky Survey (SDSS; York et al. 2000) where the access to high-quality public data has promoted a wealth of legacy research (e.g., thousands of journal papers) and public engagement opportunities (e.g., GalaxyZoo³³; Raddick et al. 2010; Lintott et al. 2011).

Euclid and LSST will each survey a significant fraction of the sky starting at the beginning of the next decade. We summarize each of these facilities below and provide more details about the intended surveys in Section 4.4.1:

1. LSST is a ground-based telescope with a primary mirror with an effective diameter of 6.7 m designed to provide a time-domain imaging survey of the entire southern hemisphere ($\simeq 18,000 \text{ deg}^2$) in six optical bands (*ugrizy*) with an average seeing of 0.7 arcsec. First light is scheduled for 2020, with survey operations commencing in 2022 and running for the following 10 years to gradually accumulate depth to $r \sim 27.5$ magnitudes (5σ point source). The science objectives of LSST are multifold, but among the primary scientific drivers is the elucidation of the nature of dark energy and dark matter through a combination of probes associated with statistical analyses of billions of galaxies and hundreds of thousands of supernovae.
2. *Euclid* is a 1.2 m space-based telescope designed to image $15,000 \text{ deg}^2$ of sky in one broad optical band ($\text{VIS} = r + i + z$ covering approximately 5400 to 9000 Å) with pristine image quality ($0''.16$ FWHM) to an expected depth of at least 24.5 AB magnitudes ($10\sigma 1''$

³¹ <http://desi.lbl.gov/>

³² <http://skatelescope.org/>

³³ <http://www.galaxyzoo.org>

extended source) and three near-infrared bands covering approximately $0.95\text{--}2\ \mu\text{m}$ down to $Y = 24.0$, $J = 24.2$, and $H = 23.9$ AB magnitudes (5σ , point source), respectively. *Euclid* will also obtain near-infrared low-resolution ($R \sim 250$) grism spectroscopy over the same area of the sky using a “red” grism ($1.25\text{--}1.85\ \mu\text{m}$). *Euclid* is officially scheduled for launch in late 2020, with science operations of a least 6.5 years. Although *Euclid* data will also address a range of scientific questions, the primary goal of the mission is the elucidation of the nature of dark energy (along with possible modifications to general relativity) through investigations of the distance–redshift relation of galaxies and the evolution of cosmic structures.

With these features, it is clear that *Euclid* and LSST are complementary. Consequently, as the scientific goals and schedule overlap, it is natural to ask whether cooperation between the two experiments can lead to both increased efficiency and scientific reach. Indeed, the limitations and risks faced by one are mitigated by the strengths of the other e.g., LSST observes the sky from the ground, while *Euclid* has a finite lifetime and constrained cadence due to the degradation of the CCDs in the space environment and limited propulsion gas needed to point the telescope.

This document aims to outline the benefits of scientific cooperation, mainly at the level of survey coordination, information interchange, and joint data processing, as well as to provide a technical framework for how the coordinated efforts might be arranged. We also seek to identify potential scientific risks associated with such coordination that result from the loss of independent analyses that would occur due to sharing the data and associated tools (see Appendix A.1 for a discussion of the merits of independence). This document should inform future discussions between the *Euclid* Consortium (EC), the LSST Project, and the associated LSST science collaborations about how to achieve maximal joint cosmological science in the next decade.

The paper is organized as follows. Section 2 outlines the benefits of possible coordination for cosmology in more detail, and some non-cosmology science and technical cases for coordination are outlined in Section 3. Section 4 describes the technical level of coordination required, and Section 5 describes several models in which we might achieve that coordination. What we still need in order to achieve the described coordination is described in Section 6, and we offer concluding remarks in Section 7. A discussion of the pros and cons of coordinating two experiments, and a warning about what a lack of coordination might mean, is included in the Appendix.

2. Benefits for Cosmology

In this section, we detail some of the benefits of coordination for cosmological science. Benefits to non-cosmological science are detailed in Section 3.3.

2.1. Object Detection

The optical images from LSST and *Euclid*’s VIS instrument differ in three main characteristics: spatial resolution, number of filters, and depth. *Euclid*’s Near Infrared Spectrometer and Photometer (NISP) differs from LSST and VIS by covering the near-infrared (NIR) regime, and by being significantly

undersampled. These differences can be exploited to optimize object detection and characterization.

The *Euclid* VIS instrument has a spatial resolution of $0''.1$ per pixel with a $\leq 0''.18$ width of the point-spread function (PSF; Cropper et al. 2016). This is compared to the anticipated median seeing for LSST of $0''.7$. This difference in resolution can enable the recognition of objects that would be unresolved by LSST, improve the separation of stars from galaxies, and improve the deblending of objects. This works for objects that are brighter than $\simeq 25$ AB magnitudes, the present design limit of VIS for a $1''$ extended source. LSST will have the advantage of a point-source detection depth of ~ 27.0 AB magnitudes in the *i*-band, which is superior for the detection of fainter objects and extended emission, e.g., low surface brightness galaxies and intracluster light (especially at wavelengths shorter than the *Euclid* VIS filter band). *Euclid* NISP will extend the wavelength range from 900 nm to $2\ \mu\text{m}$, but at a lower spatial resolution (compared to VIS) of $0''.3$ per pixel.

Euclid and LSST are thus strongly complementary for the key decision of what constitutes a valid celestial object (in the areas where the two surveys overlap). Two approaches for combining these data sets, with different levels of independence, are:

1. Catalog-level combinations. For both surveys, objects would be detected independently and then matched. The *Euclid* VIS object catalog can be used to improve and calibrate the star–galaxy separation methods in LSST. To identify so-called “ambiguous,” or unrecognized, blends (Dawson et al. 2016), for which the number of nearby overlapping objects cannot be determined in ground-based images, one can make use of both *Euclid* imagers: VIS can spatially resolve peaks with separations of $\approx 0''.3$, and NISP can distinguish between objects with different NIR Spectral Energy Distributions (SEDs).

LSST object catalogs can provide confirmation of objects at the detection limit of *Euclid* and identify faint and diffuse emission of objects whose full extent was not visible in the *Euclid* images. Furthermore, requiring corroborating detections across both surveys suppresses the contamination of spurious objects in the static catalogs, while providing additional time-domain information for transient and variable objects.

Beyond confirming objects, collaboration on catalogs can provide the opportunity to carry out “forced photometry” in both surveys, namely, measuring the flux in one survey based on the aperture of a source detected in the other survey. For resolved sources, a common aperture is not sufficient, and some form of modeling or PSF matching is necessary. For LSST, the VIS catalog could be used to measure the deblended flux in crowded fields based on the high-resolution *Euclid* data.

2. Pixel-level combination. Object detection and characterization could be performed simultaneously on images from both surveys. This is particularly important for separating blended sources, which will be about 40% of all LSST objects (Dawson et al. 2016). In contrast to catalog-level detection of blends, model-based or non-parametric deblending methods can benefit from a joint data set because of the extended wavelength range and resolution, enabling separation schemes that exploit color and morphological information. Even for isolated objects, higher-level characterizations, such as photometry or

shear measurements (Sections 2.2 and 2.3), can be performed at higher significance and fidelity when models are fit across several bands of both surveys, potentially employing priors on galaxy colors, sizes, and morphologies.

These two approaches are not mutually exclusive. On the contrary, it appears prudent to perform the catalog-level combination first to establish whether the two surveys find consistent results for isolated objects. Although this step will be limited to brighter objects, these are also the objects with the most precise shape and flux measurements. Any inconsistencies (e.g., from image calibrations or the deconvolution from the PSF) should be seen as offsets in astrometric position, magnitude zeropoint, or ensemble lensing shear estimate from a matched galaxy sample. These comparisons will establish whether common definitions, for detection thresholds or photometric apertures, are being used, or how to convert measurements between the surveys. Such information will be vitally important for the wider community when they start using the public data products.

An expected source of inconsistency is due to unrecognized blends, which may show up as a single detection in LSST but may have several detections in *Euclid* VIS (or NISP). Furthermore, the NISP NIR centroids may not correspond to the LSST centroid derived from the optical bands alone. Passing along these measurements and their errors will enable downstream data processing and science algorithms to decide how to best characterize the blended group.

Once cross-survey inconsistencies are identified (if present), a joint pixel-level analysis could extend the source catalogs in those domains that are difficult for one of the surveys but not for the other. This would, of course, require the sharing of pixel-level data from *Euclid*'s VIS and NISP instruments and the multiband LSST images (see, for instance, Section 5.5).

Both approaches for joint object detection would be made considerably easier if data processing conventions have been coordinated from the beginning to share a common coordinate system and photometric standards (e.g., from *Gaia*). While both *Euclid* and LSST will likely use their own internal photometric system, there are obvious advantages to establishing a common set of photometric standards and making sure to exchange the standard photometric measurements between projects from an early stage. There are obvious cosmological gains from improved galaxy deblending, e.g., better shape and photometry measurements that result in higher surface densities of galaxies useful for lensing, leading to improved statistical power and lowered systematics in dark energy studies. Beyond dark energy constraints, better deblending and the resulting higher surface density of lensed galaxies may also lead to a better understanding of the small-scale power spectrum, which allows for better constraints on the role of baryons, the nature of dark matter, and the neutrino masses.

2.2. Photometric Redshift (Photo- z) Estimation

A variety of cosmological measurements (most notably weak lensing) from both projects will require accurate knowledge of the redshift distribution of the source galaxies when split into tomographic redshift bins. In the cases of LSST and *Euclid*, the source galaxy samples will consist of billions of galaxies, and redshifts can only be obtained using photometric techniques. The resulting redshift estimates, generally referred to as

photometric redshifts or photo- z 's, benefit from a large coverage in wavelength. The availability of a greater number of photometric passbands spanning a broader wavelength range will yield photo- z 's with higher precision; tighter redshift distributions will also be easier to calibrate.

LSST would benefit from the addition of the *Euclid* NISP-photometric (NISP-P) passbands (Y , J , H) as they will extend wavelength coverage into the NIR, thereby breaking degeneracies between redshift and the intrinsic SED of galaxies (Benítez 2000; Hildebrandt et al. 2010). *Euclid* depends crucially on optical multiband photometry from the ground. LSST will be superior to all alternatives in the southern hemisphere in terms of depth, number of photometric passbands, image quality, and area coverage.

The combination of data for photo- z 's could be achieved at the catalog level. However, the photo- z 's from both projects would be improved via a joint pixel-level analysis as described in Section 2.1, as the effect of untreated blending on photo- z estimates is generally difficult to characterize. In addition, the photo- z probability distribution $p(z)$ for each object depends on the colors available and how they are measured (e.g., the relative weighting of the bulge and disk light in each galaxy).

A critical issue for both surveys will be calibrating photometric redshift algorithms, e.g., quantifying the error in the measured distribution of photo- z 's (see Section 3.1). Different techniques have been proposed to tackle this problem, including the use of a large, representative spectroscopic calibration sample (Lima et al. 2008; Masters et al. 2015) as well as using angular cross-correlations between the samples of interest and galaxies with known redshifts (Newman 2008; Rahman et al. 2015). All of these calibration techniques benefit from additional wavelength coverage of the photo- z and yield smaller errors when photometric redshift scatter is reduced (Newman 2008; Hearin et al. 2010).

For LSST, the addition of the *Euclid* NISP-P passbands would significantly reduce photo- z uncertainties at $z > 1.5$. Likewise, LSST has the capability of covering two-thirds of the *Euclid* survey footprint at greater depths than needed to achieve the required photometric redshift precision for *Euclid* (see Section 4.4). *Euclid* operates with a fixed integration time in all bands, which will lead to non-uniform depths across its footprint due to the zodiacal light background (in the range of 0.6 mag). While the *Euclid* northern footprint will be covered by various telescopes at the depth needed to match the dark energy constraint driven requirement of 30 source galaxies per square arcminute ($VIS = 24.5$), LSST has the unique potential of producing photometry going significantly deeper than the depth required by *Euclid*.

Another advantage of using LSST, as opposed to a broader set of telescopes, is the relative uniformity of the PSF and properties of the data (including the data processing and calibration) across the passbands, facilitating a more homogeneous photometric calibration. Finally, LSST will perform many more visits per area than needed by *Euclid*, allowing for the selection of the best subset for photometric uniformity (e.g., based on the homogeneity of the PSF, delivered image quality, and atmospheric transparency). The additional u -band from LSST would also significantly reduce photo- z uncertainties at low redshift for *Euclid*, which is useful for a proper treatment of a number of photo- z related systematics including galaxy intrinsic alignments (for a recent review of intrinsic alignments, see Joachimi et al. 2015).

2.3. Shear Measurement

Cosmological weak-lensing measurements require the ability to accurately infer the ensemble statistics of the shear field, i.e., the weak-lensing distortion as a function of position, and its spatial correlation. Typically, the inference of the shear field is based on estimating the weighted averages of individual galaxy shapes (for some definition of shape whose ensemble average has a well-defined relationship to the shear). The main challenge is to deconvolve the observed galaxy shapes from the influence of the PSF, which can vary temporally, spatially, and as a function of wavelength.

For weak-lensing measurements, three levels of coordination are advantageous. First, the cross-correlation of the shear catalogs from the different surveys would help us expose and understand any residual survey-specific systematics. Second, matching the object catalogs of both surveys, and then comparing the shear signals around a common set of foreground lenses (e.g., clusters) could reveal differences in the relative shear calibration. Finally, a joint pixel-level measurement would be more robust against the problem of objects blending together (see, e.g., Section 2.1), which is beneficial for LSST because of the wider PSF. Also, this could account for color variations within galaxies, which needs to be accounted for with *Euclid* due to the single VIS filter for shape measurements (Voigt et al. 2012; Semboloni et al. 2013).

Methods to measure the ensemble weak gravitational lensing shear distortions from PSF-corrected galaxy shapes typically need to be calibrated on simulated images containing a known shear (Hoekstra et al. 2017). Although advances have been made to self-calibrate using shear methods based directly on the survey data (Huff & Mandelbaum 2017; Sheldon & Huff 2017), additional test data need to be generated to independently verify the accuracy of the calibration (Hoekstra et al. 2017). Shear estimation methods that have been devised to avoid shear calibration biases that most shear estimation methods suffer by design (e.g., Bernstein et al. 2016) require deep data to build their priors, and still require realistically complex image simulations to confirm that they are truly unbiased, similar to the needs of self-calibration methods. Creating such realistic simulations of complex galaxy morphologies requires images of the true sky with higher resolution and that are deeper than the survey images. Calibrating the shear measurements of *Euclid* could require an appreciable fraction of the *HST* archive to statistically account for the impact of color gradients in galaxies (Voigt et al. 2012; Semboloni et al. 2013).

Calibrating shear measurements for LSST will require a similar data set, like the *Euclid* imaging or the *Euclid* shear measurements, especially from the *Euclid* deep fields, which will be of comparable depth to the LSST lensing sample. Calibrating multiplicative bias in shear measurements can potentially be done via external measurements of the lensing of the cosmic microwave background, or possibly LSST supernovae (but only at the level of a few percent; Schaan et al. 2017). This would be sufficient for some, but not all, of the tomographic redshift bins of *Euclid* and LSST; the development of robust shape measurement methods will need to continue in order for both surveys to reach their shape measurement systematics requirements.

2.4. Cluster Mass Estimates

LSST and *Euclid* will constrain cosmological parameters using galaxy clusters, primarily measuring the shape and evolution of the halo mass function. To harness the power of this approach will require a precise and accurate cluster mass calibration (Dodelson et al. 2016). Cluster weak-lensing measurements are clearly the best opportunity to deliver the required mass estimates for ensembles of clusters (Hoekstra et al. 2015; Mantz et al. 2015; Krause & Eifler 2017). The combination of *Euclid* and LSST data will also help with cluster detection.

