Final Report

AstroDeep Final report

On behalf of the ASTRODEEP consortium

ASTRODEEP

"Unveiling the power of the deepest images of the Universe"

THEME [SPA.2012.2.1-01]

[Exploitation of space science and exploration data]

Grant agreement for: Collaborative project

Grant agreement no: 312725

ABSTRACT

We present here the results of the ASTRODEEP project

Prepared by: A. Fontana, D. Elbaz, J. Dunlop, S. Derriere
Approved by: AEC
Date: 27/03/2017

Executive summary

We designed ASTRODEEP as a co-ordinated and comprehensive programme of algorithm development, data release, and scientific data analysis dedicated to study the birth and evolution of galaxies in the first few billion years of cosmic history. Our aim was to extract and exploit the full wealth of information contained within the deep multi-frequency datasets already collected by the best astronomical satellites and telescopes, and to share our results and expertise with the rest of the international community.

Our major "enemy" is *confusion*, i.e. the inevitable overlap between sources that exist in most deep astronomical imaging, due to the limited spatial resolution of several instruments.

We have taken a unified and comprehensive approach to this problem, studying how it affects different datasets, and developing methods to use the highest-quality dataset as a template with which to extract the full information contained in the lower-quality/resolution data. We have then extensively applied these methods to a large variety of datasets. Our methods and data have been made public and can be used by the entire community.

Specifically, we have obtained the following results:

Algorithm development. We have developed software tools and methods that incorporate the results of our experiments and tests, and the resulting products represent the current state-of-the-art for this kind of scientific data processing, namely:

- *T-PHOT*, a new public software product designed to extract robust photometry based on the prior knowledge of the position and shape of objects. We have used this software extensively, with *Hubble Space Telescope* (*HST*) images as inputs, to extract photometric measurements from the data supplied by other satellites and ground-based instruments.
- *EGG*, a public galaxy simulator code, which we have developed and utilised to produce simulated data against which we have tested our algorithms;
- A pipeline for the removal of the bright cluster galaxies in the Hubble Frontier Fields.
- Specific methods to clean the list of prior positions to be used in the analysis of the far-IR or sub-mm imaging as supplied by the *Herschel Space Observatory*, or SCUBA-2 on the James Clerk Maxwell Telescope (JCMT).

Data analysis and release. Using our sophisticated tools we have analysed the most recent and highest-quality data released by public surveys with top instrumentation, and made our catalogues available worldwide. Specifically, we have:

- Obtained and released imaging data sets, most notably the VLT K_S-band data from the HUGS survey and an improved reduction of the *Herschel* data set on the GOODS fields;
- Produced a new multi-wavelength catalogue for the GOODS-S field, now including 43 photometric bands from the UV to the *Spitzer* (mid-infrared) bands, that significantly improves upon the previous existing public catalogues in both the depth of the images and in the quality of the analysis;
- Created an optimally de-blended *Herschel* catalogue for the GOODS-S and GOODS-N fields;
- Created an optimally de-blended SCUBA-2 sub-mm catalogue within the CANDELS fields;
- Constructed a prior-based photometric catalogue from the *Chandra* 4Ms data in the GOODS-S field;
- Produced full photometric catalogues (including *HST*, VLT K_S-band and *Spitzer* data) of the first four Hubble Frontier Fields;
- Completed a public release of an extensive set of simulated images reproducing the CANDELS and COSMOS data sets;
- Simulated the imaging and spectroscopic data expected from the deep field of the *EUCLID* mission, and hence re-evaluated the anticipated power of the *EUCLID* Deep Survey.

Scientific analysis. A significant fraction of our effort has been dedicated to the study of the evolution of galaxies at high redshift, in particular, albeit not exclusively, using our new data analysis tools and the resulting catalogues. We have addressed many topics, including:
- The study of galaxy evolution via statistical representations of the galaxy population, such as luminosity or stellar-mass functions, and exploring the resulting implications for our understanding of the process of cosmic re-ionization;
- The search for, and study of star-forming galaxies at high redshift, especially those in which the light from young stars is strongly absorbed by dust; such galaxies evade discovery in traditional UV-based surveys, but in fact constitute a large fraction of the total star-formation rate (SFR) budget;
- The study of quiescent galaxies at high redshift;
- The assessment of the morphological types of high-redshift galaxies and its evolution;
- The study of the metallicity and ionization status of galaxies, over a range of redshifts and at stellar masses, in order to place new constraints on the process of chemical evolution and enrichment over cosmic time;
- The presence and role of faint active galactic nuclei (AGN) in apparently "normal" galaxies in the early Universe.

Dissemination. We have taken great care to properly promote and advertise the results of ASTRODEEP, and have succeeded to the point of making it an acknowledged "logo" in extragalactic astronomy.

We have organized three international conferences that have proved key to the effective dissemination of our results, and have given enhanced credibility to our public releases.

We have developed a new public interface for the interactive query of our catalogues; this has been developed by our partner CDS – the European leader in astronomical data preservation - and will remain available on the CDS servers.

We have developed and maintained an attractive and accessible website, that will remain active for a long time after the end of the project.

We have prepared a number of "trademarked" images/pictures and movies that will be used for future press releases and *James Webb Space Telescope* (*JWST*) proposals.

Overall, we believe that ASTRODEEP has largely met its original goals. Our understanding and treatment of the problems due to confusion has significantly advanced. The tools that we developed, refined and delivered are widely acknowledged as the state-of-the-art, and have set a new reference standard in the field. The data and catalogues that we have created and delivered are of the highest quality and – crucially – have been focused on data that are at the frontier of current extragalactic astrophysics, and hence are most eagerly awaited by the community.

With more than 60 published refereed papers (collecting 525 citations as of March 2017) in major international journals, the impact of ASTRODEEP in the astronomical community is now well established and should be long lasting. Many of these science papers have been led by the post-doctoral researchers that we recruited to work on the ASTRODEEP project. For most of them, ASTRODEEP has provided an important opportunity and springboard for the future advancement of their careers, as well as providing invaluable experience of collaboration within a genuinely international environment, at the cutting edge of modern astrophysical research. We look forward to working again with them in the near future.

Summary description of project context and objectives

1. The original proposal

The original ASTRODEEP proposal – that traces back to late 2011- was designed around a few very clear goals. We proposed a coordinated and comprehensive program of i) algorithm/software development and testing; ii) data reduction/release, and iii) scientific data validation/analysis aimed at making Europe the world leader in the exploitation of the deepest multi-frequency data from the major space and ground-based observatories.

A huge amount of time and resources have been, and still are being invested in obtaining ultra-deep astronomical data, in a series of key extra-galactic survey fields, over a wide range of wavelengths. These large-scale, often public observing programs are motivated by the desire to understand the formation and evolution of galaxies and super-massive black holes from the very earliest times, and to answer fundamental questions in cosmology such as how and when the universe was re-ionized.

These deep surveys have been successfully coordinated and are focused on a set of recognized key deep fields. From space, the deepest UV/optical/near-infrared data are obtained by the *Hubble Space Telescope* (with exquisite angular resolution), unparalleled mid-infrared data are delivered by *Spitzer*, far-infrared imaging by *Herschel,* and imaging and spectroscopy at X-ray wavelengths by *Chandra* and *XMM-Newton*. Remarkably, all of these space observatories are still functioning and producing world-class data, but their final major output will probably emerge over the next 2-3 years. Meanwhile, these space surveys are supported by major ground-based programs of optical-infrared imaging/spectroscopy (with ESO's VLT and VISTA telescopes in Chile, and Keck, Subaru, UKIRT in Hawaii), sub-mm observations (with Laboca & ALMA in Chile and SCUBA-2 in Hawaii), and radio imaging (with the VLA/EVLA in the US, and e-Merlin/LOFAR in Europe). The legacy value of this data set is immense.

This impressive level of multi-facility co-ordination reflects the widely-shared appreciation that a complete understanding of cosmic evolution requires a genuinely multi-frequency approach to the study of the young universe. However, while exciting results have already emerged from these surveys, it is abundantly clear that their full potential cannot be unlocked unless the data can be properly and robustly combined and delivered to the wider astronomical community, in a form that easily facilitates full scientific analysis and exploitation of the complete multi-frequency database.

Our proposal aimed to achieve this. We assembled a team of internationally-renowned scientists (from Italy, UK, France and USA) who are leaders in many of the above-mentioned surveys (with intimate knowledge of, and access to, the data), and who were already working actively to overcome the challenge of combining all of the available multi-frequency survey data to produce a robust, comprehensive, science-ready dataset.

We clearly identified that such challenge had several aspects that needed to be addressed, and hence constructed a carefully-designed program involving parallel/staged work packages defining key deliverables in:

Algorithm development and data processing. In order to reap the extraordinary scientific benefits of this unparalleled technology investment, a number of key technical problems first needed to be overcome, and a major part of the proposal was devoted to solving these technical questions. The most important challenge comes from the fact that the data are delivered by very different kinds of telescopes having widely different angular resolutions. This is most explicitly an issue for very deep surveys, which yield a high surface-density of detections, often resulting in "highly-confused" images. This results in conceptual and technical problems in a number of crucial areas, which include:

• the unambiguous identification and definition of individual objects;
• the measurement of accurate multi-frequency photometry in crowded fields from images with very different resolution and depth;
• the optimal extraction of the wealth of information contained in deep slit-less spectra obtained from space, exploiting the prior information available from ultra-deep imaging.

Before ASTRODEEP, members of our team had already developed and applied algorithms to solve or mitigate these problems. ASTRODEEP allowed us to join efforts to develop, refine and combine these algorithms, test and properly validate their effectiveness on the deepest multi-frequency data, and apply them to produce uniquely powerful homogenized data products.

Dissemination We also envisaged an aggressive dissemination plan to release our results to the wider European and world-wide astronomical community in a timely manner. The delivered data were expected to be in the form of a) final stacked images, b) multi-wavelength catalogues, spectral energy distributions, and photometric/spectroscopic redshifts, and c) derived rest-frame parameters (e.g. stellar masses, star-formation rates). We also planned to disseminate the software specifically developed for our analysis. We have ensured professional high-quality, properly-documented data products by engaging the technical expertise of units in France, UK and Italy which have vast experience of serving Virtual Observatory (VO)-compatible data products to the European community.

Scientific validation/exploitation A key feature of our approach is the recognition, based on extensive experience, that the development of new analysis methods, and data reduction must proceed in parallel with on-going attempts at scientific exploitation. This is essential for the efficient and timely recognition of the strengths and weaknesses of different techniques, and for the proper validation of science-ready data products for public release; the best-quality data are generally produced by those most motivated by their scientific exploitation.

We ourselves proposed to use our database to address several of the outstanding issues in present-day extragalactic astronomy, including

• the first galaxies and cosmic re-ionization,
• the complete history of obscured and un-obscured star formation,
• the formation of the most massive galaxies, and
• the nature of the connection between the evolution of galaxies and their central super-massive black holes.

Preparation of the future *EUCLID* Deep survey Looking further ahead, we planned to apply our data analysis techniques and tools in preparation of the next generation of space surveys, especially ESA's planned *EUCLID* Deep Survey, where the images and spectra will be subject to the same observational biases as existing deep data, but will require dedicated tools to cope with a dataset that is 10^{2-3} times larger.

2 Revisiting our plans to keep focus on the main goals

As we show in the next section, we feel proud enough of our results to claim that we have successfully met the goals of the project. However, mentioning Wellington, "even the best designed plan does not survive the impact with the enemy". This is very true even in the development of a scientific project as ASTRODEEP: the forefront of scientific research is a constantly evolving (battle-)field, and we deployed an adaptive strategy to ensure the project remained relevant and at the forefront of this rapidly evolving field. We have therefore continuously revisited our *specific goals* while keeping our focus on the *general goals* of the project. It may be instructive to discuss here how we can re-visit *a posteriori* our goals, and how we have revised and updated our specific plans to adapt the project to a continuously changing scientific landscape.

The *Herschel* mission, an ESA cornerstone that delivered unique far-IR data plagued by very poor resolution, has completed its lifecycle and has delivered the final data set. As is customary for space missions, the continuously improving knowledge of the instrument has enhanced the performance of the data reduction software adopted for the image processing. As a result, we decided to dedicate part of our resources to the re-processing of the *Herschel* maps with the Unicode software, a need that was not anticipated at the beginning of the project. This effort guarantees that our catalogues – derived from such images – can be considered as the "ultimate legacy" of this unique instrument.

Similarly, the *Spitzer* warm mission delivered unique and impressive data at shorter wavelengths. The total exposure time of *Spitzer* data over the few selected extragalactic fields that are our main focus has increased by factors of up to 5x-10x. We invested time and resources in pre-processing these data in order to include the final and deepest data set in our catalogues – again in order to maintain/ensure their "legacy" value.

A new, unprecedented effort of the *Hubble Space Telescope* – the so-called Frontier Fields Initiative – was completely unexpected at the time of the original proposal and revolutionized the field when ASTRODEEP was already ongoing. It is delivering unique ultra-deep data with *HST* (and complementary data from ground-based telescopes) over as many as 6 intermediate-redshift clusters and relevant un-lensed "parallel" nearby fields. The data are supposed to provide the astronomers with a first glimpse of ultra-faint sources, thanks to the magnification boost provided by the gravitational lensing of the clusters (i.e. bridging the gap between previous *HST* data and the future ultra-deep imaging anticipated from *JWST*). Such data are the main focus of current extragalactic deep astronomy, and represent at the same time an extraordinarily difficult data set to deal with, because of the complications resulting from the foreground galaxy clusters (such as contamination by bright cluster members, higher and irregular background, etc.). We decided to invest a significant amount of effort in the analysis of the Frontier Fields data, and the resulting images and catalogues now represent a major (and highly-cited) output from of ASTRODEEP.

Other projects have started or delivered results in the same time frame, often overlapping the list of specific datasets that we initially planned to analyze. The HELP project (P.I. S. Oliver) is another FP7-funded project that started one year after ASTRODEEP. It is focused on the analysis of *Herschel* extragalactic surveys over larger areas, including COSMOS. The VANDELS spectroscopic survey, led by two close collaborators of ours in Rome and Edinburgh (L. Pentericci and R. McLure) has been approved by ESO and has independently assembled high quality photometric catalogs over the wide CDFS and UDS fields. Considering these efforts, and our constrained manpower and resources, we decided to drop the analysis of the wider fields from our goals in order to focus on the smaller, ultra-deep fields. We considered these smaller and deeper fields as more topical, both in the light of the Frontier Fields initiative, and

because of the growing interest in connecting ultra-deep data to projects envisaged with *JWST* (now less that two years from launch).

The 3D-HST collaboration and the CANDELS collaboration have delivered in the meantime their own catalogues over most of the original CANDELS fields. In most cases the images analyzed represent the final data set available as of today, and the techniques adapted are well tested and are close to the ultimate performances that our state-of-the-art technique can obtain. As such we concluded that the improvement that ASTRODEEP was able to obtain was somewhat marginal and decided to degrade the re-analysis of the CANDELS fields to a lower priority compared to the new Frontier Fields data.

The GOODS-South field is an exception though. Here the wealth of new data gathered in the meantime – most notably the ultra-deep 100hr exposure with *Spitzer*, but also a wealth of intermediate band images from the ground, coupled with the unique *Herschel* and *Chandra* ultra-deep exposures – make this field unique. The existing public catalogues fall short of exploiting the full power of such data. We therefore decided to keep the GOODS-South field as the other main focus of our activity, both because of its scientific impact, but also because it provides in many ways the ultimate test-bed on which to best demonstrate the added value of our superior de-convolution techniques.

One way of summarizing this refinement and redirection of our goals is that we decided to focus on **proofs of concept and high-quality applications rather than quantity.**

Indeed, we invested much of our efforts into improving our de-convolution techniques in order to ensure we genuinely reached the limit in the amount of information that can be reliably extracted from the multi-frequency data. We then applied our new, but now tried and tested de-convolution software to the ultra-deep, high-impact data sets where these approaches are of most benefit, and likely to yield maximum scientific return through ASTRODEEP. We believe that this is the most important and long-lasting value that a project of the size and scope of ASTRODEEP can deliver. We note in particular that one of our main outputs (the public algorithm TPHOT, that we extensively used for our analysis) is much more than a re-written, improved version of the former T-FIT/ConvPhot routines. Thanks to extended testing and validation, we have been able to increase the confidence on this class of methods to the point where they are now much more widely accepted and applied by a much larger number of scientists. Similarly, the validation of the de-blending procedures in the extreme cases of the severe de-blending of the far-IR data (such as the imaging delivered by *Herschel* and *SCUBA-2*) made us much more confident in pushing these techniques to the limit, yielding a much larger number of detections in the *Herschel* deep fields.

During our work we also realized that the careful development and implementation of appropriate simulations for the validation of our procedures was much more important, and indeed much more demanding than originally anticipated. In the end, to overcome this challenge in a controlled way, we decided to develop a specific new simulation tool – named EGG – which we also made public. The performance and adaptability of this new software for the production of genuinely realistic and scalable simulated galaxy survey images and catalogues far surpasses that of any other tool currently on the market. This simulation work obviously required a level of additional effort that was not foreseen at the start of the project.

Finally, the design of the *Euclid* mission has significantly evolved in the intervening years. The instrumental concept has now at last been finalized and its performances are now better evaluated. The science cases have been revised and – most notably for us – only recently the strategy and organization of the Deep Survey have been finalized. Anticipating this evolution we had originally confined the activity on the *Euclid* mission to the last ASTRODEEP period. We have therefore been able to shape our data products (essentially simulated data) to mimic

the real anticipated *Euclid* data as accurately as possible, and indeed have been able to use them to re-define the scientific goals of the *Euclid* Deep Survey, and to re-establish/enhance the importance of the deep (grism) spectroscopic element of the *Euclid* mission.

Description of main S & T results/foregrounds

As already mentioned, ASTRODEEP stands over three pillars: i) algorithm/software development and testing; ii) data reduction and release, and iii) scientific data analysis. We describe here the main achievements of the project in these three main areas.

1. Algorithms and software developments

The activity related to the development of original algorithms and software has started immediately at the beginning of the project, and has been the focus of the first 2/3 of the project (mostly period #1 and #2). In the last part of the project we extensively used these tools, sometimes refining them, and dedicated much care in releasing and advertising them across the community.

1.1 EGG: the galaxy image simulator

The commonly adapted code to produce realistic high-resolution images is the SkyMaker program (E. Bertin), which we regularly used to test the various source extraction methods and algorithms developed within ASTRODEEP. In input, this program requires a simulated galaxy catalogue, which is produced by the Stuff program (also created by E. Bertin). The quality of these simulated catalogues is not optimal. In particular, the distribution of the
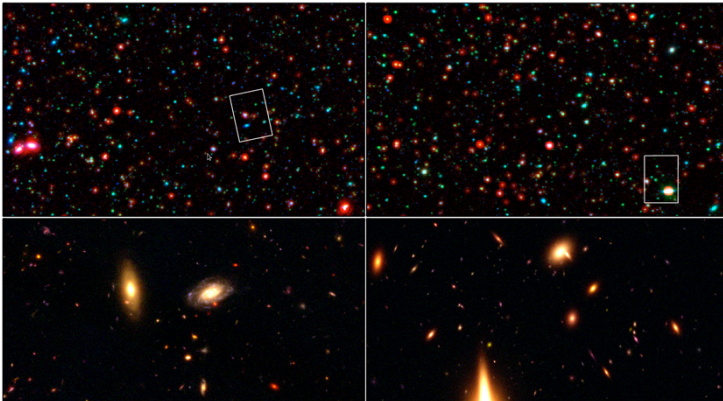

Figure 1 Left : real images of the GOODS-South field (top : false color image using Hubble H, Spitzer 4.5um, and Herschel 100um images ; bottom : false color image using Hubble B, Z and H images). Right : simulated images generated with EGG (same color bars).

simulated fluxes in some bands differ substantially from those that are observed, leading to simulated images that are not representative of the real Universe. Furthermore, Stuff can only generate fluxes in a limited, pre-defined list of broad band filters. Unfortunately, both SkyMaker and Stuff are poorly documented, and we cannot easily remedy to these problems. For this reason, we have developed a new tool to generate simulated galaxy catalogue, called **EGG** (the Empirical Galaxy Generator; Schreiber et al. 2017).

This new tool can generate catalogues in the format required by SkyMaker, and therefore can be used as a "drop-in" replacement for Stuff. Using this tool we are able not only to generate fluxes in all the photometric bands from 3000 Å to 8 µm, like Stuff, but we also merge in our technique to simulate far-IR fluxes from 8 µm to 3 mm, essentially covering, in a single tool, the whole wavelength range where stellar and dust emission dominate.

The main idea behind the generation process of this mock catalogue is that everything can be statistically inferred from the redshift, the stellar mass and the "star-forming/quiescent" flag of each galaxy. The procedure is therefore composed of two main steps: first, generate a

realistic distribution of galaxy masses at different redshifts both for star-forming and quiescent galaxies using observed mass-functions; second, estimate all the other physical properties using statistical recipes calibrated on the observed galaxies: morphology, SFR, attenuation, optical colours, and sky-projected position with clustering. Details on the precise recipes for the various steps can be found in a separate deliverable document (D3.2 – note that the code was dubbed *gencat* at the time we wrote the deliverable D3.2). In Figure 1we show an example of simulated images we have produced and illustrates the typical output of our simulations, which are now physically consistent from the UV to the far-IR.

Using these recipes, we show that we can create distribution of fluxes that are more realistic than that obtained by state-of-the-art complex semi-analytic models, particularly in the far-IR. In addition, the code is optimised to run very efficiently even on personal computers, and can generate an entire field like GOODS-South with tens of photometric bands in only a couple of minutes.

The drawback is that our galaxies are somewhat idealistic in shape: they are described as a two-components systems consisting of a smooth exponential disk (usually star-forming) and a de Vaucouleur bulge (usually quiescent). We therefore cannot produce clumpy or irregular galaxies, or differential dust attenuation (dust lanes, etc.). However this is also a limitation affecting SkyMaker, and fixing it would require a rewrite of this software as well. The major resulting limitation is that we EGG is not well suited to test morphological tools, given that the assumed models are too simplified and do not include irregular morphologies. In addition, our method to produce clustering in the galaxy positions is admittedly crude, and does not truly capture complex structures such as galaxy clusters or groups, and should not be used at that purpose.