For accurate cluster masses from weak lensing, it is essential to separate the lensed background population from foreground galaxies, especially cluster galaxies. The combination of LSST and *Euclid* provides more robust photometric redshift estimates, allowing one to minimize the roles of external priors and calibration. From a *Euclid* perspective, the combination with multiband, deep optical photometry is essential for cluster mass estimates at all redshifts in order to identify background galaxies. From an LSST perspective, the addition of NIR photometry enables precise photo- z 's beyond $z \sim 1.4$, and thus an increased density of background sources with secure redshifts for clusters at $z \gtrsim 1$.

The combination of LSST high signal-to-noise ratio (S/N) optical photometry and *Euclid* space-based shape measurements is well suited to push weak-lensing studies to higher cluster redshifts $z_{\text{cluster}} \simeq 1-1.3$. The required background source galaxies at $z \gtrsim 1.5$ are still well resolved in the *Euclid* VIS data for shape measurements. Even if individual galaxy photo- z 's may be too noisy, the faint high-redshift tail of the expected galaxy distribution can still be selected from deep LSST colors (and calibrated on the deep fields) with low to moderate contamination from the foreground and cluster galaxies (see Figure 1 and Schrabback et al. 2017). This strategy is particularly effective (i) at high ecliptic latitudes, where the *Euclid* data are deeper because of lower zodiacal background, (ii) where deep photometry is available from LSST, and (iii) if new shape measurement techniques, which yield robust results for galaxies with lower S/N, are employed (e.g., Bernstein et al. 2016).

The combination of LSST and *Euclid* data will also play an important role in selecting clusters based on photometry, for example, with the redMaPPer algorithm (Rykoff et al. 2014), especially toward higher redshifts (also see Section 3.3.4). This will make the photometric data from *Euclid* and LSST especially synergistic with other surveys across the electromagnetic spectrum, e.g., X-ray (*eROSITA*), and Sunyaev-Zel'dovich (SZ) surveys like AdvACT (Henderson et al. 2016) and SPT-3G (Benson et al. 2014).

2.5. Transients and Variable Sources

The transient survey of LSST is one key observational scientific difference with *Euclid*. The *Euclid* wide survey will be covered in only a single visit, making it inadequate for transient searches (the *Euclid* deep fields will have multiple visits over six years). However, transient and variable source searches could still be enhanced through coordination. For example, cosmological analyses based on Type Ia supernovae (SNe Ia) from LSST would benefit from NISP data on the host galaxies, e.g., *Euclid* could provide spectroscopic redshifts, via NISP spectroscopy (NISP-S), for a fraction of LSST SNe Ia

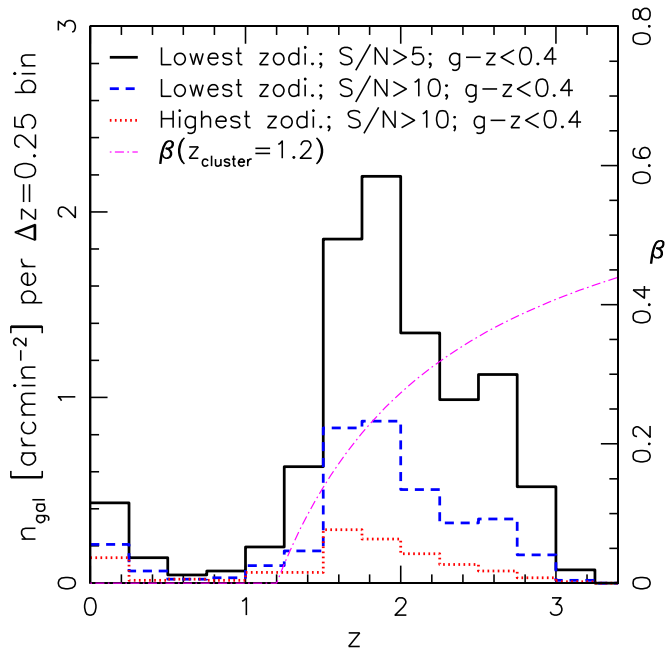


Figure 1. Weak-lensing studies of clusters at $z_{\text{cluster}} \simeq 1\text{--}1.3$ require source galaxies at redshifts $z \gtrsim 1.5$, where the lensing efficiency β (shown in magenta for a cluster at $z_{\text{cluster}} = 1.2$) is high. Optical color selection from deep LSST photometry (here $g - z < 0.4$) is sufficient for selecting most of these distant background galaxies (see also Schrabback et al. 2017). The plotted histograms show the expected source density for *Euclid* based on CANDELS/COSMOS photometric redshifts (Skelton et al. 2014). While the source density is low at low ecliptic latitudes with high zodiacal background (red dotted histogram, $|b| = 15^\circ$), it increases at high ecliptic latitudes, where the *Euclid* data are deeper (blue dashed histogram). Compared to the default analysis using galaxies with a VIS signal-to-noise ratio $S/N > 10$, a further $\times 2.6$ boost in the high- z source density (solid black histogram) could be achieved with advanced shape measurement techniques which yield robust results for galaxies with $S/N > 5$ (e.g., Bernstein et al. 2016), in areas where deep LSST photometry is available for the color selection (see Section 4.4.1).

host galaxies, helping to remove possible systematic biases when using photometric redshifts alone (Olmstead et al. 2014).

Beyond improving redshifts, the *Euclid* data would help improve the classification and photometry of LSST transients. First, *Euclid* could provide serendipitous NISP spectroscopy for many active (long-lived) LSST transients, e.g., superluminous supernovae (Inserra 2017). The low resolution of the *Euclid* grisms are still sufficient to resolve the broad spectral features we expect from these objects and thus provide a definitive classification for a wide variety of serendipitous objects. Such a random (“unbiased”) sample provides an excellent training set for the LSST transient classifiers and is complementary to planned, dedicated spectroscopic surveys of LSST transients. This will be particularly relevant in the first few years of overlap, when LSST will be spending most of its supernova-specific survey time building up a library of light curve templates with which to perform classifications later. This will be timely in that *Euclid* will have been operating for long enough that systematics should be under control and the coordination of fields, and/or including a trigger for follow-up, will be of great use in improving classifications.

Moreover, for SNe Ia, even a few serendipitous NIR data points overlapping with the LSST light curve (NISP-P and NISP-S) could improve the model fitting of these joint data, thus producing better distance estimates for a subsample of

LSST SNe. That is, SNe Ia are better “standard candles” in the NIR (Wood-Vasey et al. 2008). It is well-known that SNe Ia have a second light curve maximum in the NIR, thus extending their lifetime (in the observer frame) and increasing the opportunity for serendipitous detections by *Euclid*.

Second, the high-resolution *Euclid* VIS imaging will provide important extra information about the host galaxies. Such information can help better inform the difference imaging required to detect transients, e.g., by convolving the higher-resolution *Euclid* image to match the lower-resolution LSST template. This would not be a simple task (due to the different passbands and PSFs, and the chromatic effects from the wide *Euclid* PSF), so a coordinated effort between projects would be needed to unlock this science benefit. Moreover, there is evidence that SNe Ia can be improved as “standard candles” by correcting their peak magnitudes by the local star formation rate (Kelly et al. 2015; Rigault et al. 2015) at the location of the SN within the host galaxy (e.g., by using local measurements of the galaxy colors enabled via the high-resolution *Euclid* imaging from VIS and NISP). Likewise, the extra information on the overall color and morphology of the host galaxy will be important for making additional corrections to the standardization of SNe Ia (e.g., Lampeitl et al. 2010).

Finally, there are interesting cosmological synergies between the shear weak-lensing measurements from *Euclid* and LSST, and the magnification weak-lensing measurements possible from the LSST supernova sample (Scovaccicchi et al. 2017). Such comparisons provide interesting systematic checks of the multiplicative bias, as well as possibly improving the signal from the combination of these measurements; for example, Scovaccicchi et al. (2017) predict that the LSST SNe Ia–galaxy spatial cross-correlation function can be detected to high significance and can be used to provide additional constraints on cosmological parameters. It will be interesting to measure the spatial cross-correlation function of LSST SNe and *Euclid* spectroscopic galaxies as any systematic uncertainties would be different between these two surveys.

The LSST supernova survey will allow for novel tests of cosmology, including tests of the isotropy and bulk flows. Using bright tracers as a probe of the underlying structure and structure formation requires both an accurate estimate of the distance modulus (which will be improved with the NIR addition from *Euclid*) and from more precise host galaxy redshifts, which *Euclid* can provide. Work is underway to test the various cadence strategies of LSST on these novel probes, but coordination with *Euclid* will enable this novel science with both telescopes.

The coordination discussed in this section applies to both the “wide” and “deep” fields planned for both experiments. For the wide surveys, this coordination may happen naturally, although care must be taken to ensure that the complementary survey strategies remain the default. However, more direct, active coordination may be required to ensure the LSST and *Euclid* deep fields overlap both in area and time. Moreover, LSST and *Euclid* could consider synchronization of the observation of common deep fields, thus maximizing the number of unique, well-spaced epochs. Such an active coordination would greatly benefit a range of time-domain studies, e.g., reverberation mapping of active galactic nuclei (AGNs; see Bentz et al. 2009).

3. Other Areas of Coordination

In this section, we explore the benefits of coordination that go beyond the cosmological goals outlined in the previous section.

3.1. Spectroscopic Training Sets

Euclid and LSST are exploring two routes toward spectroscopic calibration of their photometric redshifts. One method is to calibrate the color–redshift relation using a spectroscopic sample of galaxies with an extremely high completeness; see Masters et al. (2015), who explored this strategy for *Euclid*. This method is also explored in Speagle & Eisenstein (2017a, 2017b). The alternative is to measure the angular cross-correlations between the photometric samples and large, wide-area spectroscopic surveys. The true redshift distribution for the photometric sample is then reconstructed from this cross-correlation (Newman et al. 2015; Rahman et al. 2015). The requirements on the spectroscopic samples needed by both LSST and *Euclid* will be similar because of their overlap in area, redshift range, survey depth and calibration accuracy required.

For the color–redshift method, *Euclid* and LSST have similar requirements on sample purity, quality, and knowledge of the selection function. At present, the Complete Calibration of the Color–Redshift Relation (C3R2; Masters et al. 2017) project aims to be complete to $RIZ < 25$ using spectroscopic data from Keck and VLT (located in several LSST deep fields including CDFS, SXDS, and COSMOS). This should be sufficient for the *Euclid* wide survey, which will reach to $RIZ < 24.5$ AB magnitudes. Although not as deep as the LSST lensing sample of $i < 25.3$, this *Euclid* training set will cover a significant fraction of the LSST color–color space for galaxies, so that a smaller number of additional spectra will be needed to calibrate LSST.

Based on previous studies, Newman et al. (2015) estimate that $\sim 20,000$ spectra in total would be needed to calibrate the LSST color–redshift relation down to the LSST weak-lensing imaging depth. The planned C3R2 sample is comparable in size, and the cosmic variance across six distinct 1 deg^2 fields (as envisioned for C3R2) would be comparable to that from the fifteen 0.1 deg^2 fields established as the LSST requirement in Newman et al. (2015). Therefore, C3R2 should be sufficient for both surveys.

For calibration via cross-correlations, both LSST and *Euclid* will have similar requirements on the total sky area, spectral density, and redshift range of the surveys. As described in Newman et al. (2015), LSST calibration using this technique will require at minimum 100,000 objects distributed over multiple widely separated fields of several hundred square degrees, spanning the full redshift range of the weak-lensing samples used for dark energy analyses. *Euclid* requirements have been estimated to be $0.4 \text{ galaxies deg}^{-2}$ per $\Delta z = 0.05$ bin for $0 < z < 6$ over an area of at least 2000 deg^2 , corresponding to 96,000 galaxies. Such samples should be available from overlap with the planned DESI and 4 m Multi-Object Spectroscopic Telescope (4MOST³⁴) galaxy redshift surveys, where *Euclid* NISP-S can fill in the “redshift desert” (at $z = 1.5\text{--}2.5$) which is difficult from the ground. Moreover, DESI and 4MOST should have ancillary programmes allowing

the targeting of hundreds of thousands of targets across the overlap of these many surveys (e.g., the intersection of LSST, *Euclid*, 4MOST and DESI).

More work is required to develop robust cross-correlation methods and demonstrate that they can in fact reach the calibration requirements of these dark energy experiments; the needs are similar for both *Euclid* and LSST. Both methods require extensive validation, pre-processing, and metadata tracking for the spectroscopic samples. For example, the process of validating the quality of spectra obtained will be a common problem for the two surveys. It would therefore be wise to proceed in partnership and include the planned wide-field spectroscopic surveys in this effort. Cross-checks will occur naturally within each consortium, but overall coordination will reduce the total demand for external data and provide a means of quickly validating results.

3.2. Numerical Simulations and Supercomputing

LSST and *Euclid* will each require extensive numerical simulations for a range of tasks including forecasting, developing analysis techniques, and the eventual analysis of the observational data. Carrying out the simulations, transforming them into synthetic sky maps, validating the results, and serving the data in an easily accessible way are all major efforts; they require large computing and storage resources in addition to a sufficient workforce to develop the modeling and analysis pipelines.

Quantifying the supercomputing resources required for numerical simulations for each survey is a challenging process. LSST has around 30 expert members that have been active in defining cosmological software and hardware infrastructure requirements. Two reports have been delivered that define the computational requirements for LSST cosmology for the years 2016–2022. However, these reports specifically do not include numerical simulations in their estimates, so further effort is needed to define the needs in this area.

The future large-scale computing needs for DOE High Energy Physics (HEP) have been articulated in a joint DOE Advanced Scientific Computing Research (ASCR)/HEP Exascale Requirements Review Report (Habib et al. 2016). The report contains an analysis of HEP Cosmic Frontier simulation and data analysis computing requirements for the years 2020 to 2025. The annual simulation requirements for all HEP cosmology experiments are expected to be 100 billion to one trillion (current) core hours by 2025 on ASCR High-Performance Computing (HPC) systems, approximately a quarter of the projected total HEP requirement. Although the three main ASCR facilities are expected to support this requirement, possible increases in requirements over a time-scale of 5–10 years will necessitate periodic re-evaluations.

The number, resolution, and contents of the numerical simulations that will eventually be needed for LSST and *Euclid* remain an area of study (e.g., Heavens et al. 2017; Sellentin & Heavens 2016, 2018). It is possible that the simulations required for covariance matrix calculations could be extremely challenging, with up to a million independent realizations needed per covariance matrix using a brute force approach (Taylor et al. 2013). Although there are active investigations underway to find more intelligent methods for producing covariance matrices (e.g., Friedrich & Eifler 2017), this remains an open question for both LSST and *Euclid*. It is anticipated that even with orders of magnitude reduction in the number of

³⁴ <https://www.4most.eu/>

simulations required for one covariance matrix, the total number needed for all covariance matrices will still demand significant computational resources.

Moreover, the surveys also require computationally expensive large, high-resolution N -body and hydrodynamic simulations for synthetic sky maps used for end-to-end testing of analysis pipelines and investigating systematic effects at sufficient fidelity for these Stage IV experiments, e.g., determining the impact of baryons on the weak-lensing power spectra using state-of-the-art hydrodynamic simulations. Although the computational requirements for running these simulations and storing their outputs are commonly accounted for, the resources required to perform detailed analysis, and to store the resulting data products, are not as readily acknowledged. Given the likely scale of the numerical simulations and analysis needed for the cosmological science of LSST and *Euclid*, it would be sensible to coordinate efforts to define common numerical simulations to ensure that as much of these simulated data can be used by both surveys. This may require slight modifications to some simulation requirements, but a modest overhead that enables sharing would be preferable to each survey trying to acquire the resources to run, store, and analyze all numerical simulations separately.

Such coordination would build on existing efforts already underway within the surveys and within various countries. Initial efforts were made toward formulating a plan for what numerical simulations are required as a function of time within each consortium. This work will be valuable within *Euclid* and LSST for planning purposes and should form a template for coordination between projects.

The United Kingdom Tier Zero (UKT0) group has brought together scientists from a wide range of areas in astronomy and particle physics (including gravitational waves). The primary goal of this group is to explore ways to collaborate on High-Performance Computing and High Throughput Computing. In particular, one experiment will utilize the Large Hadron Collider grid computing in the context of *Euclid*, LSST, and SKA studies to determine if this is an effective computing solution for these surveys. Likewise, on the US side, there is a Tri-Agency (NASA, NSF, DOE), Tri-Project (*Euclid*, LSST, *WFIRST*) Group (TAG) that has US representatives from each of the agencies and projects. In 2015, the TAG commissioned a US-based task force with representation from each of the projects to write a report on the current state of cosmological simulations within the projects and identify benefits of coordination. The report submitted in 2016 March noted that coordination is certainly beneficial in many areas, but the significant work effort required to produce detailed plans would require a separate task force to be formed. In early 2017, this led to the creation of a Tri-Agency Cosmological Simulation (TACS) task force, formed at the request of the US project leads for *Euclid*, LSST, and *WFIRST*. The task force has representation from US and European *Euclid* members in addition to representation from members of LSST and *WFIRST*. TACS will investigate the logistics of coordinating hardware between the agencies to enable a more effective supercomputing infrastructure from running the simulations through to their analysis, storage, and sharing with a wider community. The TACS will investigate coordination opportunities for flagship simulations, lower-resolution simulation suites, and synthetic sky generation. Perhaps most importantly, TACS will investigate methods for reducing overall computing

and storage requirements for simulations and whether simulation modeling efforts for a wide range of observational systematics are sufficient or need further work.

There is broad scope for coordination between *Euclid* and LSST for numerical simulations and many avenues are currently being explored. The significant resources required to run, analyze, store, and host the numerical simulations for any one survey are the prime reason to put effort into coordinating between surveys as much as possible.

3.3. Astrophysics

The combination of LSST and *Euclid* will transform all areas of astrophysics beyond the cosmological measurements discussed in previous sections. We do not seek to provide a comprehensive review of these synergies in general astrophysical studies but provide a few interesting examples to illustrate the range of impact envisaged.

3.3.1. Solar System Science

In the context of planetary science, the strength of LSST is the high-cadence multi-epoch astrometry and photometry catalog that will contain hundreds of detections for each of millions of solar system objects (SSOs). This will dramatically improve our knowledge of SSO populations through discoveries and orbit determination (LSST Science Collaboration et al. 2009), phase curves for composition study (e.g., Oszkiewicz et al. 2012), and 3D shape modeling (Đurech et al. 2005, 2015).

Although the *Euclid* wide survey step-and-stare survey mode suggests that only one epoch will be acquired for each target, its observing sequence is casually well-adapted for solar system needs (Carr 2017); astrometric shifts within a single visit can indicate SSOs. The repeated visible and NIR photometry (over an hour) will provide the colors of SSOs, used for spectral classification and compositional interpretation (DeMeo & Carr 2013). In particular, the NIR colors are key to resolving the present degeneracy between several spectral classes (DeMeo et al. 2009) based on only visible wavelengths (as produced by the SDSS or *Gaia*; see Ivezić et al. 2002; Delbo et al. 2012).

Nevertheless, orbital determination, which is crucial to pinpoint compositions at specific locations in the solar system, will rely on data from other facilities as the hour-long observations by *Euclid* will not suffice to constrain SSO orbits. This point was a major limitation of the SDSS census of SSOs, which detected over 400,000 moving objects but only half were linked with sources with known orbits (Ivezić et al. 2001). However, a posteriori identification is possible when the known population increases (Solano et al. 2014), e.g., from LSST.

The synergy between LSST and *Euclid* is obvious even at the catalog level; spectral characterization from *Euclid* can be linked to orbits and 3D shapes from LSST. To do so, an efficient *Euclid*-centric ephemerides cone-search tool must be available (e.g., Berthier et al. 2016). This will enable multiple investigations including studies of dynamical families and surface aging effect (Spoto et al. 2015), the source region of near-Earth asteroids (Carr et al. 2016), highly inclined populations (Petit et al. 2017), and the large-scale distribution of material linked with solar system evolution (DeMeo & Carr 2014).

There are also synergies on the operation and cadence of LSST and *Euclid*. By performing observations simultaneously, the distance to SSOs can be determined by triangulation, providing

accurate orbits from only a few observations (Eggl 2011). Trans-Neptunian Objects at 45 au will have a parallax of a few arcseconds, tying down the distances to a high certainty. This increased parallax will also significantly reduce the orbital calculation uncertainties of the hundreds of new near-Earth objects that LSST will detect per night (if coordinated observations are possible). In terms of the different passbands, doing multiwavelength imaging with LSST and *Euclid* at the same time will give very important physical information regarding transient events such as asteroid collisions and comet outbursts. Owing to the observing geometry of *Euclid*, which points close to a solar elongation of 90° , and tight planned schedule, practical implementation of coordinated observations will require LSST to observe *Euclid* fields right after dusk and before dawn.

3.3.2. Internal Analyses of Galaxies

The internal properties of galaxies at $0.2 \lesssim z \lesssim 1$ are now being studied in detail, but the number of galaxies for which detailed data are available is small. By combining LSST and *Euclid*, we will revolutionize such internal studies. For decades, we have explored galaxies as if they were single “points” and even carried out analyses of galaxy evolution in terms of gross photometric quantities such as stellar and luminosity functions, and the correlation of light and mass with other properties such as internal kinematics. A few studies, mostly from *HST* (e.g., Abraham et al. 1999; Lanyon-Foster et al. 2007, 2012; Welikala et al. 2008), showed it was possible to determine the internal stellar population content of galaxies through a pixel-level approach, i.e., treating each pixel as a distinct object and determining its stellar population parameters, such as its stellar mass, light-weighted age, and metallicity. In this way, one can utilize all potential information about galaxies which can reveal the formation history of a single galaxy based on its spatial location. This technique also allows the creation of stellar mass images of galaxies.

The combination of LSST and *Euclid* data at the pixel level will create internal maps of star formation histories, metallicities, dust content, etc. This will be achieved through SED fitting to the matched pixels across a range of wavelengths. This approach has been applied to only a few hundred galaxies, but such an approach should be applicable to 10 million galaxies up to $z \sim 1$. This approach requires the imaging data from the two surveys to be aligned to high accuracy, and likely processed together, to avoid any offsets between the two and to ensure that the same physical regions are probed by both observatories. This will require matching of PSFs and pixel scales, as well as a careful control of the signal-to-noise and photometric limits, on a pixel basis. This approach can also be applied to fitting parametric models within galaxies, e.g., bulges and disk components.

3.3.3. Morphological Classification of Galaxies and Machine Learning

Galaxies display a wide range of morphologies, which encode information about their potential energy and angular momentum, as well as their cold gas content, the influence of cosmic environment on galactic structure, and the history of mass accretion and mergers from which the galaxies were built. As morphology is considered a fundamental property of the galaxy population and provides a window into the origin and fate of individual galaxies, surveys of galaxies like LSST and

Euclid are significantly enhanced through the morphological classification of each object they discover.

LSST and *Euclid* will provide images for billions of galaxies. Although each galaxy will have many parametric measurements (luminosity, shape, size), the sheer size of this database will be prohibitive for manually (visually) classifying these galaxies using the traditional Hubble sequence. Moreover, the Hubble sequence may not be relevant at high redshift, where many galaxies are irregular.

Fortunately, new machine learning methods can rapidly characterize billions of images through so-called “Deep Learning” models. After training, based on a subset of manually classified galaxies (possibly using GalaxyZoo; Dieleman et al. 2015; Simmons et al. 2017), we can use the deep learning models to efficiently process enormous data sets with approximately the same fidelity as manual classification. The higher-resolution *Euclid* data, especially in the deep fields, will serve as an ideal training set for both *Euclid* and LSST.

Such deep learning models are already under development for surveys like LSST and *Euclid*. Figure 2 shows the results of a deep learning model (R. Hausen & B. Robertson 2017, in preparation) trained on the Kartaltepe et al. (2015) visual classifications of the CANDELS *HST* survey (Grogin et al. 2011; Koekemoer et al. 2011). Shown are six objects with inset comparisons of the Deep Learning morphological classification probabilities (left corners) and those of the human-classified morphologies (right corners). As is evident from this figure, Deep Learning classifications can reproduce accurately the visual classification schemes used by astronomers on an object-by-object basis, but can be performed extremely rapidly and objectively. Since Deep Learning is sensitive to fine details in each image and uses this information to improve the galaxy classification, a uniform processing of the data from both LSST and *Euclid* will reduce systematic uncertainties in machine learning galaxy classifications that could otherwise arise from disparate treatments in noise, deblending, mosaicking, and artifacts. The combination of deeper LSST photometry and higher-resolution *Euclid* imaging could provide a new way of classifying galaxies, e.g., *Euclid* focusing on the inner, brighter structures at high resolution and LSST detecting the outer, fainter parts including merger debris.

3.3.4. Cluster Astrophysics and Dark Matter

In addition to their use as a cosmological probe (Section 2.4), galaxy clusters can be used to investigate the interplay of all major components of matter (dark matter, hot gas, and stars). Weak-lensing measurements with a high background source density can be used to map the total mass distribution in clusters, which has provided strong evidence for the existence of dark matter (Clowe et al. 2006) and constraints on the dark matter self-interaction (Randall et al. 2008; Harvey et al. 2015). Similar investigations can be performed with the joint *Euclid* and LSST weak-lensing analysis and would benefit from the increased source density.

As explained in Section 2.4, the combination of *Euclid* and LSST data enhances cluster weak-lensing studies at redshifts $z \gtrsim 1$, particularly at high ecliptic latitudes where there is reduced zodiacal background. At faint magnitudes, the fraction of compact sources increases, and while they are typically still resolved with *Euclid*, they can be too small for reliable shape measurements from the ground. Hence, it is only through the combination of *Euclid*’s resolution and LSST’s depth that the additional statistical power of these faint and compact galaxies

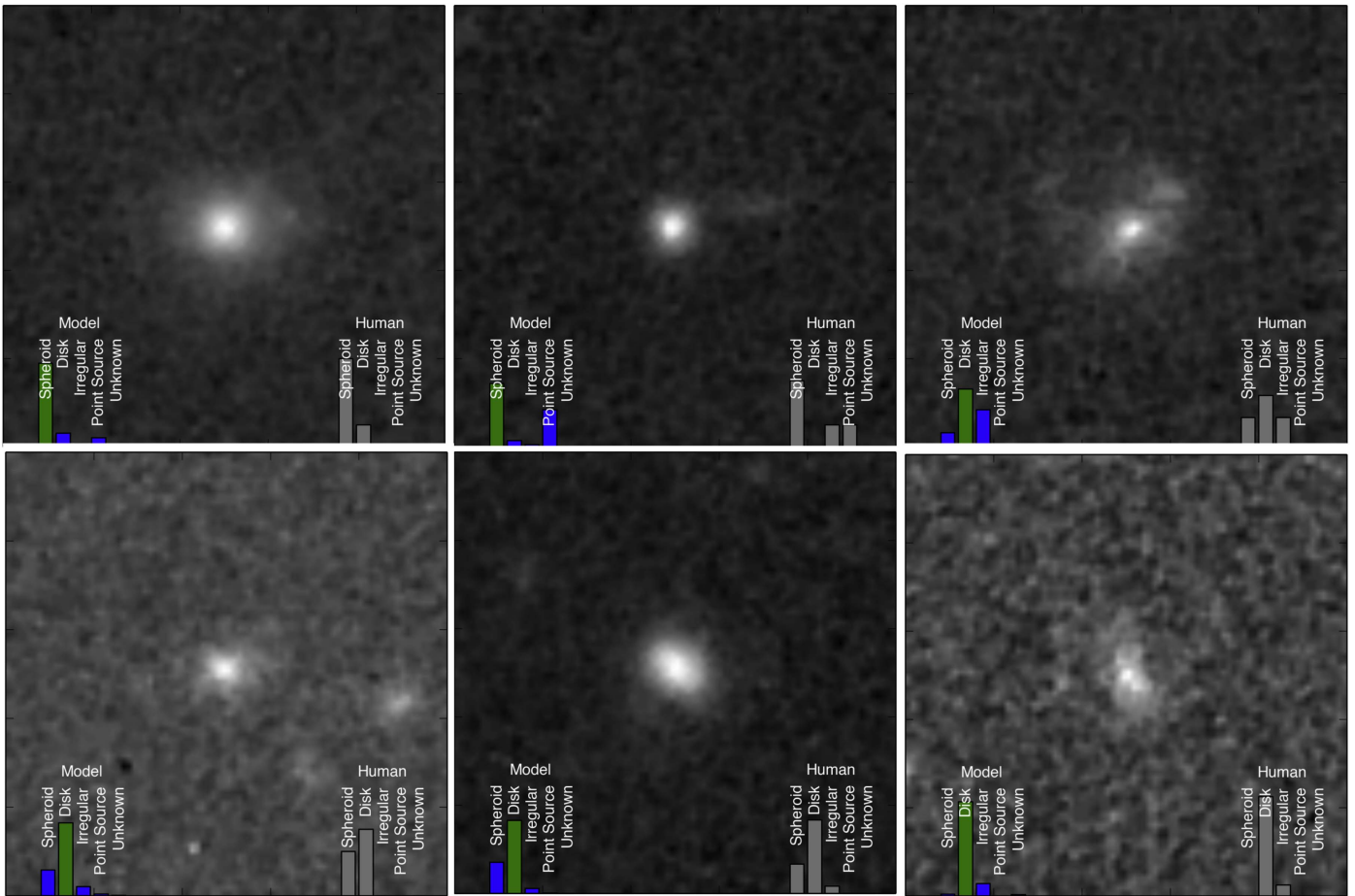


Figure 2. Representative results of a deep learning model for galaxy morphological classification (R. Hausen & B. Robertson 2017, in preparation) that will be applied to LSST and *Euclid* imaging data. Shown are *V*-band *HST* ACS images of galaxies in the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011), with visual classifications provided by Kartaltepe et al. (2015). The distribution of classifications assigned by human inspectors are shown as bar graphs in the lower-right corner of each image, reflecting the Spheroid, Disk, Irregular, Point Source, or Unknown classifications. The results of the deep learning model are shown as a bar graph in the lower-left corner of each panel, indicating the probability distribution of morphologies returned by the model. Green bars indicate when the most likely classification determined by the model matches the most frequent human classification.

can be incorporated into weak-lensing studies, amplifying the total scientific return. In addition, the gain from a joint analysis including deep LSST photometry can be increased further if advanced algorithms that provide reliable shape estimates from the *Euclid* imaging and also for galaxies with a slightly lower S/N than the default threshold $S/N > 10$ (*Euclid* will measure shapes at lower S/N) are implemented. Promisingly, Bernstein et al. (2016) already demonstrate accurate shape measurements for galaxies with $S/N > 5$ for simulated data.

The extra depth in the weak-lensing source catalog increases both the density and the average redshift of the source sample. This boosts the statistical constraining power not only for cosmological weak-lensing measurements, but also other weak-lensing studies, such as investigations of the mass properties of galaxies as a function of redshift (e.g., Leauthaud et al. 2012) or the mass calibration of galaxy clusters (see Section 2.4). The largest relative gain in the source density occurs in the high-redshift tail (see Figure 1). These extra sources increase the redshift lever arm over which “foreground” objects can be studied.