1.2 TPHOT: a general tool for deep photometry and deblending

General description

T-PHOT is the first software tool delivered to the public from the AstroDeep project. Developed at INAF-OAR by Emiliano Merlin, T-PHOT is a software package aimed at extracting reliable photometry from extragalactic multiwavelength datasets, using the PSF-matching technique to de-confuse the images and obtain robust flux measurements even for severely blended sources in low resolution images. Inspired to its direct predecessors TFIT (Laidler et al. 2007) and CONVPHOT (De Santis et al. 2007), the first public version of T-PHOT (v1.5.7, released in October 2014; Merlin et al. 2015) merged their features into a much faster and more robust code. On top of that, T-PHOT improved on such codes, both in terms of performance and in the accuracy of the results. Version 2.0 has been released in September 2016 (Merlin et al. 2016b); it incorporates a number of new options and features to remove potential sources of errors, and extend the range of cases in which it is possible to use the code. T-PHOT is therefore a completely new and versatile tool, suitable for performing detailed photometry on images taken in a very broad range of wavelengths, not only in the optical domain but also in the FIR and sub-mm regimes where its performance is comparable to, or better than, other existing codes (e.g. FASTPHOT by Bethermin et al., or DESPHOT by Roseboom et al.).
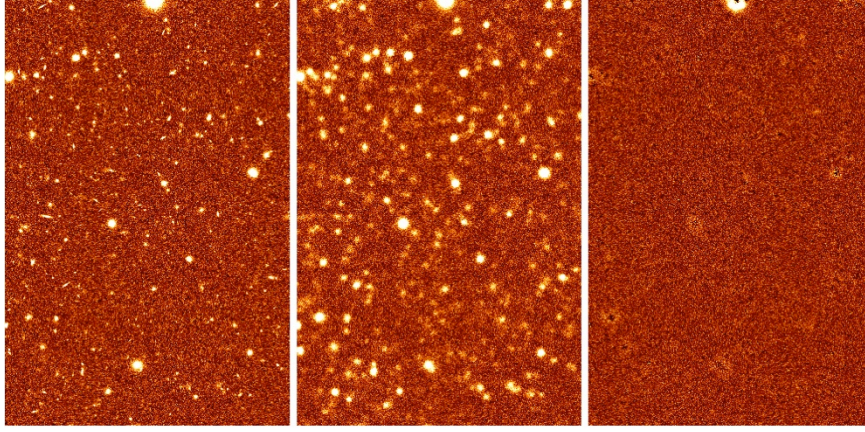
Figure 2 Example of the results of a standard T-PHOT run using real priors. Left to right: HRI (FWHM = 0.2"), LRI (FWHM = 1.66"), and residuals image for a simulated dataset. LRI and residual image are on the same intensity scale

In its pipeline, T-PHOT goes through well-defined stages, in each of which a single task is performed. It uses high-resolution priors to determine the positions and, when possible, the morphological information of the sources, and then uses this information to measure the fluxes of those sources in a lower resolution image (LRI). T-PHOT accepts three different kinds of priors: a catalog of sources from a high resolution image (HRI), and/or analytical 2-d models obtained e.g. using Galfit (Peng et al. 2010), and/or a catalogue of positions for unresolved point-sources (common practice or FIR and sub-mm band-passes). Note that the use of mixed priors is allowed, making it possible to e.g. remove foreground bright sources by modelling them as analytical 2-d profiles and simultaneously fitting them along with standard "real" cutouts (this is the procedure that has been followed to obtain K and IRAC AstroDeep catalogs on the Frontier Fields cluster images).

Figure 3 Accuracy of the flux determination in a simulation containing non-overlapping, PSF-shaped sources and with "perfect" detection. Relative measured flux difference (f_meas-f_true)/f_true is plotted versus logarithm of the input flux f_true, for a simulated image populated with PSF-shaped sources (FWHM=1.66"). Each dot corresponds to a single source, with different symbols and colors referring to various diagnostics

A normalized low-resolution model (*template*) of each object is created degrading its HRI cutout, or its model profile, using a PSF-matching kernel - or just the LRI PSF if unresolved priors are used. Then, to overcome the problem of the blending of sources, a Chi-square minimization problem is solved, fitting all the sources at once in a chosen region. The fit can be performed on the whole LRI, giving the most reliable results, or constructing "cells" around each source including all its potentially contaminating neighbours in the fit. The standard TFIT approach, consisting in dividing the LRI in a regular, arbitrary grid of cells, is still allowed but it is strongly discouraged, since it has proven to introduce non negligible errors

11

due to the potential contamination from the light coming from objects just outside the considered cell.

Nominal statistical uncertainties are assigned to each measurement from the covariance matrix of the problem. However, systematic errors may affect the measurements in some particular cases (e.g., saturated or blended priors and border sources): in such cases, a flag in the output catalogue highlights the problem. Also, strongly covariant objects can have badly measured fluxes: a "covariance index" offers a qualitative indication about this risk.

After the fitting stage, T-PHOT can perform a spatial cross-correlation between the LRI and the model image constructed from the templates, to obtain locally registered kernels which can be used for a second pass in order to minimize spatial inaccuracies. The main final products of the run are a catalogue including positions, measured fluxes, uncertainties and diagnostic flags, and residual image obtained subtracting the model image from the LRI, useful to check at a glance the overall goodness of the fit. Other sub-products and diagnostics are also produced, in particular those related to detailed statistical analysis of the residuals (e.g. variance of the residuals for any object, etc.). The overall quality of the fitting procedure can be visually appreciated on the images as shown in Figure 2, and analysed statistically with dedicated simulations, as shown in Figure 3. An example of the improvement between TFIT and T-PHOT is shown in Figure 4.

T-PHOT v2.0

The new options included in version 2.0 of T-PHOT are:
– background estimation, with two methods: a global subtraction of a constant fitted value on the whole fitted region, and a local fit of individual "background templates" corresponding to each source (this option was used to obtain the AstroDeep Frontier Fields IRAC and K-band catalogues, fitting a local background in the central regions of the clusters to properly remove residual light);
– local / individual kernel fitting: it is possible to associate a different convolution kernel to each source to optimize the fit, coping with local variations of the PSFs (this feature has been used to obtain the new AstroDeep IRAC GOODS-South catalogues);
– individual source registration (dance): after the fitting stage, the refinement of the spatial registration of the objects is performed on an individual basis rather than on arbitrary regions (this feature has been used to obtain the new AstroDeep IRAC GOODS-South catalogues, see Figure 5);
– flux prioring: the flux of selected or all sources can be constrained to a given desired value within a chosen uncertainty limit, e.g. to remain consistent with any expected prediction on the SED of the galaxy;
– statistics on the residuals: the output includes a new text file with diagnostic statistics for each source, based on the residual image produced after the fit;
– r.m.s. threshold to exclude sources from the fit: if the central pixel of a source has rms uncertainty exceeding a chosen value, the source will be excluded from the fitting procedure;
– model building with selected sources : it is possible to build a model image (and a residual image) including only a selection of sources from the priors list.

Summary

The performance of T-PHOT has been checked extensively on a wide set of simulated data and on real datasets. T-PHOT proves to be much faster than its predecessors, up to a factor of hundreds in the most favourable situation; it can deal with large datasets and give more accurate results with an appropriate choice of the input parameters. Input parameters are

easy to modify in a SExtractor-style file, and default options for the most common cases is also provided.
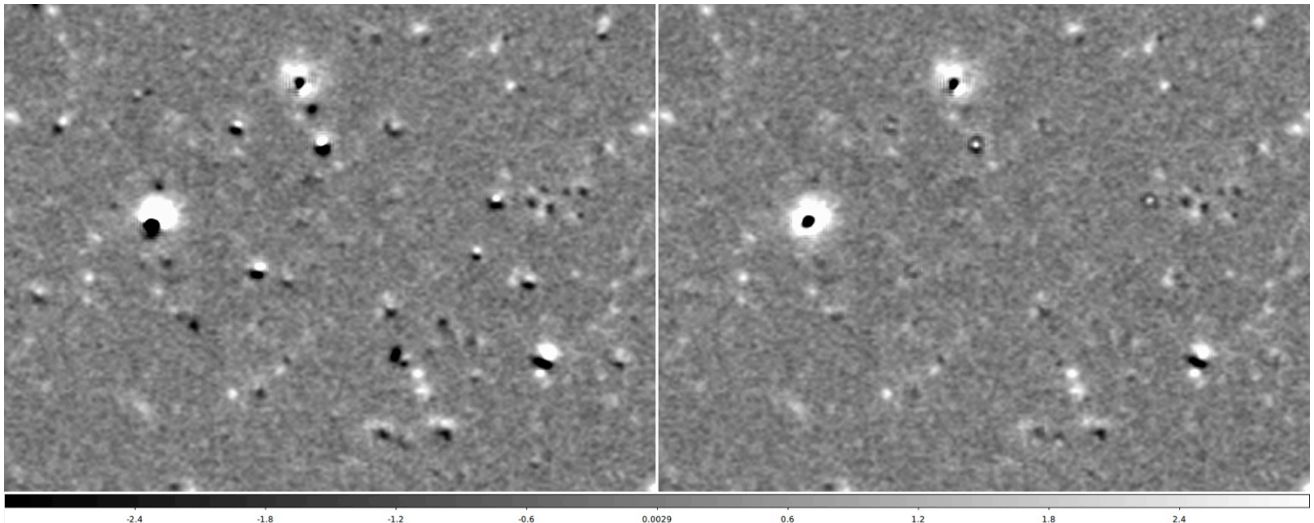

Figure 4 UDS I band TFIT vs. T-PHOT comparison. The panel on the left shows a small patch of the "official" CANDELS residual image obtained using TFIT. The residual image of the same regions is showed in the right panel, this time obtained using T-PHOT with "cells-on-objects" method and improved local kernel registration. Note the disappearance of many spurious black spots

T-PHOT is a public software and can be downloaded from the ASTRODEEP website; subscription to a mailing list is recommended. It comes as a tarball including documentation, installation scripts, and the source code. It consists of Python envelopes calling fast C and C++ codes, only needing a few external dependencies (some standard Python modules, the CFITSIO and FFTW3 libraries). It is easy to install and to use on UNIX and MAC-OS, with a user-friendly parameter file and a straightforward command line from a terminal.
T-PHOT has been presented and advertised by E. Merlin at the ADASS XXIV conference held in Calgary (CA) in October 2014, and at the conference "The spectral energy distribution of high redshift galaxies" held in Sexten (IT) in January 2015. It is indexed on the ASCL server (http://ascl.net/1609.001).

For all these reasons, T-PHOT is already being used by many researchers both within the ASTRODEEP consortium and not (>100 downloads of the code, most of which from other team). In our project it has been used to obtain photometric catalogues of K and IRAC Frontier Fields images and SCUBA CANDELS images, and to re-analyse IRAC CANDELS GOODS-S data. Other teams have used T-PHOT to work on the preparation of the JWST Guaranteed Time (C.Willmer, G.Rieke, Steward Observatory - University of Arizona *"TPHOT is a critical component in this endeavour, particularly given the range of PSF sizes for the JWST Near Infrared Camera"*), the South Pole Telescope and Dark Energy Survey (I.Chiu, ASIAA, Taiwan), Spitzer surveys (SMUVS using Spitzer-IRAC, Kaputi et al.), surveys of galaxy clusters (Lemaux et al., submitted; Shen et al., in prep; Tomczak et al., in prep), the preparation of satellite projects and low surface brightness measurements (D.Valls Gabaud, Paris Observatory), tests on undersampled images (e.g. W.Green).
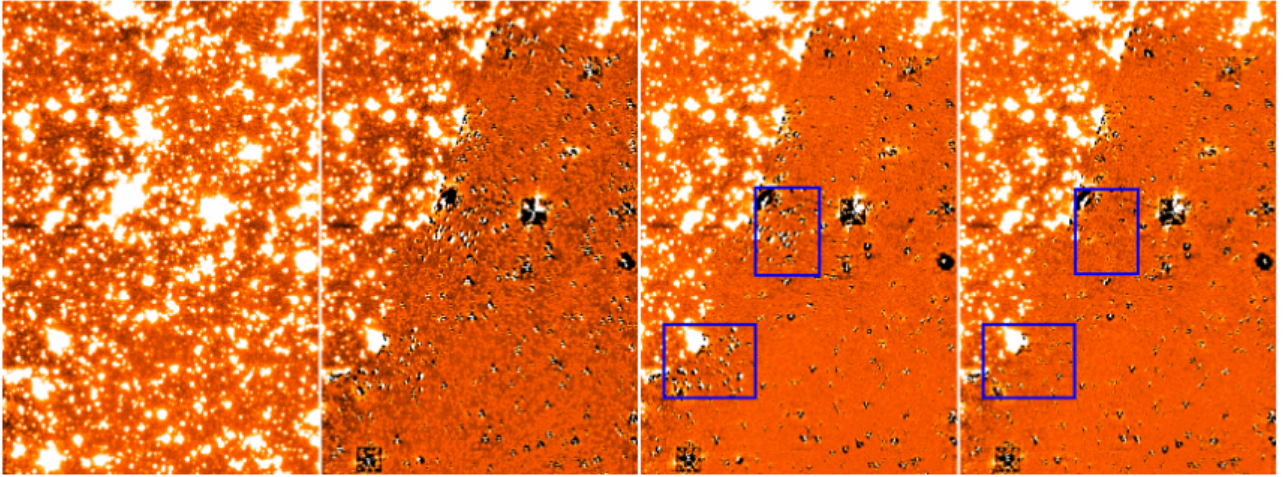
Figure 5 original LRI (a portion of GOODS-South observed with IRAC at 3.6 m, mosaic by R. McLure); residual after standard fitting using a single convolution kernel; residual after fitting using a different individual kernel for each source, tailored on the basis of the positional angles of the pointings used to build the mosaic (global background subtraction has also been applied); residual after including the individual kernel registration.