There are numerous motivations for studying galaxy clusters at $z > 1$, including their ability to tighten constraints on dark energy from clusters and the longer redshift baseline for exploring the physics of the intracluster medium and the evolution of galaxies. LSST data alone will be sufficient to

identify systems with strong red sequences out to $z = 1$; however, at $z > 1$, the NIR data from *Euclid* will be needed to detect and characterize such clusters. This is underlined by Chiang et al. (2014), who find that secure (proto)cluster detection will require photometric redshifts with at least $\Delta z/(1+z) \sim 2.5\%$ precision at $z > 1$.

4. Details of Coordination

The main data products needed for the weak-lensing cosmological analyses of LSST and *Euclid* are highly accurate photometric and shape measurements for as many galaxies as possible. A number of algorithms have been developed (and are still being developed) to perform those measurements, some of which work on individual exposures, others on co-added images, “co-adds,” that are aggregated from all suitable exposures for a given area of the sky. The main trade-offs between those approaches are computational efficiency versus flexibility and accuracy.³⁵ For maximum accuracy, one may want to avoid the co-addition process as it can introduce artifacts from objects that

³⁵ In principle, it is possible to create an optimal, statistically sufficient “likelihood co-add” (Kaiser 2001; Zackay & Ofek 2015), but the adaptation of this concept to survey data processing as well as downstream algorithms has not yet been successfully demonstrated.

are noticeably moving on the timescale of one to a few years; naive processing can also cause difficulties in interpreting the PSF across the boundaries where the number of co-added exposures varied or where some of the images have masked pixels. However, in practice, the data volume of LSST will probably demand that at least some data processing steps be performed on co-adds, e.g., object detection and initial characterization (Section 2.1).

4.1. Photometric Redshift Estimation

The quality of photometric redshifts depends critically on accurate, high-S/N photometry in multiple bands. In addition, galaxy photometry needs to be measured consistently across bands, i.e., a galaxy’s stellar population needs to contribute to the integrated flux with the same weight in each filter. Even with a fixed measurement scheme, this color consistency is not guaranteed when the observing conditions (most importantly the PSF) vary between bands.

One way to account for that is by generating PSF-homogenized co-adds, where each contributing exposure has individually been convolved with an extra kernel to match a desired homogeneous output PSF. If done for all exposures and all bands, integration apertures can be chosen in such a way as to optimize the S/N or to avoid contamination of adjacent objects. The alternative approach is to determine apertures or fit models that account for the respective PSFs in each band, or even in each exposure, as well as the presence of any adjacent objects (i.e., that handle deblending).

The following questions have to be answered for LSST and *Euclid* independently. (i) Which bands are used for source detection? (ii) Can the PSF be characterized sufficiently well on co-adds or is that only possible on individual exposures? (iii) Are fluxes measured on co-adds or on individual exposures? (iv) Is the PSF variation accounted for by homogenizing the co-adds or by adapting apertures and models to the PSF differences between the bands?

The decisions made for these questions will in some cases affect further coordination efforts, all of which will at a minimum require the adoption of a common astrometric system. Aperture fluxes are only consistent between both surveys if the PSFs have been matched to the same output PSF. Model fluxes are consistent only if they adopt the same model parameterization and parameters for each object *and* those models accurately describe the data. One must decide on a common aperture or model based on *Euclid* VIS detections and subsequent object characterizations must then be consistently applied to the LSST and *Euclid* NISP images. If, however, PSF-homogenized co-adds are adopted by LSST, a joint pixel-level analysis at the co-add level could negate any advantage of the higher spatial resolution of *Euclid*. As a result, aperture and model fluxes of blended galaxies would remain susceptible to undetected leakage between the blends.

Once accurate multiband photometry is available, the estimation of photo- z ’s and their calibration is relatively independent of the project. Photo- z codes having a long history of development have matured rapidly in recent years. The requirements on scatter and outlier rates for *Euclid* have been achieved in real-world data sets with existing codes (Hildebrandt et al. 2010; Laigle et al. 2016), although the demonstrations to date at *Euclid*’s depth have utilized substantially more bands of photometry than *Euclid*+LSST would provide. These codes also meet requirements for LSST on

simulations (Graham et al. 2017), albeit with the assumption of perfect knowledge of templates or very large representative training sets. Calibration requirements are also similar for both surveys (Section 3.1).

Apart from algorithms, a key requirement for photometric redshifts for both projects will be enormous training and calibration samples that may only be obtained with highly multiplexed optical and infrared spectrographs on large telescopes. Sharing and combining these calibration samples would certainly make sense but would likely not have a major impact on the data flow for either project.

Previous works (e.g., Dahlen et al. 2010; Newman et al. 2013) suggest that using high-resolution imaging data (such as those *Euclid* would provide in areas that overlap the LSST survey footprint) could reduce catastrophic redshift failures due to blending. This is supported by as yet unpublished *WFIRST* work (S. Hemmati & P. Capak 2017, private communication), which suggests having a prior on galaxy shapes as a function of magnitude and other parameters would significantly improve LSST photometry. The largest impact of the *Euclid* NIR photometry on LSST photometric redshifts will be for galaxies where $1.5 < z < 3.0$, for which the sharpest features in galaxy spectral energy distributions (SEDs), the rest frame $0.1216 \mu\text{m}$ ($\text{Ly}\alpha$) and the $0.4 \mu\text{m}$ (Balmer) break, are not constrained by the LSST filters. This lack of spectral coverage decreases the precision of photometric redshifts and increases the outlier fraction. We attempt here to provide some quantitative rigor to this claim based on simulated photometry, but leave a full analysis of the improvements of *Euclid* and LSST photometric redshifts enabled by co-processing the data to a future work.

First, we use the COSMOS data (Laigle et al. 2016) scaled to approximate the LSST+*Euclid* data. In Figure 3, we show a self-organizing map (Masters et al. 2015) of the LSST and LSST+*Euclid* data, which provides a 2D projection of the higher-dimensional color space. One can clearly see additional features in the LSST+*Euclid* color space compared with the LSST color space. This indicates a significant increase in the information content. Next, we simulate the photo- z performance using the methodology described in Stickley et al. (2016) but only simulate the end-result LSST and *Euclid* photometry with estimated Gaussian noise and do not perform end-to-end image simulations. In brief, the simulations start with the COSMOS2015 catalog and photometric redshifts (Laigle et al. 2016). We then fit a combination of Brown et al. (2014) galaxy and Salvato et al. (2011) AGN templates modified with additional emission line ratios and dust obscuration to the non-stellar photometry. This normalized template set is then used to estimate LSST and *Euclid* photometry, which is then degraded to the expected sensitivity levels in those surveys. In Figure 4, we show a simulation of the expected improvement in the LSST photo- z performance for the LSST “gold” sample $i_{AB} < 25.3$. In these simulations, the scatter decreases by a factor of ~ 2 in the $1.5 < z < 3$ redshift range (the so-called “redshift desert”) and $\sim 30\%$ at other redshifts. The improvement of photo- z ’s in the redshift desert provided by the combination of *Euclid* and LSST will also increase the precision from photometric redshift calibration techniques that rely on spectroscopic cross-correlations/clustering redshifts, as tighter redshift distributions yield both smaller calibration errors and a decreased sensitivity to bias evolution (Newman 2008). This will correspondingly reduce systematic uncertainties in the large-scale structure and

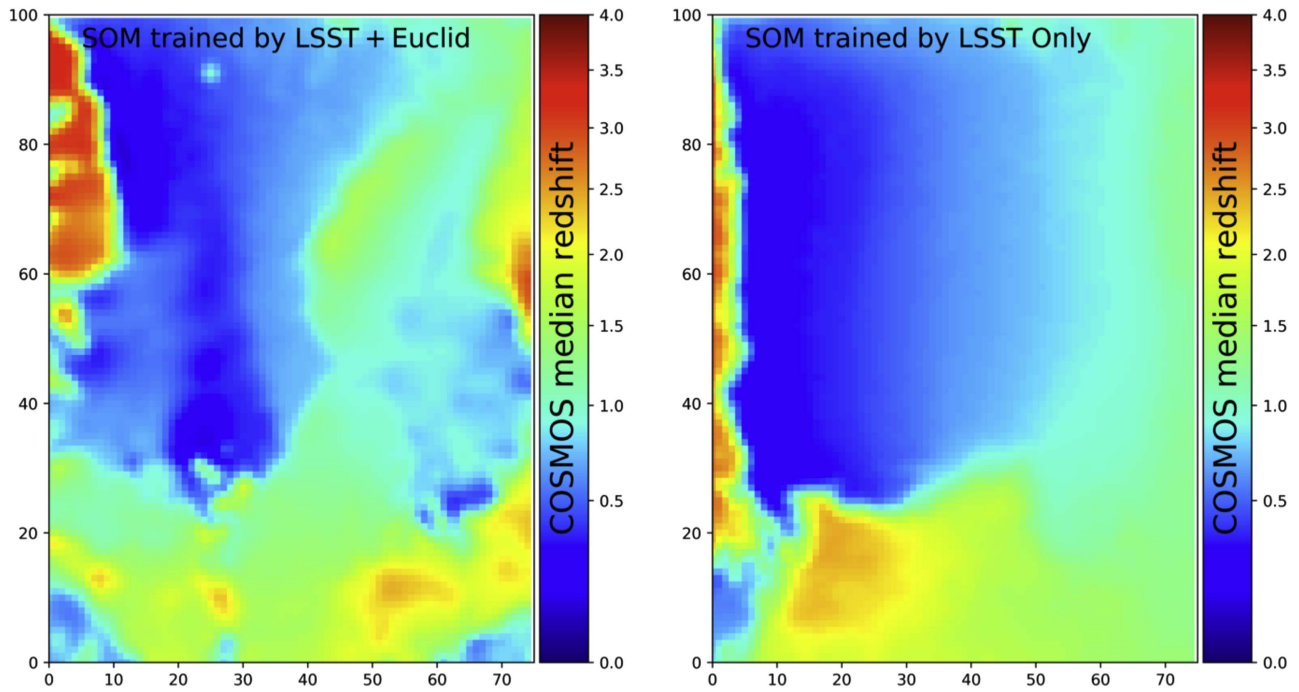


Figure 3. Self-Organizing Maps (SOMs) trained by colors of COSMOS galaxies (to the *Euclid* depth) in the LSST+*Euclid* filters (left) and the colors in the LSST filters only (right) color-coded by the median thirty-band redshift of the COSMOS galaxies mapped to each cell. SOMs are a class of unsupervised neural networks that reduce dimensions of data while preserving the topology. Each cell on these rectangular grids has a weight vector with the dimensions of the input data and therefore represents a point in the multidimensional color space (see Masters et al. 2015 for details). The greater complexity of features in the SOM shown on the left reflects the increase in the information content when *Euclid* bands are incorporated in the training. These features show that the photo- z resolution in a multidimensional color space has improved.

weak-lensing probes of cosmology from each experiment. However, the true performance is sensitive to the photo- z method, depth, data quality, selection function, and other properties of the LSST photometry and photo- z pipeline and so our projected improvement should be taken as a potential scaling rather than a definitive result (for instance, gains are smaller although still significant in simulations where magnitude priors are incorporated; S. Schmidt 2017, private communication).

Although this improved performance will not directly apply outside of the overlapping area, it will provide statistically robust priors for the full survey. For instance, the effects of blending on the photo- z distribution can be quantified as a function of LSST color and magnitude space and applied to the full survey. Similarly, the improved photo- z distributions in the overlapping area will provide redshift priors for the full LSST sample where the infrared data are not available.

4.2. Shear Measurement

As already discussed in Section 2.1, joint analysis of LSST and *Euclid* images could be important to improve deblending issues in LSST, which has been identified as an important factor for LSST weak-lensing measurements, given the depth of the images. More generally, shear estimation hinges on highly robust and well-understood algorithms to estimate the PSF and to remove the impact of image defects and of the PSF on the galaxy shapes to enable a robust ensemble shear estimate with systematics removed or properly marginalized over. Here, the complementarity of LSST and *Euclid* is primarily in the fact that the issues in PSF estimation in the two surveys are quite different (and independent), image defects do not occur at the same levels, and they experience different levels of chromatic effects (which

cause the effective PSF to differ for each galaxy or even for different points in the same galaxy if it has color gradients). For this reason, many of their shape systematic errors can be treated as independent. *Euclid* and LSST will also measure shapes using different spatial weights, which, in addition to being a possible cross-check on systematic errors, can be relevant in handling intrinsic alignment effects or in using intrinsic alignments as a probe of cosmology (Chisari et al. 2016). For example, the combination of two shape measurements from the same sample of aligned galaxies can aid the mitigation of cosmic variance (similarly to Seljak 2009) when constraining anisotropic models of primordial non-Gaussianity.

Because of this complementarity, one useful high-level coordinated analysis would be the cross-correlation of shear fields in some region of the sky (see Jain et al. 2006 for an early example of this type of analysis in the context of separate exposures within a survey). This does not require coordination at the image or catalog levels. Nearly all sources of coherent additive biases should cancel out with such a cross-correlation, meaning that this joint analysis should provide a cleaner shear power spectrum requiring less marginalization over nuisance systematics than auto-correlation.

To test for multiplicative offsets in the shear calibration of the two surveys would require the sharing of catalog-level products (e.g., per-object weights, photometric redshifts, and galaxy shapes). For this test, we would select a sample of massive foreground objects, such as photometrically detected galaxy clusters, and compare the galaxy–galaxy or cluster–galaxy lensing amplitudes. If each survey calculates the shear or surface mass density using internal data products, a comparison of the results will not cleanly tell whether there are disagreements in shear calibration or photometric redshifts.

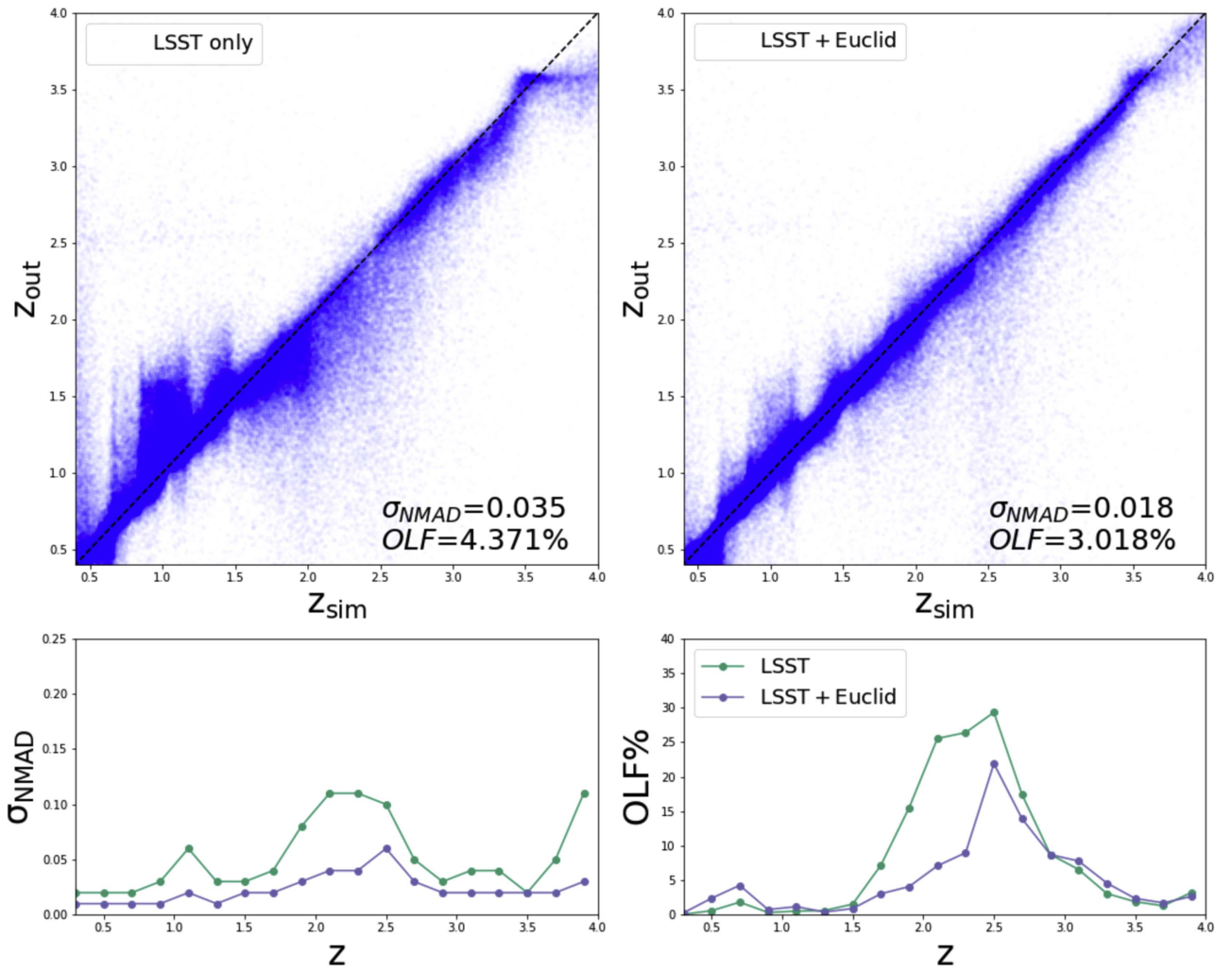


Figure 4. Top: comparison of simulated photo- z vs. spec- z performance for LSST and LSST+*Euclid* using the methodology described in Stickley et al. (2016). Clear improvement in the performance can be seen. Bottom: σ_{NMAD} defined as $1.48 \times \text{median}(|\Delta z|/(1 + z_{\text{spec}}))$ and outlier fraction (defined as the fraction of objects with $|\Delta z|/(1 + z_{\text{spec}}) < 0.15$) in redshift bins of 0.2 are shown for the simulation. In these simulations, both the dispersion and outlier fraction improve by a factor of ~ 2 between $1.5 < z < 3$.

However, if one survey uses a matched sample of objects to infer the surface mass density around these lenses, and the other uses the cross-matched catalog to adopt the weighting scheme and photometric redshifts from the other survey, while using internal galaxy shape estimates, then a comparison of the inferred survey mass density should agree in the absence of relative shear calibration biases. Amon et al. (2017) offer a recent example of this type of comparison using KiDS i - and r -band data with very different depths. All such comparisons should be done at the level of inferred shears (or inferred surface mass densities), not per-object galaxy shapes. Per-galaxy shapes measured using different algorithms should not necessarily agree depending on differences in weighting schemes and resolution of the imaging data, so comparison must be done using the quantity that is really of interest—the ensemble shear estimate.

4.3. Weak-lensing S/N and Photo- z Accuracy: An Example

We provide here an example calculation (with some simplifying assumptions), demonstrating that the weak-lensing S/N is increased due to the improved photo- z accuracy in the overlap area between *Euclid* and LSST. Future efforts will do more complete calculations

and full joint dark energy forecasts for the *Euclid* and LSST combination. We include this calculation as a preliminary demonstration of the power of combining these two surveys.

To this end, we consider the S/N for measuring the shear power spectrum from the *Euclid* and LSST data and the cross-correlation spectra between sources in the common area. For this analysis, we first assume that one is not sharing data between the two surveys, but only use the final catalogs so that photo- z accuracy is the standard for each survey taken on its own. The data vector is then

$$\mathbf{D}^T(\ell) = \{C_{ij}^{EE}(\ell), C_{ij}^{EL}(\ell), C_{ij}^{LL}(\ell)\}, \quad (1)$$

where $C_{ij}^{XY}(\ell)$ is the shear power spectra for the redshift bin combination (i, j) at the multipole ℓ for sources in the the X and Y catalogs (with $X = E, L$ for *Euclid* and LSST). The covariance matrix $\Gamma_{ij}(\ell)$ can be computed as detailed in Harrison et al. (2016), to which we refer the reader for further details. For a given ℓ , we the define the S/N as $\mathcal{S}_+ = \{\tilde{\mathbf{D}}^T \Gamma^{-1} \tilde{\mathbf{D}}\}^{1/2}$, where quantities with a tilde are computed for a given cosmological model and using the photo- z specifications for each single survey.

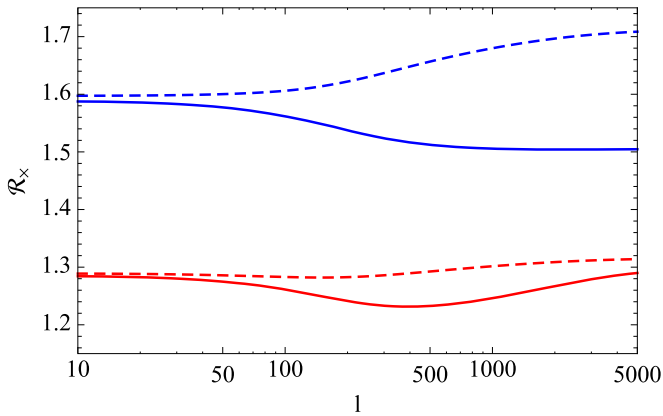


Figure 5. S/N ratio of the measurement of the lensing data vector \mathbf{D} defined in the text (including contributions from both the auto- and cross-correlation power spectra) with and without photo- z improvement from survey combination. Blue (red) lines refer to the case with 11,000 (7000) deg^2 of overlap between the two surveys, while dashed and solid lines are for $f_E = 0.75$ and 1.0, respectively. We use 100 logarithmically equispaced bins in ℓ over the range $10 \leq \ell \leq 5000$, and 10 approximately equipopulated redshift bins over the range $0 \leq z \leq 2.5$. We set $i = j = 6$ so that the results refer to the bin combination corresponding to the median of the galaxy samples from each survey. Changing the values of (i, j) gives qualitatively similar results.

We now recompute the S/N under the assumption that the two surveys share photometric data so that the photo- z accuracy is better for the sources in common. The *Euclid* and LSST shear catalogs will then be split in two parts depending on whether sources are in the overlap area. Both the theoretically expected values and the covariance matrix must be recomputed accordingly. We can then define a new S/N ratio as $\mathcal{S}_x = \{\hat{\mathbf{D}}^T \hat{\mathbf{\Gamma}}^{-1} \hat{\mathbf{D}}_{ij}\}^{1/2}$, where now the hat quantities are computed with improved photo- z accuracy parameters assuming a fraction f_E of *Euclid* sources are also detected and useful for lensing³⁶ by LSST.

Figure 5 plots the ratio $\mathcal{R}_x = \mathcal{S}_x / \mathcal{S}_+$ as a function of ℓ for the autopower spectra of the most populated redshift bins. We consider two different cases for the overlap area (namely, 7000 and 11,000 deg^2) and for the fraction of *Euclid* galaxies in common with LSST (i.e., $f_E = 0.75$ and $f_E = 1$). As is apparent, the boost in the S/N from sharing data between the two surveys can be quite large even under the most conservative assumptions. Although we are well aware that a boost of a factor b in the S/N does not automatically translate into a similar improvement in the constraints on the cosmological parameters, this preliminary result (obtained by ignoring the systematics and the intrinsic alignment contribution) clearly suggests that the improved accuracy in photo- z 's coming from the joint use of *Euclid* and LSST photometric data offers the fascinating possibility to significantly increase the amount of information from the two surveys.

4.4. Survey Characteristics

4.4.1. Sky Coverage

The main LSST survey will cover a total area of 18,000 deg^2 located between an equatorial decl. of $+2^\circ$ and -62° to ensure all data contributing to its weak-lensing science are acquired under an airmass of 1.4. This defines a large sky area denoted

³⁶ We do not expect $f_E = 1$ since LSST has a lower expected n_{eff} (number density of galaxies useful for weak lensing) than *Euclid*.

by the green box on Figure 6. There will be secondary LSST surveys such as the South Celestial Pole survey that will bring a shallower exploration of the southernmost part of the sky (decl. -62° to -90° ; red box in Figure 6) as well as the ecliptic plane at the northern decl., up to and above $+30^\circ$ (Figure 7). LSST will use the same observing strategy for all of its surveys with a pair of short fixed integrations (two 15 s exposures) per single visit of each sky area in all of its six available photometric broad bands (u, g, r, i, z, y).

There are three key constraints to the visible area for *Euclid*. First, stellar contamination, coupled to stray-light constraints from the integrated light from the galactic plane, stops *Euclid* from observing at low galactic latitude. Second, with its constant integration time, *Euclid* has to avoid areas with significant attenuation from galactic reddening. Finally, the zodiacal light from L2 causes a significant background, which excludes the ecliptic plane (the S/N would be too low for the fixed integration time, especially for the NISP). Therefore, to ensure *Euclid* can achieve its requirement of at least 15,000 deg^2 of extragalactic sky over the entire sky, the survey area is defined simply as galactic latitude $> \pm 25^\circ$, ecliptic latitude $> \pm 15^\circ$, and reddening $E(B - V) < 0.08$ (and up to $E(B - V) < 0.15$ to avoid holes or islands in, or near, the main footprint).

These different areas are shown in Figure 6 on an equatorial coordinate projection, with the *Euclid* Wide Survey shown in yellow, alongside the DES footprint and other various points of interests. This simple definition presents the clear advantage of focusing on the best parts of the extragalactic sky. There are, however, areas at lower galactic latitudes near $l = 180^\circ$ that *Euclid* could explore since requirements on scattered light might still be met; this is currently being explored by *Euclid* using various weight maps that draw a longitude-dependent exclusion zone.

LSST could reach further north than its present survey region due to its location (-30° latitude). Therefore, if we wished to maximize the LSST–*Euclid* overlap, nearly 11,000 deg^2 of the *Euclid* Wide Survey could be covered by LSST (considering a longitude-dependent exclusion zone, not the depiction shown in Figure 6), or nearly three-quarters of the required 15,000 deg^2 . This would require a new, currently unplanned, dedicated LSST extension survey of 3000 deg^2 from a decl. of $+2^\circ$ to $+30^\circ$. Such an extension would however only need to reach the modest depths required by *Euclid* shown in Table 1, and would exclude the u -band, which would otherwise require excessive integration times at such high airmass.

In Figure 7, we illustrate this extension with a simulation using the Operations Simulator (Delgado et al. 2014) where we extend the LSST footprint survey with decl. $< +30^\circ$, galactic latitude $> \pm 25^\circ$, and ecliptic latitudes $> \pm 15^\circ$. For this case, we allocate 43 visits spread across the filters, excluding the u -band. Figure 7 shows the effective r -band depth achieved by this new survey. This extension only requires ~ 222 hr (or 27 full nights) to cover the additional 3000 deg^2 in the g, r, i, z bands.

4.4.2. Depth Requirements

LSST will observe its survey areas in all of its six available photometric filters (u, g, r, i, z, y) using a fixed integration approach (two 15 s exposures per single visit of each sky area), leading to different depths across the various bands. There will be approximately 825 visits per field area over the 10 year long survey, leading to a net gain of more than 3 magnitudes over

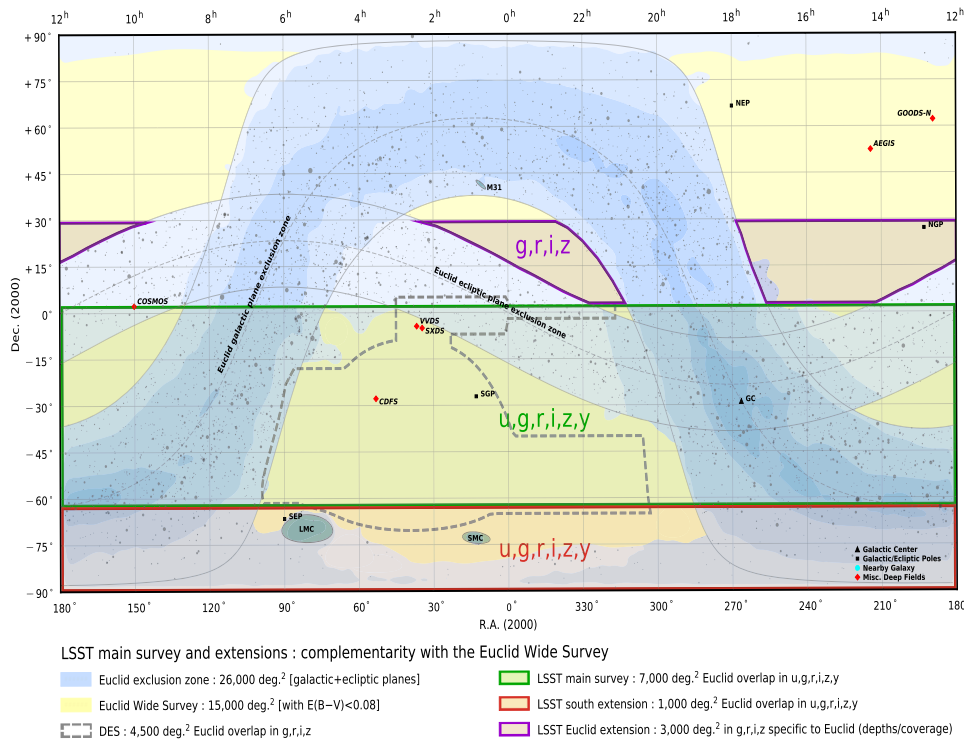


Figure 6. LSST survey areas and photometric bands and the *Euclid* Wide Survey with its exclusion zone (blue: galactic plane + ecliptic plane + reddening). We indicate in the legend the number of square degrees from the LSST surveys that overlap the *Euclid* Wide Survey in the relevant photometric bands.

the single visit depths listed in Table 1, assuming a point-source 5σ PSF photometry metric/performance (Ivezic et al. 2008).

The *Euclid* Wide Survey average depth requirements remain anchored in the ESA mission definition study report (Laureijs et al. 2011). The critical need for ground-based data is driven by the required photometric redshift accuracy. The spacecraft brings three infrared bands (Y, J, H) and one broad optical band (VIS) of limited use for the photometric redshifts. A full optical broadband data set such as the LSST g, r, i, z is critical to reach the required photometric redshifts accuracy across the redshift range explored by *Euclid*. The addition of the LSST u -band further improves the quality of the photometric redshifts, especially at low redshift. The required depths are reported in the same metric as LSST in Table 1.

We can now answer an important question: how many visits are needed by LSST to fulfill *Euclid* depth requirements? For all five bands, from u to z , 56 LSST visits are needed, for a total of approximately 28 minutes of integration time, or 37 minutes when considering the overheads (readouts plus slewing). However, *Euclid* will suffer from a zodiacal diffuse light background that varies from the ecliptic poles to the ecliptic plane by more than 1.2 mag. Since *Euclid* will use a fixed integration time, these space data will not be uniform in depth across the survey footprint and, as a consequence, ground-based data should as much as possible match the space data depth or anchor on its deepest part, which is possible with LSST.

The numbers listed above present the average solution across the entire 15,000 deg.² *Euclid* survey footprint (not all of which could be observed with LSST). Assuming uniform survey progress for the LSST wide-fast-deep survey, it would take about 10 months to complete the $g, r, i,$ and z observations, and

four years for the u -band observations. Therefore, for all bands except u , the required depths in the areas of overlap between *Euclid* and the LSST main survey will be achieved in about the first year of the LSST survey. For the south celestial pole, with the current allocation of time, LSST will not reach the required depths in eight years for the u -band and will require four to five years for the i - and z -bands, and within two years for the g - and r -passbands.