1.3 Far-IR deblending (SCUBA-2 + Herschel)

The effects of confusion are particularly severe in the case of far-IR images obtained by instruments like SCUBA-2 or PACS and SPIRE onboard the Herschel satellite. These data have a unique value in establishing the amount of dust-enshrouded star-formation rate and AGN activity, and are therefore a necessary complement of the deep optical-nearIR data. Unsurprisingly, they represent a major area of interest in our project and a field where our team has dedicated significant resources. Clearly, the major problem to overcome stems from the poor spatial resolution of the far-IR data, that have PSF typically in excess of 5 arcsec.

We have used our deconvolution methods to analyse in particular two data sets: the deepest 450 and 850μm imaging from SCUBA-2 CLS, covering 230 arcmin2 in the AEGIS, COSMOS and UDS fields, and the HERSCHEL PACS and SWIRE images from 100 to 500μm in GOODS-S and GOODS-N fields.

In both case we have used our de-convolution method (i.e. the TPHOT code) to correct the effects of de-blending, but we paid a particular attention in applying tailored recipes to identify the most suitable sets of priors to feed TPHOT.

Indeed, it is possible to show that using the full H-selected catalog obtained from the HST CANDELS images the surface density of sources become too large compared to the size of the PSF (and to the error on the centre of the emission, that especially for low S/N sources is significant). This would lead to a large chance of incorrect association between the near- and far-IR sources, or to instabilities in the fitting method. Trimming the input catalog in a sensible manner is hence absolutely needed to correctly identify the counterparts and measure the IR flux properly. For this reason, a significant effort has been dedicated to develop the optimal strategy (that depends on the specific data set analyzed) for prior selection.

1.3.1 The SCUBA-2 catalogue of selected CANDELS fields

In the case of the SCUBA-2 data we have selected priors from the 3D-HST CANDELS catalogue (which is effectively *H*-band selected) by first imposing limits in AB magnitude of *Ks* < 24 or IRAC [3.6]< 24, which are chosen to maximise the completeness of massive galaxies at *z* < 6. The combined *Ks* and [3.6] photometric se- lection ensures that the rest-frame selection wavelength is > 5000Å over the full redshift range, which is primarily sensitive to stellar mass, and relatively insensitive to variables such as star-formation history or dust obscuration. Finally, we applied a stellar mass cut of $\log M_\star > 9$, in order to avoid over-crowding of low-mass priors at low redshifts.

Provided with these prior lists and the sub-mm maps, T-PHOT runs an optimization routine to obtain flux measurements and uncertainties at all positions given in the input list. With such dense and deep prior lists, many of the output measurements have low signal-to-noise ratio and are not individually detected. The analysis in Bourne et al. (2017) therefore focuses on two samples. The first of these consists of the subset of galaxies detected with signal/noise (S/N) >3 at 450 μm, while the second is the full stellar-mass selected sample. To study the latter, the technique that we adopted in this work was to combine the measurements in pre-defined bins in order to obtain significant detections of average flux as a function of various galaxy properties (such as redshift, stellar mass, and UV luminosity). This method is analogous to image stacking, except that the explicit de-confusion means that it is not subject to the potential bias introduced by blended and clustered sources.

We detect 165 objects from our prior list, 130 of which have $M_\star > 10^{10} M_\odot$, and 66 % of which have spectroscopic redshifts (the rest have photometric redshifts). The T-PHOT residuals show that there are no remaining significant 450μm sources missing from our priors. Figure 6 shows the basic properties of the selected sources.
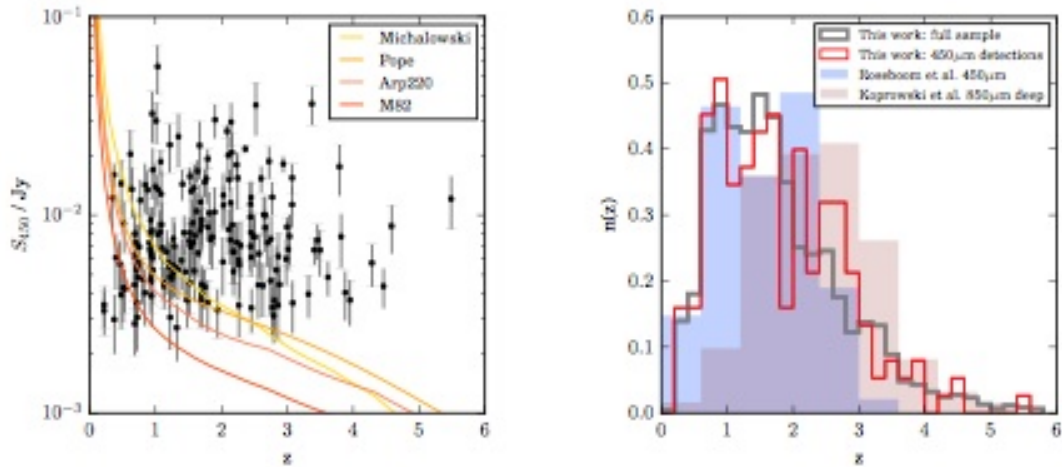

Figure 6 Properties of 450-μm sources detected at *S/N* ≥ 3 by the T-PHOT algorithm in the SCUBA-2 images. Left: 450-μm flux as a function of the redshift of the prior. Solid lines show the tracks of various SED templates scaled to a constant SFR$_{IR}$= 100 M$_\odot$yr^{-1} (L$_{IR}$ = 6.7 × 10^{11}L$_\odot$); templates are for SMGs from Michalowski et al. (2010) and Pope et al. (2008), and for Arp220 and M82 (Silva et al. 1998). Right: normalized redshift distribution of detections (red line) and of the full prior sample with $M_\star$ > 10^{9}M$_\odot$ (grey line). We also show, for comparison, the 450-μm-selected sources (with identifications) in the COSMOS and UDS deep fields of S2CLS (Roseboom et al. 2013); and the 850-μm-selected sources (with identifications) in the COSMOS deep field of S2CLS (Koprowski et al. 2016). Figure credit: Bourne et al. (2017).

1.3.2 Deblending the Herschel images

With PSFs that range from ~6" (at 70 μm) to ~35" (at 500 μm) the Herschel images, and especially those obtained with the SPIRE instrument, represented the toughest challenge to our de-blending techniques. We have developed an incremental methodology to produce

catalogs of sources in images where the dominant source of noise is the confusion noise, due to a large density of sources with large beam sizes.

This recipe, described in Tao et al. (2017, to be submitted), has been used to produce the Herschel catalogs that will be released by ASTRODEEP. It takes advantage of the following major improvements with respect to our previous catalog release:

1. the use of the covariance matrix to derive flux uncertainties in the Herschel images. This method was calibrated using realistic mock Herschel images.
2. determination of the optimum number of priors to be used for source extraction
3. improvement of the photometric accuracy on "non-clean" sources by "freezing" faint sources.
4. search for 24μm dropout sources, i.e. sources detected with Herschel SPIRE with no 24μm prior detection(using a blind source extraction on the residual maps).
5. optimization of the identification of distant optical counterparts to 500μm sources and validation of the technique using SCUBA2 450μm data (see also Shu et al. 2016).

In the case of the Herschel data the choice of priors to be adopted is even more crucial than at the other wavelengths. For source extraction in PACS maps, we have used all the 24 μm-detected sources as priors (H-band positions). We have first tested this approach also in all the SPIRE bands, using simulated images. We have found that while it works for the SPIRE-250μm, the outlier rate increases substantially at input fluxes below ∼10 mJy at 350 and 500 μm. This suggests that simply using all the 24 μm priors is not optimized for the source extraction at SPIRE bands. We conclude that a balance must be obtained between including as many priors as possible to include all the detectable sources, yet avoiding adding priors for sources that will not be detected and will unnecessarily increase the flux uncertainties of their neighbours. We have extensively studied this problem with simulated images. We find that the optimum number density of priors is around ∼ 0.7 beam/source, that for a field with size of the GOODS fields, i.e., 170 arcmin2 yields a number of priors of ∼1600, 840 and 400 priors at 250, 350 and 500 μm, respectively.
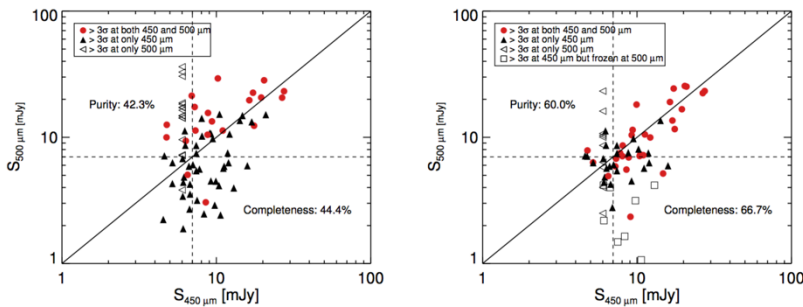

Figure 7 Comparison of extracted SPIRE 500μm and SCUBA2 450 μm flux densities from observed images in CANDELS-COSMOS. The left panel shows the comparison without freezing fainter sources while the right panel is the comparison after freezing fainter sources. 450 μm fluxes for sources only detected at 500 μm are assigned to 6 mJy, which is approximately the detection limit at 450 μm.

Another crucial improvement that we have introduced in the flux measurements of the brightest sources is to put constraints on the flux densities of the surrounding (potentially) fainter sources, and only fit the (potentially) bright one. This significantly reduces the risk of spreading the flux of the bright sources toward the faintest ones. The similarity of the infrared SEDs of star-forming galaxies or the exis- tence of the main sequence of star-forming galaxies allow us to predict their flux densities with a reasonable accuracy in the SPIRE bands based on information from shorter wavelengths (including MIPS and PACS bands), as was also shown in Leiton et al. (2015). For simplicity, we call this approach "freezing faint priors".

We have implemented this approach in a statistical robust manner, by modifying T-PHOT such that it allows also for "flux prioring" in the computation of the χ^2. This proceedure has been extensively tested with simulated images, and an example if shown in Figure 7. A

comparison with higher resolution 450 μm image from SCUBA 2 also suggests that these approaches yield better source extraction results than those used in previous work.

1.4 Prior photometry in deep X-ray images

We have applied the prior-based photometry for the first time also on the X ray data, with a specific and novel approach applied to the Chandra 4Ms data in the Chandra Deep Field South region.

For the first time we produced a catalogue with a Maximum likelihood PSF fitting technique based on prior HST galaxy detections. This has not been done using TPHOT (that cannot be straightforwardly applied to X-ray data, due to the different characteristics of the noise and background in the images), but a custom-modified version of a public software named *CMLDETECT.*

While several authors have used *cmldetect* for analysing Chandra surveys (see e.g., Puccetti et al. 2009; Krumpe et al. 2015), the major step forward here is the employment of WFC3-HST galaxies as priors to improve the efficiency on faint sources and to facilitate the identification process.

The tool allows us to analyse the data even with a strong gradient in the PSF, as typical of the Chandra images. We created at this purpose an ad–hoc PSF library by averaging over all the azimuthal angles the PSF templates in energy and off-axis angle bins.

Given an input list of source positions, the fit is performed simultaneously on all sources, with a maximum likelihood PSF fits to the events distribution on the detector are performed in all energy bands at the same time. The most important fit parameters are: the source location, source extent (beta model core radius), and source count rates. Sources with overlapping PSFs are fitted simultaneously.

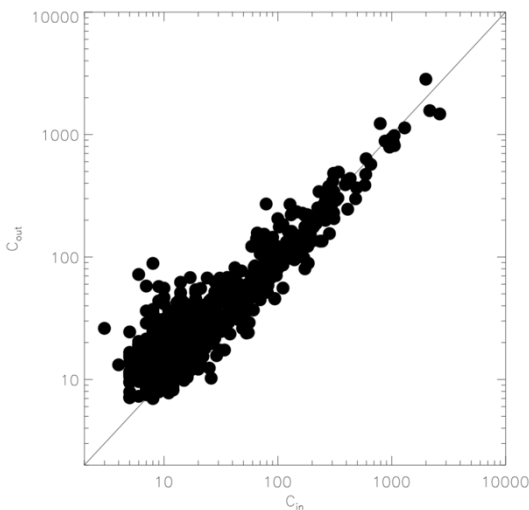To better scrutiny the reliability of our procedure and the origin of possible systematic effects we have designed a set of simulations and a comparison with other approaches to source detection. In order to reproduce in a realistic way our mock sample we created artificial X-ray fluxes of CANDELS galaxies from the estimated L8–1000μm by using ad-hoc scaling relations between LIR and LX (see below). Infrared luminosities (LIR, from 8 to 1000μm) are predicted for all galaxies in the catalogue starting from their observed photometric redshift, their stellar mass (Santini et al. 2014), their U V J rest-frame colours and their observed (or extrapolated from the SED) UV luminosity (1500˚A). A fraction of CANDELS galaxies could be AGNs, which are powerful X–ray sources. In order to include AGN X–ray emission in our sample, we divided the sample in $\Delta(z)=0.1$ redshift bins, and in every bin we assigned an AGN flux (SAGN) to a fraction of galaxies consistent with that expected by the Gilli et al. (2007) population synthesis model down to 10−20 erg s−1 cm−2. This method correctly reproduces the number founts and luminosity functions of AGNs.

Figure 8 Photometry efficiency test on the simulations. The input versus output source counts.

The simulated images performed in this way have been analysed as the real one and the recovered flux has been compared with the input flux. The agreement is quite good, as shown

in Figure 8. Further details can be found in Cappelluti et al 2016. The catalogue derived with this technique is described in the next section.

1.5 A pipeline for the removal of bright cluster sources in the Frontier Fields images

We describe here the procedure defined to extract high quality catalogues of the HST Frontier Fields. This project builds on the experience gathered with the analysis of the CANDELS fields but faces new challenges in the dense FF cluster fields. Its goals are:

a) To extend the detection of even the faintest galaxies down into to the very central regions of the clusters, in order to exploit the full power of lensing magnification, beating as much as possible the contamination from cluster members;
b) To obtain multi-wavelength photometric catalogues in all available bands from UV to NIR, i.e. adding available ground-based and Spitzer data to the ACS and WFC3 images.

The conceptual and technical challenges are substantial to extract maximal information from the ultra-deep images obtained in the field of a dense field dominated by bright, extended sources of the lensing cluster. To this end we have developed a complete procedure to perform a careful subtraction of bright cluster members, including also an estimate of the Intra Cluster Light (ICL) contribution. We then perform a properly optimized detection and photometry on the "cluster-subtracted" images using a combination of public codes (SExtractor and T-PHOT, the latter developed within our ASTRODEEP project).

The ASTRODEEP Frontier Fields catalogues include photometry in all available bands, photometric redshifts, stellar masses and other intrinsic parameters derived from the photometry. All catalogues are all made publicly available. The procedure is described in Merlin et al. 2016a (A&A 590, A30) where the photometric catalogues of the first two fields, Abell2744 and MACS-J0416 are also presented. Physical properties and derived quantities for the same fields are presented in the companion paper Castellano et al. 2016 (A&A 590, A31). A paper presenting the catalogues of the 3rd and 4th Frontier Fields, MACS-0717 and MACS-1149 is in preparation (DiCriscienzo et al.).
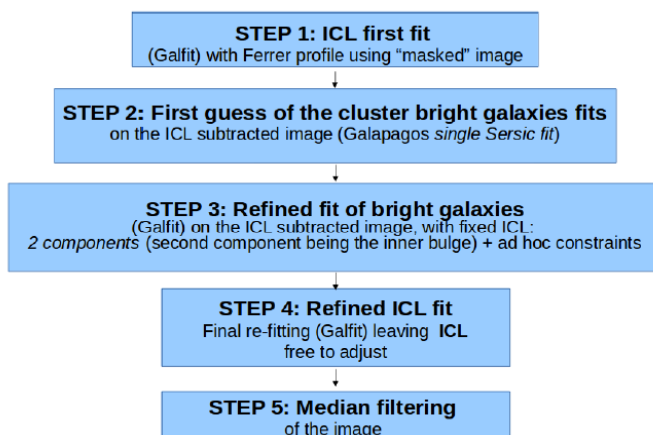

Figure 9 Schematic description of the ASTRODEEP procedure applied to "clean" the FF images, removing the light from the foreground bright sources.

Our analysis presented in the aforementioned works confirms that the FF initiative can effectively lead to the detection of high redshift galaxies with intrinsic magnitudes fainter than H~32, anticipating some of the scientific results expected from JWST.

Optical and NIR HST images

The complete procedure applied to the HST H-band image can be summarized in as follows (Figure 9):

- We use the HST-F160W (H) band to obtain a first model for the ICL. After masking-out all the objects in the field above a given threshold we use GALFIT to model the ICL emission of the cluster with one or two modified Ferrer profiles which follow the overall shape of the emission with the help of a few bending modes.

- We use the ICL subtracted image of the cluster in the H band to fit the bright cluster galaxies using GALAPAGOS (Barden 2009), a public code which wraps SExtractor and Galfit in a single working pipeline.
- We use GALFIT to fit both bright galaxies and ICL iteratively. This procedure produces the final models and residuals.
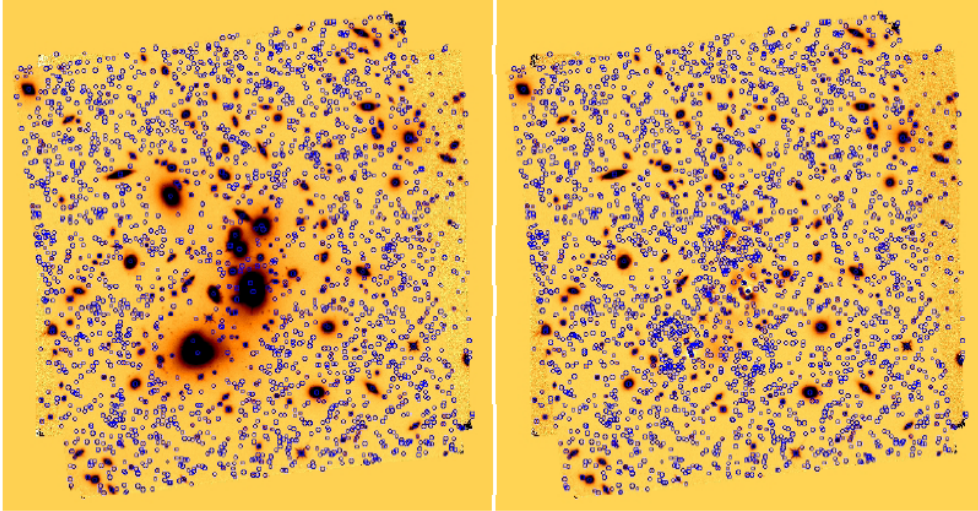

Figure 10 The blue circles show the objects detected in the A2744 cluster field. Left: detection on the original H160 image. Right: objects detected on the same image after removal of the fitted ICL and bright galaxies.

The residual image, i.e. the image of the field without the light of the bright galaxies and the ICL, is finally median filtered to mitigate the effects of remaining residual features after the GALFIT fitting.

The final output is a processed image of the cluster, which is used as a detection image.

Thus, all galaxies in the field have been detected in the F160W (H) band, after removing the ICL and bright cluster members as described above. As shown in Figure 10, the detection on these "subtracted" images can reveal many faint galaxies that are otherwise hidden in the bright halos of cluster members. The number of galaxies detected this way is much larger than using even an optimized detection on the original images. Between H=24 and H=28 we increase the number of detected sources (over the whole WFC3 field) by nearly a factor of two. The catalogue is extracted using an HOT+COLD approach, where the "COLD" mode corresponds to the standard CANDELS-HOT parameters set, and the "HOT" mode parameter choice is a modification of the standard CANDELS-HOT with a more aggressive choice for background subtraction. Performing the detection in a single band provides clear advantages in terms of a selection function that is more robust and easier to estimate. However, the resulting catalogue is not as deep as the one that could be obtained out of a stack of all the HST-IR images. For this reason we complement our catalogues with lists of objects detected in a weighted mean of the Y105+J125+JH140+H160 bands, while undetected in the H band only. This IR-stack is built from the processed *Galfit* residual images and used as a detection band in the same way as the H160 one.

Another application of the overall processing for the removal of bright galaxies is shown in Figure 11.
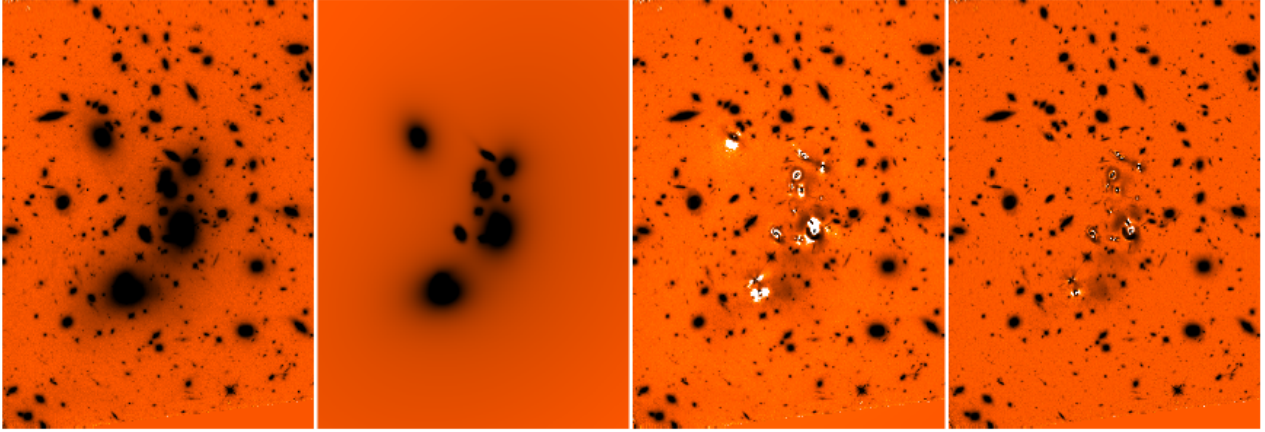
Figure 11 Application of the foreground light removal procedure to the A2744 H160 image. Left to right: original image, models of galaxies, and ICL after STEP 4, processed image after subtraction of the models, final processed image after median filtering (STEP 5). All images are in logarithmic greyscale with the same cuts