The *Euclid* Wide Survey will be deeper by ~ 0.3 mag near the ecliptic poles, compared to the average survey depth, due to the dependence on the zodiacal background. To exploit this extra depth, approximately 1.7 times more LSST visits would be needed at high ecliptic latitudes compared to the average requirement. Even deeper LSST photometry would be beneficial for future joint analyses that could explore the inclusion of galaxies with lower S/Ns, e.g., Section 4.4.1 and Figure 1.

5. Modes of Coordination

In this section, we discuss various modes in which the experiments can coordinate with each other. These can be broadly separated into several different types of cooperation, based on the types of dependency that can occur as discussed above, e.g., methodological, calibration, and data exchange. In addition, within each of these categories, gains can be made that lead to improved/new science, improved reproducibility, and improved efficiency.

5.1. Methodological

There are several ways in which methodological coordination can lead to improved reproducibility and new science. In developing methodologies for cosmological probes and

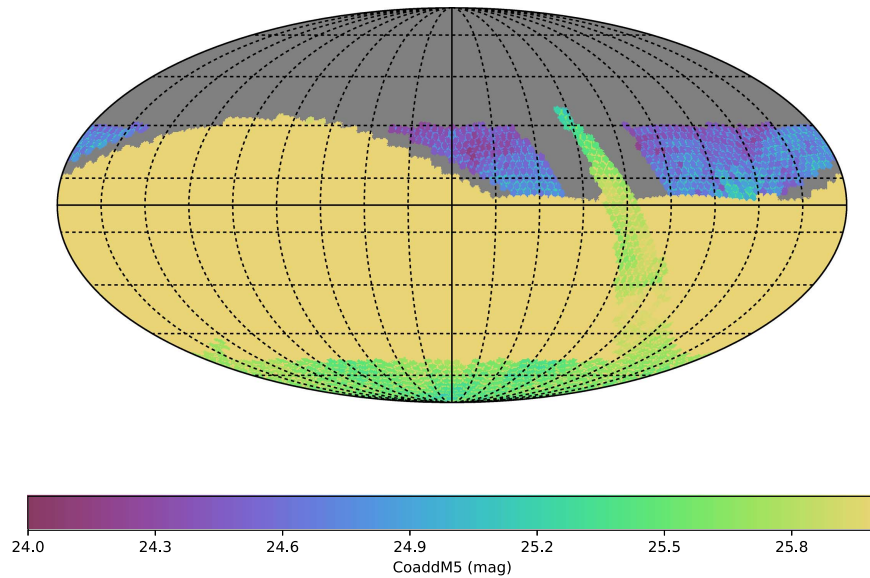


Figure 7. Simulation of an extended LSST survey at high latitude (decl. above $+2^\circ$ and up to $+30^\circ$) that increases the overlap between *Euclid* and LSST by approximately 3000 square degrees (using the target LSST depths described in the text). The decrease in image quality and the increase in atmospheric extinction for these high airmass observations reduces the effective exposure times at the high decl. limit of this survey. A more detailed study will be required to determine how to minimize the impact of such effects on the LSST–*Euclid* science cases, e.g., photometric redshift estimators. The simulation also shows the planned LSST ecliptic plane survey at northern latitudes for reference.

algorithms to measure these observables, inter-experiment collaboration is already helping in many ways. As it is difficult to develop common algorithms to work on the raw data analysis, catalog-level manipulation to create summary statistics is more commonly done. There are many case studies from current surveys in which common inter-experiment algorithm development leads to more robust and validated code, for example, CAMB (Lewis et al. 2000), GalSim (Rowe et al. 2015), and SNANA (Kessler et al. 2010).

In addition to individual-experiment cosmological probes, the multi-experiment combination enables additional statistics to be applied to the joint data set. Synergies between experiments will help, for example, by probing different parts of the cosmic expansion history so that in combination, Stage IV experiments will probe a larger redshift range than any experiment individually, leading to improved dark energy constraints, e.g., combinations of SN (LSST) and BAO (*Euclid*) distances (see Aubourg et al. 2015). By taking the ratio of observables from different experiments, common statistically limiting effects (e.g., cosmic variance) may be mitigated, a method known as the multitracer approach (Seljak 2009; Abramo et al. 2016). Multimessenger (i.e., multiple wavelength) observations of the same area of sky also allow for additional cosmological probes, e.g., through cluster studies and kinetic-SZ measurements. Such multimessenger observations also allow for consistency tests to check for replicability and reproducibility of results, e.g., measuring the weak-lensing power spectrum from two entirely distinct data sets to be consistent (e.g., Schaan et al. 2017). *Euclid* Deep Fields may also provide extra epochs for the LSST cadenced observations important for transient science as well as proving high-resolution imaging (and grism spectroscopy) of the host galaxies; such data will improve SNe Ia as “standardizable candles.”

5.2. Calibration

As previously noted, the precise calibration (e.g., photometric, astrometric, and shape) of both surveys will require considerable external information (see Section 2). This includes astrometric information to establish a coordinate grid on the sky, galaxy spectra to calibrate photometric redshifts (Section 2.2), high-resolution imaging to calibrate measurements of weak gravitational lensing (Section 2.3), and simulations of what the universe would look like in different cosmological models. Both surveys will also require significant effort to achieve their desired photometric accuracy, and coordination on cross-calibrations between the surveys, and with other large-area experiments (e.g., *Gaia*), will be of huge potential benefit to each survey and the whole astronomical community. All of these require an expensive investment of time and resources (other telescope observations and supercomputers), plus effort to analyze and reduce the data into usable form. A well-organized effort aimed at the sharing and comparison of calibration data will be needed to reach the full potential of cross-survey calibration.

A key component of both surveys is the use of gravitational weak lensing to measure cosmological parameters. This is at the heart of both *Euclid* and LSST but remains a significant observational challenge given the potential systematic uncertainties and its low S/N per galaxy. Present cosmological weak-lensing measurements have made huge strides in demonstrating the significant promise of this technique but have also found intriguing tensions with other cosmological measurements, which could be new physics or unaccounted systematic errors (e.g., Abbott et al. 2016; DES Collaboration et al. 2017; Efsthathiou & Lemos 2017; Harnois-Déraps et al. 2017; Hildebrandt et al. 2017; Joudaki et al. 2017a, 2017b). The comparison of LSST and *Euclid* weak-lensing measurements seems to be an excellent risk mitigation given the accuracy demanded by both experiments. Direct galaxy-by-

Table 1
LSST and *Euclid* Survey Depths

	<i>u</i>	<i>g</i>	<i>r</i>	<i>i</i>	<i>z</i>	Note
LSST	23.7	24.9	24.4	24.0	23.5	Depth reached per single visit
<i>Euclid</i>	25.4	25.6	25.3	25.3	24.9	Average requirement for ground-based complement
LSST visits	23	4	5	11	13	Number of visits to reach the <i>Euclid</i> depths
LSST exp. (mn)	11.5	2	2.5	5.5	6.5	Total integration needed (open shutter time)
LSST time (mn)	15	2.6	3.3	7.2	8.5	Total budget (integration plus overheads)

Note. We quote the total magnitude for a point source at 5σ with a PSF fitting photometry.

galaxy comparisons of shapes are likely to be less useful than overall cross-calibration of shear; the measured “shape” of a galaxy is highly dependent on the method used to measure that shape, and the methods for the space-based *Euclid* data and the ground-based LSST data will differ in significant ways.

5.3. Data Sharing

A higher level of coordination can occur when the collaborations agree to exchange or share the actual data that they collect and/or produce. In experiments such as *Euclid* and LSST, where data processing is a rather long chain of actions, this can take many forms that lead to various degrees of scientific benefit and mutual dependencies. Section 2 covered a number of science cases for data sharing between the collaborations. These are explored in more practical terms here, starting from the deepest level of data exchange. Constraints coming from the respective observation schedules of the experiments are not discussed here; they would give rise to a higher level of practical complexity.

The deepest level of data sharing would occur with pixel-based coordination. At this level, collaborations agree to share access to the lowest stage of their data products. It is likely that, practically speaking, calibration of raw data is still within the remit of each experiment as it requires detailed expertise of instrument scientists which is not easily shared. Nevertheless, coordination must take place between these instrument scientists, if only to define common calibration references. Astrometric reference coordination in the post-*Gaia* era is probably going to be straightforward, but photometric calibration is more demanding, especially since both experiments have extreme requirements on the knowledge of galaxy colors for photometric redshift determination. Thus, to achieve coordination at the pixel level, collaborations should identify a forum to exchange information on standards, instrumental characteristics, and calibration approaches. In such a framework, a very large number of the scientific goals identified in Section 2 can be reached. A single catalog can be built from detection applied to the complete data set. This catalog then benefits from both the higher spatial resolution of *Euclid* to achieve high-efficiency deblending on the brighter sources and the greater depth of LSST to provide reliability to the fainter sources. In such a configuration, we probably reach the optimal case for photometric redshift determination as well. If coordination is put in place at the pixel level, then it also allows all cases of shear measurement calibration/correlation between the two experiments, as the source lists and pixel information are fully compatible.

A slightly less deep level of coordination could be achieved with catalog-only combination. This case is more restricted than the above as here data coordination would occur when both collaborations have created their source catalogs. A driver

for this case is to maintain a separation for each experiment as catalogs are built from two independent data sets. The main purpose of coordinating at the catalog level is to increase (or quantify) the reliability of sources in each catalog. On the LSST side, the high-resolution *Euclid* source list can provide prior information for deblending, while on the *Euclid* side a deeper source list will provide better control of source reliability at low S/N. Practically, a coordination at catalog level would likely mean two stages of object detection and cataloging, one independent and one using the other catalog for prior information. However, to allow one to combine the catalogs, a number of conditions have to be satisfied; these conditions are similar to those set above. A shared astrometric reference must still be defined so that cross-matching at the catalog level can be performed. This must be defined early on as astrometric corrections are applied before cataloging takes place. Coordination at the methodological level should also occur to clarify the particular approaches used for object detection and thus significance. Beyond increasing the reliability of sources inside the catalogs, the benefit of coordinating at the catalog level can be to allow the creation of common source lists between the two data sets, for instance, to build samples on which shear measurement calibration can be inspected/compared. A limit of coordination at the catalog level is that it is not sufficient to guarantee the compatibility of photometric measurements, unless explicit coordination has also happened for the definition of a photometric reference system and for the measurement approach itself.

Coordination in photometric measurements is also a possibility. As Section 2 mentions, both collaborations have an interest in accessing each others’ photometric capacities: *Euclid* to cover its wide visible band with a set of narrower filters and LSST to increase its wavelength coverage with the *Euclid* NIR bands. In that context, one can envision a coordination at the photometric measurement level, each collaboration providing, as a service, forced photometry on its imaging data. This is probably the minimum level of coordination (above no coordination at all). It is probably also the level where data independence is maximum. To achieve this, however, once again a number of steps have to be taken so that the imaging data sets are compatible with one another. Quite obviously, these will include defining common astrometric and photometric references, but also sharing information on data processing steps up to the construction of calibrated images (individual and co-added) so that uncertainties in the photometry are properly taken into account.

The above cases thus all show that if coordination between *Euclid* and LSST should cover mission data, a number of steps have to be taken whatever the case. For instance, common calibration references should be defined (at least for the astrometric reference) and information should be exchanged on the early stages of data processing, including catalog

construction. As both collaborations have already clearly defined processes for data processing, implementing these coordination requirements should be quite feasible.

5.4. Survey Planning

There are a number of opportunities for coordination within the planning of the LSST and *Euclid* surveys. These include common survey footprints, depth as a function of passband, time of observations (including contemporaneous observations), software for processing and calibrating the pixel data, common calibration catalogs, and interfaces for sharing catalogs and resources. Many of these coordination activities can be accomplished prior to the launch of *Euclid* or the engineering first light for LSST.

As described in Section 4.4, for the main LSST survey area, comprising 7000 deg² of overlap, the LSST will reach the equivalent depth of *Euclid* within the first ~ 10 months of observations for all except the *u*-band observations. Prioritizing a uniform sampling of the sky by the LSST (as opposed to a rolling cadence where less of the sky is observed but more frequently) and optimizing the scanning strategy for *Euclid* would maximize the speed at which this joint data set could be derived (although the impact of such a change on LSST transient science would need to be studied). For the 1000 deg² patch of overlap at the South Celestial Pole (SCP), a comparable depth would require between 2 and 8 years of overall LSST operations, dependent on the passband; the SCP is classed as a mini survey within LSST and gets fewer visits than the main survey. Increasing LSST sky overlap with the *Euclid* survey beyond these two areas would require extending the LSST footprint with a northern 3000 deg² (up to a decl. of +30°), costing approximately 1% of the LSST overall survey time.

Early *Euclid* science coordination with the commissioning activities of the LSST would provide substantial opportunity for the early calibration of the joint data sets. During commissioning, LSST expects to survey a small (~ 100 deg²) region to a ten-year depth and a wider (~ 1000 deg²) region to a two-year depth to validate its source detection and characterization algorithms. *Euclid* plans to make regular data releases to the public to promote legacy science as well as stimulate follow-up observations. *Euclid* will have four “quick” (Q) release fields (some possibly selected via community input) and three official data releases (DR) spaced regularly throughout the mission with the first quick release 14 months after the start of survey operations. The quick release products will have undergone Level 2 data processing (co-added images, PSF model and distortion maps, co-added spectra) with the best available calibration at that time, but will be restricted in area ($\simeq 50$ deg²) and location. The official data releases will grow in area over the survey lifetime (e.g., 2500, 7500, and 15,000 deg² for DR1, DR2, and DR3 respectively) and also provide Level 2 and Level 3 (e.g., cosmology-related data products) data products. An obvious area of coordination would be to ensure that LSST observations are available for the *Euclid* quick release fields at the same depth as (or greater than) the *Euclid* visible and NIR photometry.

Coordination of early *Euclid* and LSST observations would clearly promote tests of joint processing algorithms and calibration of the photometric redshifts and shear measurements. This could occur early on in both surveys, although given the constraints on the commissioning time for the LSST it is unlikely

that any extended-footprint surveys could be undertaken within that commissioning or early operations period.

The existence of detailed survey simulation frameworks for both experiments provides a mechanism by which common commissioning geometries and the methodologies for synchronization of the survey strategies (including any advantages from contemporaneous observations) could be evaluated soon. Making such tools widely available would benefit both communities.

In summary, we should aim to exploit as much as possible the complementarity of the two survey strategies, namely, with LSST going wide first, building up depth over 10 years, while *Euclid* goes to full depth first, while collecting area over time. The two surveys also have complementary data release philosophies, with both planning to release worldwide significant data sets after the first few years of operations, with additional releases thereafter. Some coordination in these releases would greatly benefit the global community of astronomers, especially in the location of the LSST “deep drilling fields” and the *Euclid* deep fields, as well as the *Euclid* “quick release” fields.

5.5. Data Products and Information Exchange

As described above, scientific coordination between *Euclid* and LSST has many aspects and can be pursued to different depths. In order to maintain flexibility and facilitate information exchange, a joint effort might want to design the implementation of joint processing of data from the two missions in tiers of scientific and technical complexity. The tiers lend themselves naturally to sequential implementation. This allows different joint processing tasks to be spun off at different tiers, with the more complex tasks addressing more intricate science questions relying on more advanced tiers.

Tier 0 or Data Infrastructure would be aimed primarily at basic comparison of the survey data products for sanity checks. This tier would start with organizing data for faster access, for instance by coordinating the indexing schemes or regridding the images so pixels align across surveys. This tier would also include cross-checks on the photometry, astrometry, and depths of surveys to detect any systematic offsets, understand them, and correct for them. This tier would rely primarily on the extraction catalogs from the two surveys to compare astrometry and photometry, but would verify the image product alignments in image space, i.e., in pixel space. This tier may well be limited to point-source photometry in order to avoid the complications of extended object estimation. The products from Tier 0 would include astrometric distortion maps that represent the differences between the two surveys and photometric offset maps that represent the coefficients needed to convert photometry between the two surveys. The challenge here is to design data structures that capture the departures and systematics without forcing either one of the two data sets to conform to the other.