Detection completeness

We assess detection completeness as a function of the H-band magnitude by running simulations with synthetic sources. We first generate populations of point-like (i.e. PSF-shaped) and exponential profile sources, with total H -band magnitude in the range 26.5- 30. Disc-like sources are assigned an input half-light radius R_h randomly drawn from a uniform distribution between 0 and 1.0 arcsec. These fake galaxies are placed at random positions in our detection image, avoiding positions where real sources are observed on the basis of the original SExtractor segmentation map. To avoid an excessive and unphysical crowding of simulated objects, we include 200 objects of the same flux and morphology each time. We then perform the detection on the simulated image, using the same SExtractor parameters adopted in the real case. In the case of Abell2744 and MACS-J0416 we find that the 90% detection completeness for point sources is at H $\sim$27.7-27.8, and decreases to H $\sim$27.1-27.3 and H$\sim$26.6-26.7 for disk-like galaxies of R_h = 0.2 arcsec and R_h = 0.3 arcsec respectively, the lower values referring to M0416, as expected from its slightly shallower H -band depth. Comparable values are found in the MACS0717 and MACS1149 fields.
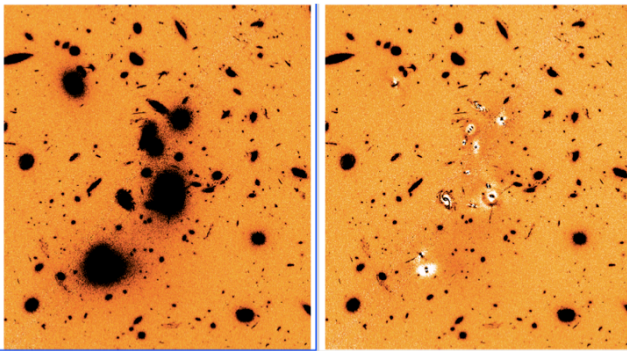
Other HST bands

To remove foreground galaxies on the other (bluer) HST bands, we have adopted the same procedure used for the H band in an iterative fashion. Starting from the J125 band, we have used the best-fit solution of the image immediately redder as a starting guess for the fit, and let the ICL and bright galaxies free to vary. The fit converged well in all bands from J125 to B435. The net computing time of the whole process is of the order of 4 days for each cluster field. One example are

Figure 12 original image and processed image in F606W

shown in Figure 12, where in each band the original images are on the left, and the processed ones are on the right.

K-band and Spitzer IRAC images

Including the K and especially the IRAC 3.6 and 4.5μm images is an even tougher challenge, because of the poorer image quality and increasingly large ICL contribution. To properly derive a robust photometry in these cases we adopt a prior-based approach. Briefly, the position and shape of the objects detected in the H band (including the ICL) are used as priors to fit the flux in the K and IRAC bands. This is accomplished using our photometric code T-PHOT, using the option that allows us to use real high-resolution cut-outs of sources as priors together with analytical models. We took advantage of this option to simultaneously fit the faint sources with the models of the bright cluster members. As a first step we estimate the PSF of the IR images and correct them by subtracting constant background components. The
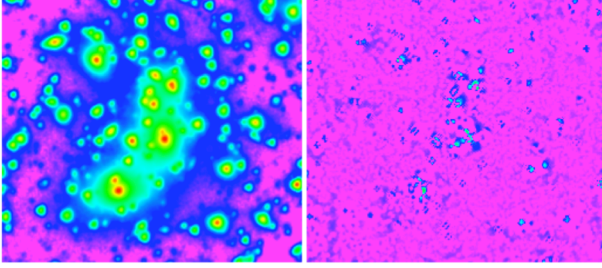

Figure 13 Deep K-band image from IRAC CH1 image from Spitzer (Capak et al.) and residual image after fitting the H-detected sources with the T-PHOT code.

r.m.s. maps are also consistently corrected after estimating the noise distribution from random empty positions in the images. The T-PHOT runs are then performed using source H-band cutouts and the Galfit models of bright galaxy as priors. We also performed a local background estimation independently for each source, building a large-scale background image that has been subtracted from the scientific image before measuring the fluxes. An example of the images after fitting all the
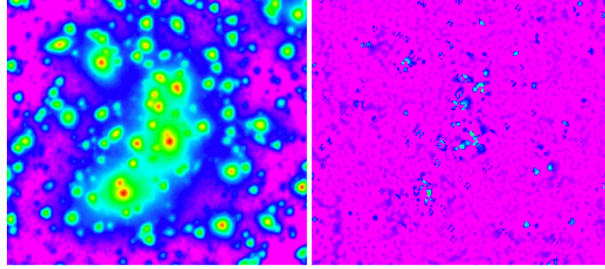
sources is shown in Figure 13

2. Catalogues and Data

2.1 The final GOODS-South catalogue

The GOODS-South field is the subject of an on-going massive observational effort by several groups worldwide. Over this field data have been collected from various telescopes and satellites, including HST, VLT, Spitzer, Subaru, Magellan, and others. Most of the data have been taken initially within the GOODS survey and later by the CANDELS survey. The latest public releases of a multiwavelength catalogue has been made by Guo et al (2013) and Skelton et al. (2014). The former was released by the CANDELS team itself and contains 17 bands, essentially ACS@HST, WFC3@HST and Spitzer. The latter has been later released by the 3d_HST team and adds also intermediate band images taken with Subaru. Both papers used prior-based, "forced" photometry of the sources detected in the H band image from HST (F160W).

With respect to such releases, we have performed a complete re-analysis of *all* the data existing on this field, motivated by the improved techniques that we have developed in the meantime (TPHOT, Merlin et al 2015) and by a wealth of new data that have been made available in the meantime.

We list here the major features of the new catalogue, highlighting the major improvements w.r.t the existing versions.

- **An improved detection for red sources.** The detection is still based on the same H band used by Guo et al (2013) and Skelton et al (2014), since no new data have been acquired in the meantime. In particular we have kept exactly the same target list of Guo et al (2013). However, we have added an independent search of objects in the K band, that is one of the ASTRODEEP deliverables (Fontana et al. 2014), that yielded an additional list of $\sim$ 200 sources. We have also included the objects detected in Spitzer (4.5μm) only described in Tao et al. (2016), again an ASTRODEEP result. The catalogue contains 34930 H-selected objects and 215 K-band selected objects, plus 6 objects that are detected in IRAC only.
- **New ACS data.** We have added the new ACS images presented by Illingworth et al. 2016. In some case they are deeper than the previous releases, most notably the F850LP and the F450W.
- **New Spitzer data.** The follow-up of the GOODS-S field with Spitzer along the recent years (the so-called warm cycles) have delivered an impressive amount of new data. We have included here the full dataset at 3.6 and 4.5μm, which improves by a factor 5-10x the depth of the images used in Guo et al 2013, as described in Labbe' et al 2016.
- **New ZFOURGE data** The ZFOURGE survey has delivered intermediate-band images in the near IR (across the J and H bands, the so-called J1, J2, J3, H1, H2) over a half of the GOODS-field, that help in the redshift determination of high redshift evolved objects.
- **A new set of spectroscopic redshifts.** Several hundredths of new redshift determinations have been obtain in the field, during a number of observational campaigns. They have been added to our catalogue.
- **The Herschel images.** The addition of the Herschel photometry over the GOODS-S field is a key improvement of our catalogue. It is better described in the following section.
- **A revised photometric technique.** All photometric measurements (except those on HST images) have been performed using our T-PHOT code, that significantly improves the reliability and accuracy of previous photometric measurements with TFIT and similar codes.

The final result of the analysis is a multiwavelength catalogue combing 43 bands in the optical- near IR, 6 from Herschel in the far—IR and the X-ray photometry on the 4Ms data of the Chandra deep field south. This is the most complete and up-to-date catalogue that has ever been obtained from an extragalactic deep survey so far. The completeness of the multiwavelength coverage and the sophisticated techniques adopted for the deblending make it one the best *manifesto* of the ASTRODEEP goals and results.

We have eventually obtained standard parameter from the SED of each object. The most straightforward and important is the photometric redshift, that uses the full power of the 43 bands to reach a better accuracy.
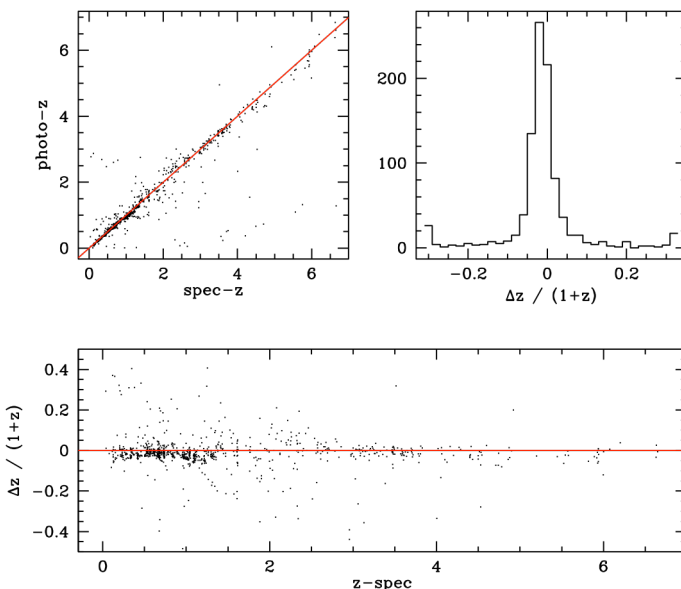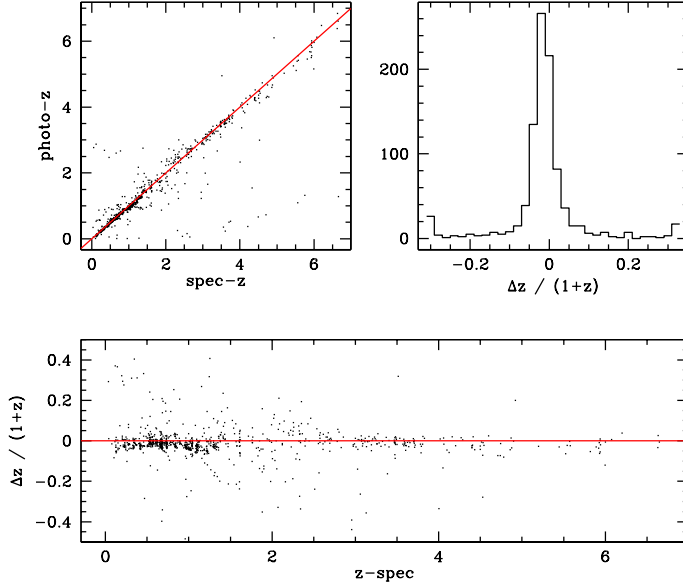
Figure 14 Relation between spectroscopic and photometric redshifts, on a asample of ~2000 objects with know redshift.

22

The   comparison   with   the   spectroscopic   sample

(Figure 14) shows a good correlation
between the spectroscopic and the photometric redshift. The average offset is Δz=0.009
scatter and the r.m.s. after sigma-clipping 0.036, better than what previously obtained in the
GOODS-S field with the same photometric redshift technique.

The average offset $\Delta z/(1+z)$ is 1×10^{-2}, the r.m.s. is 0.037 and the fraction of outlier is 8%, in
line or better than similar surveys at these faint limits (we note that our spectroscopic
sample is much fainter than those coming from wider surveys like COSMOS).
We have also obtained other rest-frame parameters (like stellar mass, star-formation rate,
rest-frame colours) that will all be released along with the photometric catalogue.

2.2 Herschel on GOODS-S and GOODS-N
Based on these newly reduced images, we have re-produced a new Herschel catalog in the
GOODS-South field (Wang et al., 2017, to be submitted), using the methodology described in
Section 3.3. Using the 24um-detected sources in GOODS-South (Magnelli+2013) as priors, we
performed PSF-based source extraction using FASTPHOT and TPHOT at the position of the
24um sources. Benefitted from the newly-reduced images, we obtained ~10% more sources
with S/N > 3 in the PACS bands. The most significant gain is in the SPIRE bands. With our
``freezing'' approach, we obtained 2 to 7 times more sources in the SPIRE bands (Table 4.1).
As an example, we show the redshift and flux distributions of the SPIRE 500μm-detected
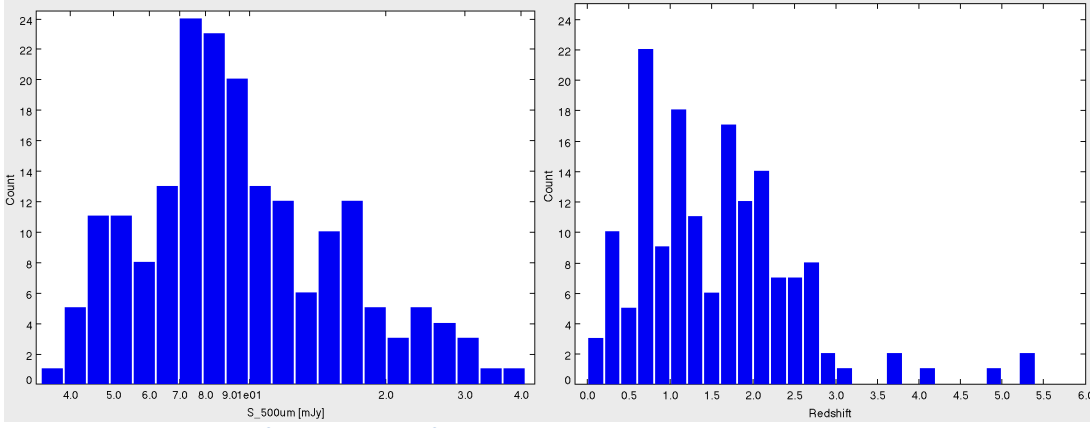sources in our new catalog (Figure 15).

Figure 15 Flux and redshift distributions of the 500um-detected sources in GOODS-South. The total number of sources detected in this case is 190, to be compared with the only 20 sources detected in previously released catalogues (Magnelli et al 2013, see Table 4.1).

The IR SED of the detected objects is typically well behaved. We show in Figure 16 some example.
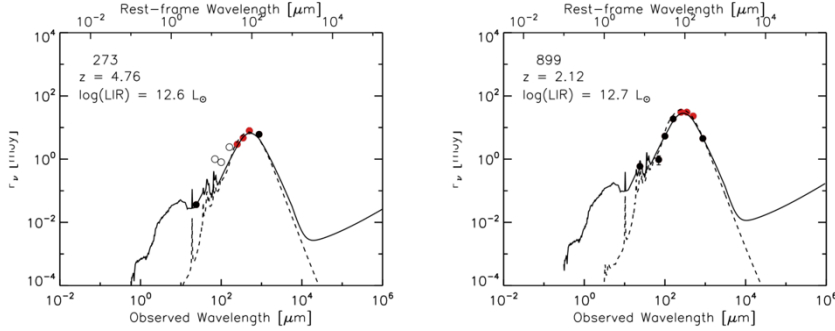

Figure 16 Example infrared SEDs of newly-detected 500um sources in our new Herschel catalog. Filled circles denote detections at certain bands while open circles denote non-detections. Photometry at the three SPIRE bands are denoted by red filled circles.

As a first application of this new and deep Herschel catalog, we explore the contribution of individual sources to the cosmic infrared background (CIRB) down to our detection limit. In particular, compared to previous measurements based on resolved galaxies (Oliver et al. 2010; Bethermin et al. 2012), our new approach allows us to resolve up to ~ 5 times more extragalactic back-ground light in individual galaxies at SPIRE 250, 350, and 500 μm.

	PACS 70μm	PACS 100μm	PACS 160μm	SPIRE 250μm	SPIRE 350μm	SPIRE500μm
GOODS-Herschel release (Elbaz+2011,Magnelli+2013)	507	1112	990	~400	~170	~20
This work	684	1367	1101	756	495	191

Table 1: Number of 3σ detections at HERSCHEL/PACS and SPIRE bands in the old GOODS-Herschel release and this work.

2.3 SCUBA-2 Survey

The SCUBA-2 Cosmology Legacy Survey (CLS) is the major branch of the James Clerk Maxwell Telescope (JCMT) Legacy Survey. It provides sub-mm imaging of several extragalactic deep fields using the world's largest single-dish sub-mm survey telescope equipped with the sensitive, wide-field SCUBA-2 camera. The CLS was designed to provide wide-field 850μm imaging over key degree-scale fields such as the equatorial COSMOS and UDS fields, and also

to exploit the very best (driest) observing conditions on Mauna Kea to deliver deep 450+850µm imaging within the CANDELS fields.

The relevant data set for ASTRODEEP is provided by the deep tier of SCUBA-2 imaging in the CANDELS fields (COSMOS, UDS, EGS and GOODS-N). The first three of these fields are particularly deep, reaching an rms noise level of ~1 mJy/beam at 450 µm (7.5 arcsec resolution), and ~0.2 mJy/beam at 850 µm (14 arcsec resolution), and have been exploited in an ASTRODEEP study published in MNRAS (Bourne et al. 2017).

As described in Bourne et al. (2017), the improved angular resolution of these maps (in comparison with Herschel maps) allows for effective de-confusion of much denser prior lists. This has made possible an analysis of obscured star formation in a complete sample of stellar-mass selected galaxies between redshifts of 0.5 and 6. The most significant contribution of this work is therefore to provide new, precise constraints on the cosmic star formation density at higher redshifts than has been possible in previous work with Herschel (which could probe ordinary galaxies only up to z~3).

In addition to the deep SCUBA-2 imaging at 450 µm, this work relies on the sophisticated de-confusion techniques of the ASTRODEEP software product T-PHOT (Merlin et al. 2015; 2016), which ensures accurate estimation of fluxes and errors for dense prior lists in confused maps, crucially taking full account of covariance between blended priors. We explored the effects of using different positional prior catalogues with varying source densities. This is an important choice, since a catalogue with too few positions risks not fully describing the full list of sources contributing flux to the map (hence flux measurements provided by T-PHOT may be biased, and reported uncertainties are too small). Conversely, too many positional priors lead to degenerate solutions and large uncertainties in the covariance matrix, which renders the flux measurements unusable. We found that the following selection provided an optimal set of priors for this study: we begin with the 3D-HST catalogue of the CANDELS fields (Skelton et al. 2014), which is primarily H-band (F160W) selected from the CANDELS HST WFC3 imaging (Grogin et al. 2011; Koekemoer et al. 2011). The advantage of the 3D-HST data set is that all galaxies have full optical/near-infrared (IR) photometric measurements, with derived stellar masses and photometric redshifts, in addition to HST grism and ground-based spectroscopic redshifts where available (Momcheva et al. 2015). From this parent catalogue, we select galaxies in the near-IR with K<24 or [3.6]<24. Selection at these wavelengths is primarily sensitive to stellar mass at z<6, hence reducing any bias against redder galaxies. We also remove any objects not flagged 'USE' in the 3D-HST catalogue, which ensures good photometry for stellar mass and photometric redshift measurements, and does not lead to significant incompleteness (see Bourne et al. 2017 for details). Finally, we applied a stellar mass cut of $\log M_\star > 9$, in order to avoid over-crowding of low-mass priors at low redshifts. This is justified on the basis that low-mass galaxies have negligible contribution to the cosmic IR background since they have both lower star formation rates (SFRs) and a lower fraction of obscured to unobscured SFR.

Provided with these prior lists and the sub-mm maps, T-PHOT runs an optimization routine to obtain flux measurements and uncertainties at all positions given in the input list. With such dense and deep prior lists, many of the output measurements have low signal-to-noise ratio and are not individually detected. The analysis in Bourne et al. (2017) therefore focuses on two samples. The first of these consists of the subset of galaxies detected with signal/noise (S/N) >3 at 450 µm, while the second is the full stellar-mass selected sample. To study the latter, the technique that we adopted in this work was to combine the measurements in pre-defined bins in order to obtain significant detections of average flux as a function of various galaxy properties (such as redshift, stellar mass, and UV luminosity). This method is analogous to image stacking, except that the explicit de-confusion means that it is not subject to the potential bias introduced by blended and clustered sources. The same techniques were

applied in lower-resolution maps from Herschel at 100, 160, 250 μm, and at 850 μm from SCUBA-2, in order to constrain the full far-IR spectral energy distributions (SEDs). While these measurements are naturally more noisy due to the increased confusion (with resolution ranging from ~9 to 18 arcsec in these maps), they provide useful constraints on the shape of the SED, allowing for reduced model dependence in the measurement of IR luminosity and obscured SFR

2.4 The 4Ms Chandra

The catalog on the X-ray data in the CDFS has been obtained applying our de-convolution technique to the 4Ms data obtained in the CDFS, overlapping with the CANDELS GOODS-South data.

The catalog uses as input the position of the 34930 sources detected in the CANDELS H-band images, and delivers the X-ray flux.

The sources detected at S/N>5 are 781 X-ray sources selected in the [0.5-7] keV energy range. Among this 757 can be considered secure detections with a secure counterpart in the CANDELS catalog. For 96% of the sources the actual Optical/NIR counterpart is the original prior source. The remaining 4% can be chance coincidences or off-axis sources. We tested our method through Monte Carlo ray-tracing simulations by using the state of art knowledge of the SFR-LX scaling relation for star-forming galaxies and recent AGN CXB population synthesis models. Our method significantly improves the efficiency in detecting faint X-ray sources in deep X-ray
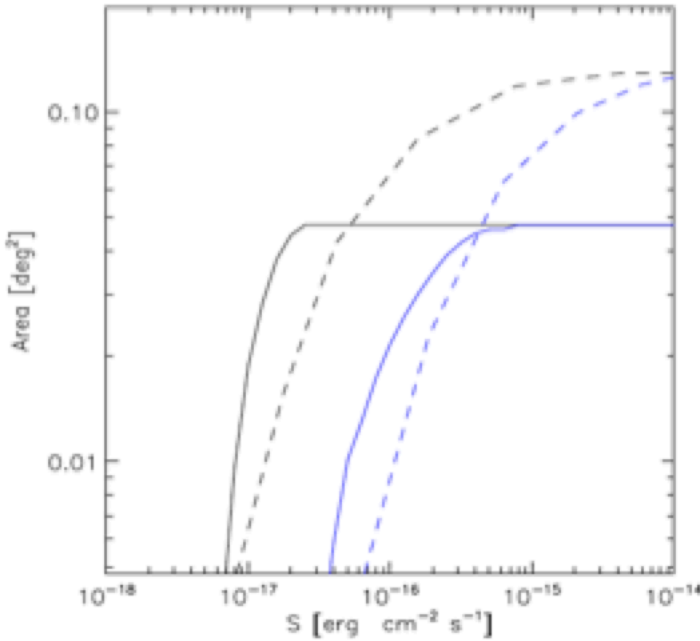

Figure 17 The sky area vs flux selection function plot for our sample in two sub-bands compared with that of Lehmer et al (2012). The black-continuous and the blue-continuous line represent the selection functions in the [0.5-2] keV and [2-10] keV energy bands, respectively. The black-dashed and the blue-dashed line represent the selection functions from Lehmer et al (2012) in the [0.5-2] keV and [2-10] keV energy bands, respectively

surveys (see Figure 17). Indeed, we discovered 423 new X-ray sources down to a flux of $\sim 0.8\text{-}0.9(2\text{-}3) \times 10^{-17}$ erg cm^{-2} s^{-1} in the [0.5-2] keV ([0.5-10] keV) energy band. These new optically faint sources have spectral signature typical of X-ray binary dominated star forming galaxies or faint highly absorbed AGN. By cross correlating our catalog with photo-z catalogs, we determined that we almost double the number of candidate z>4 AGN. This has extremely valuable applications in the framework of SMBH seeding since these candidate high-z AGN could witness the late stage of accretion onto massive seeds SMBH Volonteri et al. (2010). By cross-matching all the existing catalogs published in literature with ours we were able to estimate that the number of unique X-ray sources in the CANDELS GOODS-S area sums up to 961. A direct comparison with previous catalogues directly obtained from X-ray detections shows that we are able to recover objects at lower fluxes with high completeness, exploiting all the power of prior-based photometry. This paper and the relevant catalogue has been published in Cappelluti et al 2016.