Tier 1 or Cross-survey Associations would connect individual entries in the catalogs across the two missions, establishing one-to-one as well as one-to-many mappings. This would require coordination on the databases and their designs. In establishing the mappings, positional and photometric comparisons would be used in the first round, and other parameters would be added if needed in later rounds. These associations would take into account the effects of varying spatial resolution

in the different surveys and the astronomical properties of galaxies, such as their extent on the sky. They would characterize the severity of confusion for objects or parts of the sky. The products from Tier 1 would include association tables listing for each object in survey A all of the related objects in survey B, with annotation about the details of the relations, e.g., significance of spatial coincidence, match in brightness, index of local crowding in either survey. Ideally, a single such table would collect all of the information in a symmetric treatment of both surveys.

Tier 2 or Higher-level Checks would address the comparison of more advanced extracted information such as shape measurements or colors of objects. The products from Tier 2 would include estimation offset maps equivalent to those described in Tier 0 for the photometry.

Tier 3 or Science Analysis would implement pixel-level joint processing, running algorithms that operate simultaneously on images from both surveys, using a variety of algorithms from multi-image estimators to model parameter estimation based on multiple images and a sky and object model. There are clearly many different branches in this Tier, and they will yield different products, depending on the question being asked. Obvious examples include blended object disambiguation; photometric/morphological decomposition, with application to photo- z estimation; or using colors to improve shape estimation in very-wide-band images. Tier 3 products could include the products of scientific analyses and could be in the form of parameters attached to objects in the images or in the form of synthesized images resulting from joint pixel-level processing. Much of the benefit of LSST/*Euclid* coordination described in this paper will rely on eventual coordination at the Tier 3 level, including the possibility of updating the association tables constructed in Tier 1 with the joint result of higher-level analyses.

Lower Tiers (0 and 1) of processing feed into the higher Tiers, and could be implemented infrequently to benefit all subsequent processing, if care is taken to keep the focus on the basic properties of the data rather than specializing to specific science goals. Tiers 0 and 1 products will support all research using both surveys, not just dark energy cosmology. By contrast, Tiers 2 and 3 will tend to be specialized by scientific question and will need to accommodate many different ways of processing the data.

6. Computational Issues

In this section, we discuss the computational resources and infrastructure that might be needed to perform joint pixel-level processing of *Euclid* and LSST data.

The value of *Euclid* data to LSST includes combining the space-based image data with the LSST images. The addition of *Euclid* data might be reasonably straightforward for LSST processing, where it could be added into the processing as if it were a few extra visits, albeit with higher resolution and significant undersampling. The computational overhead for LSST might be minimal, with the main issues based on the technical understanding of the *Euclid* data. This technical understanding would be best enabled by close cooperation of the *Euclid* data processing “Science Ground Segment (SGS)” and the LSST data center(s).

In the standard external data processing paradigm of *Euclid*, the data are first “Euclidized”—converted to a common

photometric and astrometric calibration standard, quality assessed, weights and masks are generated, and converted to a *Euclid*-compatible format. This is a standard data combination procedure of building the “transfer” function that will make external data compatible with the reference systems that will be used in *Euclid* and is consistent with the quality requirements for such data. These data then enter the rest of the *Euclid* pipeline.

Processing images that have already been co-added through the LSST pipeline independently from *Euclid* data is not optimal; however, the benefit of a full co-processing still needs to be assessed quantitatively. To realize those benefits, we will need to carry out multi-image source fitting, where the characteristics of a source are simultaneously fit on all images, even those where the source is below the detection threshold, whereas the source position is imposed from a global astrometric solution with optimized deblending algorithms (see, for instance, Section 2.1).

For *Euclid*, the LSST data could be the largest ingestion of external data, if all (or even many) of the LSST visits are ingested. To process at the individual image level would require around a factor 50 increase in raw data volume and a factor of around 20 increase in the required processing power compared to the present DES. This is possible within the context of the *Euclid* SGS but would require an increase in capacity that would depend on how much of the LSST data were ingested. However, it is likely that several of the benefits of joint pixel-level processing can already be realized if *Euclid* works with co-added LSST data, and only to the depth of the *Euclid* data. This imaging depth would be available within the first few years of LSST, which would reduce the computational requirements for *Euclid* ingesting LSST data considerably.

The Euclidization process could be carried out in two ways. The standard Euclidization steps could be integrated into the LSST pipeline so that LSST produced the appropriate calibrated, quality-controlled, weighted, and masked LSST data for *Euclid*, achieving software reuse on the *Euclid* side, and resource reuse for a mostly i/o dominated process. An alternative would be to deploy specific LSST software within the *Euclid* SGS. Again, this would allow software to be reused, but potentially require significantly more processing in *Euclid*, unless only stacked data are used; even then, custom *Euclid* code might be required depending on LSST’s choice of stacking algorithm. Either approach would benefit from some synchronization of LSST and *Euclid* processing to control computing costs.

The primary difficulties for LSST in handling *Euclid* data is the undersampling of the NISP images and the probably correlated noise resulting from resampling as well as the nature of NIR detectors. Other effects (e.g., chromatic PSFs, band-dependent morphologies) are already present in LSST images and will be accounted for by the LSST pipelines. For VIS data, LSST would ingest individual images with instrumental effects removed (bias, flatfields, nonlinearity, brighter-fatter, cross-talk, etc.) and would probably also employ the *Euclid* astrometric and photometric modeling, along with the PSF model. For NISP data, LSST would ingest the combined dithered images and possibly also the processed individual images. In either case, LSST would employ *Euclid* calibrations including the PSF models—PSF estimation in undersampled data would be a significant extension of the LSST data

processing and would presumably duplicate work already performed by *Euclid*.

LSST is considering a variety of algorithms for combining images, and some of the options would not easily handle the individual NISP exposures; fortunately, hybrid schemes that could accommodate NISP data appear simple enough to implement. The LSST image-processing algorithms already need to handle a factor of 2.5 in wavelength (350 nm–1 μ m) and extending this to NIR data may be merely a matter of degree, depending on the difficulties imposed by the undersampling in NISP data. Additionally, the LSST data are deep enough that adding *Euclid* data is not expected to add significantly to the number of objects being characterized.

It is likely that both data sets will need to be ingested into both pipelines, given that the processing methodologies, driven by the different scientific use of the data, could well be very different. This approach would also address some of the concerns about loss of independence of the two pipelines and cross-checking of results and cosmological constraints. A deciding factor here will involve the trade-offs between additional costs of code adaption and integration, processing and storage, and the savings from sharing code, storage, and processing. How the mutual ingestion of data will be carried out depends on some currently uncertain issues:

1. Does *Euclid* need to Euclidize all of the LSST data, or if the main application is colors for photometric redshifts, does it only require a subset (overlapping on sky and depth, stacked data) of the LSST data?
2. Does the Euclidization step for LSST data most naturally lie within the *Euclid* or LSST pipelines? What algorithms will need to be adapted or developed in LSST to ingest *Euclid* data, and what will need to be adapted or developed in either pipeline for Euclidization?
3. How much duplication of effort is both desirable and inevitable?
4. What are the difficulties associated with deblending and undersampling?

The first issue will drive the subsequent items and so a firm understanding of what is needed from both surveys needs to be clarified. The size of the LSST data to be ingested into *Euclid* will also determine the additional storage and processing needs of the *Euclid* SGS.

One way to address those issues is to create a common *Euclid*/LSST software working group composed of software and computing experts from both projects and to carry out a realistic set of tests using simulated data. This group would then interact with both *Euclid* and LSST software groups in order to adapt the existing pipelines or develop new ones. Some of the tests should also be designed in order to assess the impact of such a common processing on the hardware infrastructure. The common *Euclid*/LSST software working group should also work in close contact with both projects' science working groups in order to design the tests for realistic use cases and to maximize the scientific return.

7. Conclusions

We outlined in this paper many reasons why *Euclid* and LSST may wish to collaborate and coordinate, going well beyond an earlier white paper on dark energy mission coordination (Jain et al. 2015). We outlined the numerous scientific benefits that will come from such coordination in pixel-level processing, cadence

and survey overlap, calibration data acquisition, and resource allocation (especially in high-performance computing). The benefits we explored are primarily in dark energy cosmology, but we touched on other areas of scientific benefit, including galaxy evolution, dark matter studies, and solar system science. We offered a plausible path forward that will require further scientific development by the community to realize, as well as coordinated discussions between the two projects and the relevant funding agencies. The two projects will have to come to some agreement on the political issues surrounding data access to enable the benefits outlined in this paper. We feel that this path will lead to the maximum scientific return from *Euclid* and LSST in the 2020s and into the 2030s.

J.R. and A.K. were supported by JPL, which is run under a contract for NASA by Caltech. J.R., A.K., P.C., R.B., R.L., and R. Mandelbaum were supported in part by NASA ROSES grant 12-EUCLID12-0004. A.N.T. is supported by a UK Space Agency *Euclid* grant and a Royal Society Wolfson Research Merit Award. Y.M. was supported by the French CNES Space Agency and Institut National des Sciences de l'Univers of CNRS. T.S. acknowledges travel support from the German Federal Ministry for Economic Affairs and Energy (BMWi) provided via DLR under project No. 50QE1103. T.K. is supported by a Royal Society University Research Fellowship. V.F.C. is funded by the Italian Space Agency (ASI) through contract *Euclid*—IC (I/031/10/0) and acknowledges financial contribution from the agreement ASI/INAF/I/023/12/0. R. Massey acknowledges support from UK STFC grant ST/N001494/1. B.N. is supported by a UK Space Agency *Euclid* grant. This paper benefited from ideas developed in the UK Dark Energy Strategy 2020 Document.³⁷ We thank Michael Brown, Elisa Chisari, Henk Hoekstra, Željko Ivezić Bhuvnesh Jain, Roberto Scaramella, and Peter Schneider for useful feedback on this paper. We thank Katrin Heitmann for useful discussions on high-performance computing. We thank the anonymous referee for useful suggestions on adding quantitative rigor to sections of this paper and useful feedback on the structure of the paper.

Appendix Thoughts on Coordination

A.1. Interdependency and Independence

A key question is what level of cooperation is appropriate at each stage of the development of these experiments. One extreme scenario is that the two experiments work in complete isolation and simply exchange final results, already fully analyzed independently, for combined scientific constraints. Many recent breakthroughs in physics were carried out this way; for example, CMS and ATLAS (Aad et al. 2012; Chatrchyan et al. 2012), 2dFGRS and SDSS (Cole et al. 2005; Eisenstein et al. 2005), and the supernova dark energy experiments of Riess et al. (1998) and Perlmutter et al. (1999).

However, before combining the results of two experiments, the consistency between the individual results must be established. Such analyses may not have been “blinded” (see below), thus adding more possible uncertainty in such a comparison. In the event that an inconsistency arises between the experiments, then a joint study must be done to disentangle

³⁷ https://www.ucl.ac.uk/mssl/astro/Documents/UK_Dark_energy_strategy_2020/

the reasons for this inconsistency. Such tests, after the fact, then rely on the full coordination of both teams, who may have changed over the many years of these long-term projects.

At the other extreme, the two experiments could merge all parts of their data processing, thus entirely removing independence and creating a de facto single experiment.

Between these extremes lie a broad spectrum of possibilities that balance independence against potential gains of closer cooperation. Some potential gains are:

1. The combination of data with different unknown biases can lead to improved science through the reduction of systematic uncertainties in the measured parameters. Likewise, coordination in survey overlap and cadence can improve science output.
2. New science can be enabled through the combination of data and through cross-correlation statistics that are not possible with a single experiment alone.
3. An increase in the reproducibility of results. If two experiments share data analysis stages, then potential systematic effects arising from implementation differences between algorithms (for example) are mitigated.
4. The additional data from each experiment can enable (blind) validation tests which could not be carried out in isolation and without access to both data sets.
5. There can be efficiency savings in terms of sharing resources, particularly data or computationally intensive tasks.

The potential cost of cooperation between experiments is primarily a loss of independence. Independence is a critical aspect of the scientific methodology and underpins our confidence in our body of knowledge. If a Stage IV experiment discovers that our Standard Model of cosmology is incomplete, or uncovers some deviation from general relativity, then external replication of this result will be critical for scientific credibility.

Any consideration of cooperation between experiments is therefore a trade-off between replicability versus reproducibility, plus gains in new and improved science and efficiency.

Although it is only possible to independently replicate a measurement via an independent experiment, it is first necessary to clearly define “dependency” and “independency.” In this context, dependency between experiments is primarily gained via three routes. First, the experiments may analyze their data using the same methods, algorithms, and codes. Second, the experiments may use the same external calibration data or numerical simulations to test or train data analysis techniques. Finally, raw data products from the two experiments may be analyzed simultaneously (for example, in the expectation that information from one experiment may help to reduce systematic errors in the other).

For the specific case of LSST and *Euclid*, it is crucial to recognize that some interdependency between experiments is inevitable, by virtue of the sociology and size of the cosmology community, and the restricted set of available data. The two experiments already have a large (and growing) overlap in personnel and have been shaped by some of the same people. Arranging non-overlapping teams would now be impossible without major restructuring of consortia.³⁸ Moreover, *Euclid* and LSST will observe the same sky—there is only one

universe to observe—because the survey areas are large enough that overlap is inevitable.

Shared cosmological simulations are an obvious place for cost savings. Also, they will have to use similar external calibration data. For instance, only the *Hubble Space Telescope* (*HST*) has the capability of observing higher-quality data over a smaller area (prior to the launch of *WFIRST*), and both surveys will use *Gaia* (*Gaia* Collaboration et al. 2016) data for stellar SEDs and astrometric and photometric information. Both surveys will depend on deep spectroscopic data sets for photo-*z* calibration (e.g., Masters et al. 2015). Therefore, traditional methods of running separate experiments on different underlying data are not possible.

Nonetheless, we argue that the design characteristics of LSST and *Euclid* mean that coordination can be managed in a way that not only avoids some of the downsides of dependency, but addresses the uncontrolled dependencies that have emerged naturally.

1. The two experiments will use different instrumentation and observe from different environments and with different strategies: LSST from the ground at high cadence, and *Euclid* from space at high resolution. Therefore, even though personnel are overlapping, the analysis of raw data will necessarily require different algorithms.
2. Within each experiment, independence is already maintained between cosmological probes. For example, the methods, algorithms, and calibration data required for weak lensing, galaxy clustering, and supernova studies are to a large degree mutually exclusive.
3. *Euclid* and LSST have already established separate data analysis pipelines, built on quite different platforms and philosophies. Therefore, combined data analysis will be performed in addition to, not instead of, separate analyses of both data sets.

Furthermore, a commonly agreed strategy between the experiments on blinding methods, implementing validation tests, and open-source publishing of algorithms and raw data will lead to verifiable results. This may be challenging given the different cultures of the two experiments, but this argues for closer immediate cooperation to discuss and agree upon such procedures.

We must also address confirmation bias in cosmology, namely, that results tend to confirm investigators’ prior beliefs through unconscious selection of data and results. This is not merely a theoretical supposition, but has been observed in the literature (Croft & Dailey 2011). However, the perception of the extent of this problem is somewhat subjective, and scientists are arguably more immune to such psychological effects than most, because they are trained to be skeptical and question results even if they appear to agree. Blind analysis may be able to mitigate against this type of confirmation bias (e.g., Kuijken et al. 2015; Abbott et al. 2016; Hildebrandt et al. 2017; Wong et al. 2017; Zhang et al. 2017). These serious issues of scientific credibility and veracity of results are areas of common concern to experiments and therefore common cooperation on solutions is desirable.

Within the context of veracity of cosmological results, one important aspect of coordination is that of risk mitigation. Through coordination, various risks can be mitigated that could threaten to undermine the cosmological results from either

³⁸ At the time of this writing, the *Euclid* Consortium has over 1400 members and the LSST Dark Energy Science Collaboration has over 600.

experiment alone, e.g., if the two surveys yield inconsistent results, a natural step would be to compare the individual measurements in detail, to find out where, and why, they disagree. A risk in this scenario is that lack of coordination will lead to a protracted period of comparison during which the perception of the robustness of both results is undermined. This comparison could become acrimonious, increasing the possibility that neither team would want to appear to have the “wrong” answer.

This risk is mitigated by anticipating this scenario and planning for comparison steps that have been coordinated from the outset, e.g., by sharing a common coordinate system, defining a common patch of the sky, and/or even a common source detection catalog. The minimum level of cooperation required is therefore one of scenario planning and developing risk-mitigation strategies, as well as developing the political desire to promote such coordination.

A.2. What if We Do Not Coordinate

If the communities behind LSST and *Euclid* do not engage in an organized effort to conduct joint processing, the world will not stand still. There will be many small efforts attempting to fill this obvious gap, and they will meet with various degrees of success. The certain outcome of this situation will be the wasted effort resulting from the many independent attempts to recreate the same basic information (Tiers 0 and 1 as described in Section 5.5 at least, and definitely more). To the further detriment of the field, the lack of such a collaborative effort will limit the exchange of deep expertise about the surveys, so that each independent effort at attempting the joint processing will fall short in its understanding of either one or the other data set, resulting in slower progress, distorted approaches, and potentially to wrong conclusions. Ultimately, the outcome will be less science on both sides, and less science from the combined data. Given the pressure to march on to the next stage of dark energy experimentation, this opportunity of joint processing may be lost forever, along with potential insights that would have informed the design of the next stage.

A.3. Culture

Both LSST and *Euclid* are large, international collaborations with many common managerial issues (e.g., publication policies, membership and reward tracking, meetings, etc.). Both experiments have developed different cultures partly due to their origins. We would hope that some sharing of good practice and experience in these cultural issues would be beneficial for both experiments.

Moreover, both experiments will have a long-term need for well-trained people to help analyze and understand the data and results. Coordination in the training of such early career researchers would be beneficial. Both projects will benefit from interchanges and joint training of future leaders in the advanced statistical and computational techniques required (i.e., data scientists). Such skills are also required for the global economy, and astronomy provides an excellent training environment.

ORCID iDs

Jason Rhodes <https://orcid.org/0000-0002-4485-8549>
 Christopher J. Conselice <https://orcid.org/0000-0003-1949-7638>

Andrew J. Connolly <https://orcid.org/0000-0001-5576-8189>
 Shoubaneh Hemmati <https://orcid.org/0000-0003-2226-5395>
 Lynne Jones <https://orcid.org/0000-0001-5916-0031>
 Robert Lupton <https://orcid.org/0000-0003-1666-0962>
 Rachel Mandelbaum <https://orcid.org/0000-0003-2271-1527>
 Peter Melchior <https://orcid.org/0000-0002-8873-5065>
 Jeffrey A. Newman <https://orcid.org/0000-0001-8684-2222>
 Brant Robertson <https://orcid.org/0000-0002-4271-0364>
 Michael A. Strauss <https://orcid.org/0000-0002-0106-7755>
 Anja Von Der Linden <https://orcid.org/0000-0002-3881-7724>

References

- Aad, G., Abajyan, T., Abbott, B., et al. 2012, *PhLB*, 716, 1
 Abbott, T., Abdalla, F. B., Allam, S., et al. 2016, *PhRvD*, 94, 022001
 Abraham, R. G., Ellis, R. S., Fabian, A. C., Tanvir, N. R., & Glazebrook, K. 1999, *MNRAS*, 303, 641
 Abramo, L. R., Secco, L. F., & Loureiro, A. 2016, *MNRAS*, 455, 3871
 Aihara, H., Arimoto, N., Armstrong, R., et al. 2017, arXiv:1704.05858
 Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12
 Albrecht, A., Amendola, L., Bernstein, G., et al. 2009, arXiv:0901.0721
 Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591
 Amon, A., Heymans, C., Klaes, D., et al. 2017, arXiv:1707.04105
 Aubourg, É., Bailey, S., Bautista, J. E., et al. 2015, *PhRvD*, 92, 123516
 Bacon, D., Bridle, S., Abdalla, F. B., et al. 2015, arXiv:1501.03977
 Benítez, N. 2000, *ApJ*, 536, 571
 Benson, B. A., Ade, P. A. R., Ahmed, Z., et al. 2014, *Proc. SPIE*, 9153, 1
 Bentz, M. C., Walsh, J. L., Barth, A. J., et al. 2009, *ApJ*, 705, 199
 Bernstein, G. M., Armstrong, R., Krawiec, C., & March, M. C. 2016, *MNRAS*, 459, 4467
 Berthier, J., Carry, B., Vachier, F., Eggl, S., & Santerne, A. 2016, *MNRAS*, 458, 3394
 Bonvin, V., Courbin, F., Suyu, S. H., et al. 2017, *MNRAS*, 465, 4914
 Brown, M. J. I., Moustakas, J., Smith, J.-D. T., et al. 2014, *ApJS*, 212, 18
 Carry, B. 2017, *A&A*, submitted (arXiv:1711.01342)
 Carry, B., Solano, E., Eggl, S., & DeMeo, F. E. 2016, *Icar*, 268, 340
 Chatrchyan, S., Khachatryan, V., Sirunyan, A. M., et al. 2012, *PhLB*, 716, 30
 Chiang, Y.-K., Overzier, R., & Gebhardt, K. 2014, *ApJL*, 782, L3
 Chisari, N. E., Dvorkin, C., Schmidt, F., & Spergel, D. N. 2016, *PhRvD*, 94, 123507
 Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, *ApJL*, 648, L109
 Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, *MNRAS*, 362, 505
 Croft, R. A. C., & Dailey, M. 2011, arXiv:1112.3108
 Cropper, M., Pottinger, S., Niemi, S., et al. 2016, *Proc. SPIE*, 9904, 99040Q
 Dahlen, T., Mobasher, B., Dickinson, M., et al. 2010, *ApJ*, 724, 425
 Dawson, W. A., Schneider, M. D., Tyson, J. A., & Jee, M. J. 2016, *ApJ*, 816, 11
 Delbo, M., Gayon-Markt, J., Busso, G., et al. 2012, *P&SS*, 73, 86
 Delgado, F., Saha, A., Chandrasekharan, S., et al. 2014, *Proc. SPIE*, 9150, 915015
 DeMeo, F., & Carry, B. 2013, *Icar*, 226, 723
 DeMeo, F. E., Binzel, R. P., Slivan, S. M., & Bus, S. J. 2009, *Icar*, 202, 160
 DeMeo, F. E., & Carry, B. 2014, *Natur*, 505, 629
 DES Collaboration, Abbott, T. M. C., Abdalla, F. B., et al. 2017, arXiv:1708.01530
 Dieleman, S., Willett, K. W., & Dambre, J. 2015, *MNRAS*, 450, 1441
 Dodelson, S., Heitmann, K., Hirata, C., et al. 2016, arXiv:1604.07626
 Āurech, J., Carry, B., Delbo, M., Kaasalainen, M., & Viikinkoski, M. 2015, in *Asteroid Models from Multiple Data Sources* (Tucson, AZ: Univ. Arizona Press), 183
 Āurech, J., Grav, T., Jedicke, R., Denneau, L., & Kaasalainen, M. 2005, *EM&P*, 97, 179
 Efstathiou, G., & Lemos, P. 2017, arXiv:1707.00483
 Eggl, S. 2011, *CeMDA*, 109, 211
 Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
 Friedrich, O., & Eifler, T. 2017, arXiv:1703.07786
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, 595, A1
 Graham, M. L., Connolly, A. J., Ivezić, Ž., et al. 2017, arXiv:1706.09507
 Grogan, N., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35

- Habib, S., Roser, R., Gerber, R., et al. 2016, arXiv:1603.09303
- Harnois-Déraps, J., Tröster, T., Chisari, N. E., et al. 2017, *MNRAS*, **471**, 1619
- Harrison, I., Camera, S., Zuntz, J., & Brown, M. L. 2016, *MNRAS*, **463**, 3674
- Harvey, D., Massey, R., Kitching, T., Taylor, A., & Tittley, E. 2015, *Sci*, **347**, 1462
- Hearin, A. P., Zentner, A. R., Ma, Z., & Huterer, D. 2010, *ApJ*, **720**, 1351
- Heavens, A., Sellentin, E., de Mijolla, D., & Vianello, A. 2017, *MNRAS*, **472**, 4244
- Henderson, S. W., Allison, R., Austermann, J., et al. 2016, *JLTP*, arXiv:1510.02809
- Hildebrandt, H., Arnouts, S., Capak, P., et al. 2010, *A&A*, **523**, A31
- Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, *MNRAS*, **465**, 1454
- Hoekstra, H., Herbonnet, R., Muzzin, A., et al. 2015, *MNRAS*, **449**, 685
- Hoekstra, H., Viola, M., & Herbonnet, R. 2017, *MNRAS*, **468**, 3295
- Huff, E., & Mandelbaum, R. 2017, arXiv:1702.02600
- Insera, C. E. A., Nichol, R. C., Scovacicchi, D., et al. 2017, arXiv:1710.09585
- Ivezic, Z., Lupton, R. H., Jurić, M., et al. 2002, *AJ*, **124**, 2943
- Ivezic, Z., Tabachnik, S., Rafikov, R., et al. 2001, *ApJ*, **122**, 2749
- Ivezic, Z., Tyson, J. A., Abel, B., et al. 2008, arXiv:0805.2366
- Jain, B., Jarvis, M., & Bernstein, G. 2006, *JCAP*, **2**, 001
- Jain, B., Spergel, D., Bean, R., et al. 2015, arXiv:1501.07897
- Joachimi, B., Cacciato, M., Kitching, T. D., et al. 2015, *SSRv*, **193**, 1
- Joudaki, S., Blake, C., Johnson, A., et al. 2017a, arXiv:1707.06627
- Joudaki, S., Mead, A., Blake, C., et al. 2017b, *MNRAS*, **471**, 1259
- Kaiser, N. 2001, Addition of Images with Varying Seeing, Tech. Rep. PSDC-002-011-00
- Kartalpe, J. S., Mozena, M., Kocevski, D., et al. 2015, *ApJS*, **221**, 11
- Kelly, P. L., Filippenko, A. V., Burke, D. L., et al. 2015, *Sci*, **347**, 1459
- Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2010, SNANA: A Public Software Package for Supernova Analysis, Astrophysics Source Code Library, ascl:1010.027
- Kitching, T., Bacon, D., Brown, M., et al. 2015, arXiv:1501.03978
- Kitching, T. D., Verde, L., Heavens, A. F., & Jimenez, R. 2016, *MNRAS*, **459**, 971
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, **197**, 36
- Krause, E., & Eifler, T. 2017, *MNRAS*, **470**, 2100
- Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, *MNRAS*, **454**, 3500
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, **224**, 24
- Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, *ApJ*, **722**, 566
- Lanyon-Foster, M. M., Conselice, C. J., & Merrifield, M. R. 2007, *MNRAS*, **380**, 571
- Lanyon-Foster, M. M., Conselice, C. J., & Merrifield, M. R. 2012, *MNRAS*, **424**, 1852
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv:1110.3193
- Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, *ApJ*, **744**, 159
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, **538**, 473
- Lima, M., Cunha, C. E., Oyaizu, H., et al. 2008, *MNRAS*, **390**, 118
- Lintott, C., Schawinski, K., Bamford, S., et al. 2011, *MNRAS*, **410**, 166
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
- Mantz, A. B., von der Linden, A., Allen, S. W., et al. 2015, *MNRAS*, **446**, 2205
- Masters, D., Capak, P., Stern, D., et al. 2015, *ApJ*, **813**, 53
- Masters, D., Stern, D., Cohen, J., et al. 2017, *ApJ*, **841**, 111
- Newman, J. A. 2008, *ApJ*, **684**, 88
- Newman, J. A., Abate, A., Abdalla, F. B., et al. 2015, *Aph*, **63**, 81
- Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, *ApJS*, **208**, 5
- Olmstead, M. D., Brown, P. J., Sako, M., et al. 2014, *AJ*, **147**, 75
- Oszkiewicz, D. A., Bowell, E., Wasserman, L. H., et al. 2012, *Icar*, **219**, 283
- Peacock, J. A., Schneider, P., Efstathiou, G., et al. 2006, arXiv:astro-ph/0610906
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, **517**, 565
- Petit, J.-M., Kavelaars, J. J., Gladman, B. J., et al. 2017, *AJ*, **153**, 236
- Raddick, M. J., Bracey, G., Gay, P. L., et al. 2010, *AEdRv*, **9**, 010103
- Rahman, M., Ménard, B., Scranton, R., Schmidt, S. J., & Morrison, C. B. 2015, *MNRAS*, **447**, 3500
- Randall, S. W., Markevitch, M., Clowe, D., Gonzalez, A. H., & Bradač, M. 2008, *ApJ*, **679**, 1173
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, **116**, 1009
- Rigault, M., Aldering, G., Kowalski, M., et al. 2015, *ApJ*, **802**, 20
- Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015, *A&C*, **10**, 121
- Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, *ApJ*, **785**, 104
- Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, *ApJ*, **742**, 61
- Schaan, E., Krause, E., Eifler, T., et al. 2017, *PhRvD*, **95**, 123512
- Schrabback, T., Applegate, D., Dietrich, J. P., et al. 2017, *MNRAS*, stx2666
- Scovacicchi, D., Nichol, R. C., Macaulay, E., & Bacon, D. 2017, *MNRAS*, **465**, 2862
- Seljak, U. 2009, *PhRvL*, **102**, 021302
- Sellentin, E., & Heavens, A. F. 2016, *MNRAS*, **456**, L132
- Sellentin, E., & Heavens, A. F. 2018, *MNRAS*, **473**, 2355
- Semboloni, E., Hoekstra, H., Huang, Z., et al. 2013, *MNRAS*, **432**, 2385
- Sheldon, E. S., & Huff, E. M. 2017, *ApJ*, **841**, 24
- Simmons, B. D., Lintott, C., Willett, K. W., et al. 2017, *MNRAS*, **464**, 4420
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, **214**, 24
- Solano, E., Rodrigo, C., Pulido, R., & Carry, B. 2014, *AN*, **335**, 142
- Speagle, J. S., & Eisenstein, D. J. 2017a, *MNRAS*, **469**, 1186
- Speagle, J. S., & Eisenstein, D. J. 2017b, *MNRAS*, **469**, 1205
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757
- Spoto, F., Milani, A., & Knežević, Z. 2015, *Icar*, **257**, 275
- Stickley, N. R., Capak, P., Masters, D., et al. 2016, arXiv:1606.06374
- Taylor, A., Joachimi, B., & Kitching, T. 2013, *MNRAS*, **432**, 1928
- Voigt, L. M., Bridle, S. L., Amara, A., et al. 2012, *MNRAS*, **421**, 1385
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *PhR*, **530**, 87
- Welikala, N., Connolly, A. J., Hopkins, A. M., Scranton, R., & Conti, A. 2008, *ApJ*, **677**, 970
- Wong, K. C., Suyu, S. H., Auger, M. W., et al. 2017, *MNRAS*, **465**, 4895
- Wood-Vasey, W. M., Friedman, A. S., Bloom, J. S., et al. 2008, *ApJ*, **689**, 377
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, **120**, 1579
- Zackay, B., & Ofek, E. O. 2015, arXiv:1512.06879
- Zhang, B. R., Childress, M. J., Davis, T. M., et al. 2017, *MNRAS*, **471**, 2254