2.5 The Frontier Fields catalogues

The Frontier Fields is an initiative originated by the Hubble Space Telescope Science Institute, that is delivering ultra-deep images with WFC3 and ACS over 6 different intermediate redshift clusters, and relevant "parallel" fields. The initiative is intended to anticipate the kind of science that will be made possible by JWST exploiting the amplification provided by the lensing clusters. This data set is clearly the frontier of the exploration of the high redshift Universe, filling the bridge between the Ultra Deep Field (that is marginally deeper but on a single pointing) and the future JWST surveys. Considering their great scientific interest, we have decided to dedicate a significant amount of time to the analysis of the Frontier Fields data, and to release high quality catalogue of these fields.

The major challenge that we had to overcome – that makes this data set conceptually challenging compared to previous - is the presence of bright and dense systems in the foreground – the bright galaxies of the lensing cluster. We have developed a dedicated and sophisticated technique - described in sect. 1.5 – to remove these objects and to obtain a clean catalogue. Due to the timing of the data acquisition, only 4 of the 6 Frontier Fields have been delivered in time for us to analyse the images. We describe here the properties of the 4 resulting catalogues, and of the relevant 4 "parallel" fields.
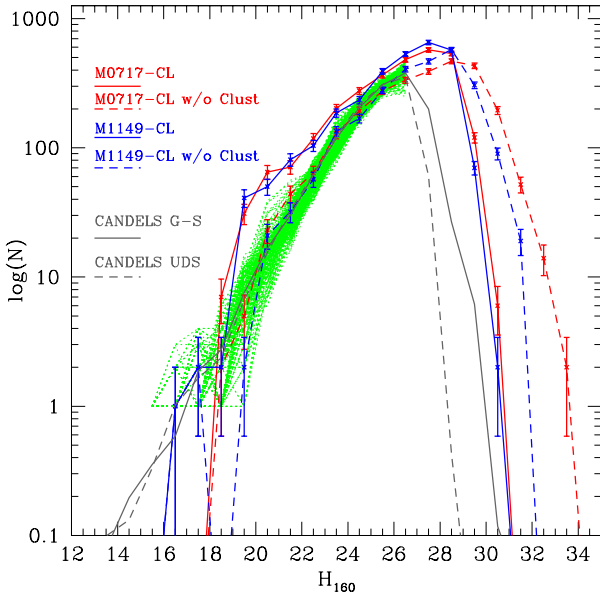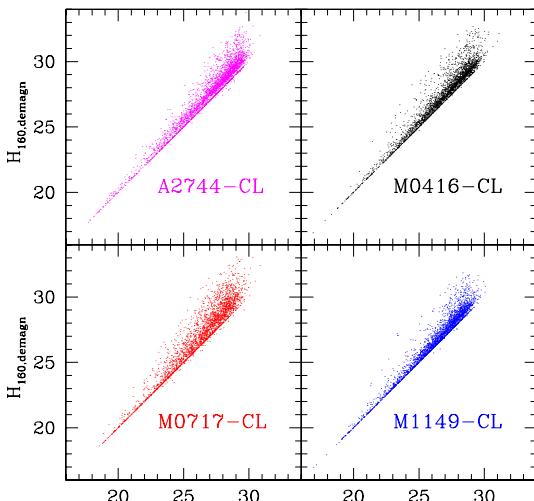

Figure 18 A comparison between the number counts in two Frontier Fields cluster w.r.t the CANDELS field. Solid coloured lines show the observed counts in the fields ; the dashed line show the same counts after de-magnification ; black like are the avera counts in the CANELS fields ; the green area show the number counts extracted in 100 random position of the CANDELS fields, over an area as large as the FF data, to show the variance of the counts over such a small area.

We have analysed and released the catalogues obtained in the first 4 Frontier Fields: Abell2744, MACS_J0416, MACS-J0717, MACS-J1149.

We perform the detection on the processed HST $H160$ image to obtain a pure H-selected sample, that is the primary catalog that we publish. We also add a sample of sources which are undetected in the $H160$ image but appear on a stacked infrared image. Photometry on the other HST bands is obtained using SExtractor, again on processed images after the procedure for foreground light removal. Photometry on the Hawk-I and IRAC bands is obtained using our PSF-matching de-confusion code T-PHOT. A similar procedure, but without the need for the foreground light removal, is adopted for the Parallel fields.

We have performed basic sanity checks on the reliability of our results. In particular, number counts in the Frontier Field are definitely consistent with previous results from CANDELS and HUDF (see Figure 18).

27

The power of the lensing analysis can be appreciated by de-magnified number counts that are shown in Figure 18, compared both with total number counts normalized to the FF area from the CANDELS GOODS-South surveys. At bright magnitudes the FF number counts are consistent with the CANDELS ones once magnification is taken into account and, in the case of the cluster pointings, sources with redshift compatible with being members of the A2744 and M0416 clusters (z*phot* within $\Delta z=0.1$ from the cluster redshift) are removed. At faint magnitudes the Frontier Fields cluster pointings allow us to detect sources up to ~3-4 magnitudes intrinsically fainter than objects in the deepest areas of the CANDELS fields. A similar result is shown in Figure 19, where we show the observed magnitude in the four clusters against the de-magnified one. It is clearly shown that galaxies that are intrinsically fainter than m~30 become accessible at m~27-28 in the FF data.

We have computed photometric redshifts for all galaxies in the survey, averaging over 5 different recipes (as computed by several partner of the ASTRODEEP team). They turn out to be quite reliable in the survey, as shown in Figure 20, with a r.m.s. error of ~0.045, that is comparable to CANDELS despite the lower number of filters available. The total redshift distribution of the 4 joined fields show a large number of galaxies at z>4, that are the primary targets of this survey.

Figure 19 Comparison between observed magnitude in the four clusters against the de-magnified ones.

The Frontier Fields catalogue have been published in Merlin et al. 2016 and Castellano et al. 2016 (A2744 and M0416), and in Di Criscienzo 2017 (A&A subm.)
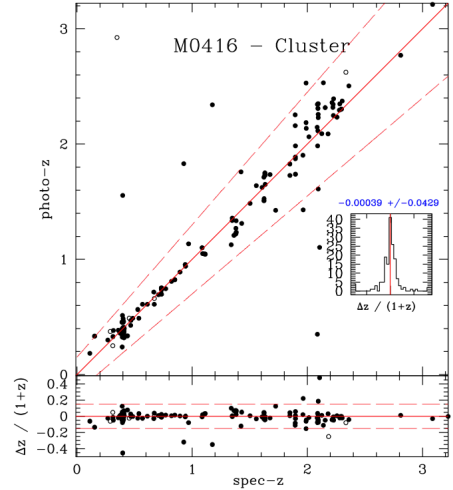
Figure 20 Comparison between photometric and spectroscopic redshifts in the MACS-0416 (right) cluster.

3 Simulated data

The role of simulations has been crucial in assessing the performances of the various recipes. As we have mentioned above, we have used simulated data in essentially all the steps of the development of our tools – and in the validation of the catalogues delivered. We will not repeat such tests here.
We have however decided to deliver and distribute some of these simulated data to the general public, to allow further tests to other interested scientists. This is particular important for the Euclid mission, and for the relevant workpackage in our project. We describe here the simulated data that we made public on the AstroDeep website.

3.1 Euclid imaging data

We decided to produce and release sky images (i.e. including noise) normalized to counts/s and background subtracted, and noise-subtracted images (from which the RMS map can be easily obtained), plus the PSFs.

We release three simulated datasets of images, created using the procedure described in D6.1.

- a deep field with area equal to the GOODS-South CANDELS field, including all the 19 passbands used in the CANDELS survey (all images have HST pixel-scale of 0.06"/pix);
- a wide field of area 1 sq.deg., i.e. comparable to a half the COSMOS wide survey area, including *HST H*160, ground-based and *Spitzer* passbands (all images have a pixel-scale of 0.15"/pix, as in Capak+2007);
- a simulated Euclid FOV (~0.7° per side) composed of 4 dithered single exposures, each in the 4 Euclid passbands (16 images in total, each of ~28000^2 pixels, at the VIS pixel-scale of 0.1"/pix).

3.2 Euclid spectroscopic data

We have performed simulations of the spectroscopic component of the Euclid mission. Detailed simulations are being performed within the Euclid science team with the aim of assessing the spectroscopic calibration of the photometric redshifts; however, we are the only group who have performed such extensive simulations tailored to assess the potential of the high-redshift legacy science case. The main scientific goal of this work was to estimate the number counts for emission lines galaxies, with a particular emphasis on Lyα, in the Euclid Deep Survey by combining our bespoke ASTRODEEP galaxy catalogue generation software EGG (Schreiber et. al. 2016), with the STScI grism simulation software aXeSIM. Below we describe the simulation methodology, before presenting some initial results which we have contributed to the Euclid Blue Grism Working Group. Finally, we describe the various simulation data products we intend to release to the community through the AstroDeep website.

3.2.1 Galaxy Catalogue Generation

To generate the galaxy catalogues for the simulation, we used the bespoke galaxy catalog generator EGG developed within ASTRODEEP (Schreiber et. al. 2016). The catalog was generated down to a limiting magnitude of m_{AB}=28 in the HST F105W band over 0.5 deg². Since EGG is based on a set of empirical prescriptions derived for galaxies at 0 < z < 4, we first had to check that these prescriptions provided an accurate description of the observed properties of galaxies at z > 6, since one of the main science goals of our work was to investigate the possibility of observing Lyα emission from these galaxies in the Euclid Deep survey. After some modification to the assumed M/L ratio, that now give a better fit to the high-z luminosity functions, and that have been eventually included in the current public version, we were able to generate galaxy catalogues which could successfully reproduce the UV luminosity function of galaxies at z > 6.
This ensures that EGG can accurately predict the number of galaxies at z > 6 as a function of absolute UV magnitude within current observational constraints.

Another important feature of the EGG catalogs is an implementation of the size evolution of galaxies, since all line profiles are convolved with galaxy morphology in slitless spectroscopy, and the resulting line S/N is dependent on galaxy size. As input to the spectroscopic simulations we were able to specify object size, in pixels, on the Euclid detector, based on the EGG prescriptions.

3.2.2 Realistic Emission Line Spectra

EGG outputs a best-guess stellar SED for each galaxy (based on the galaxies' star-formation rates, colours etc.), however it does not account for nebular emission lines. Therefore we have implemented a prescription for adding nebular emission lines to each spectrum using the following method (for the purposes of these simulations we focussed only on the Lyα, Hα, Hβ, [OIII] and [OII] lines):

- **Hα and Hβ:** the intrinsic Hα flux was calculated from the EGG star formation rate using the Kennicutt et. al. 2012 conversion. The intrinsic line flux was corrected for dust assuming $A_{H\alpha}$ =1, i.e. the median correction commonly adopted at these redshifts (e.g. Sobral et. al. 2013). The intrinsic Hβ flux calculated assuming an intrinsic Hα / Hβ ratio of 2.86 (Osterbrock et al. 2006), and we calculated $A_{H\beta}$ assuming $A_{H\alpha}$ =1 plus a Calzetti attenuation law (the standard assumption for high redshift galaxies).

- **Lyα:** the intrinsic Lyα flux was calculated from the intrinsic Ha flux assuming an intrinsic Lyα / Ha ratio of 8.1 (Osterbrock et al. 2006). In the case of Lyα we had to account for both internal dust attenuation and the Lyα escape fraction; in reality these parameters are degenerate however we found that by assuming $A_{Ly\alpha}$ = 1.2 (based on recent simulations of galaxies at z ~ 5, Cullen et. al. 2017) and an escape fraction of 30% based on the Hayes et. al. 2011 results, we were able to successfully reproduce the latest Lyα luminosity functions at z ~ 6.6 (Matthee et. al. 2015; see Figure 21). The faintest lines in Figure 21 correspond to the faintest objects (m~28) in the parent luminosity function. Since no Lyα luminosity functions exist at higher redshifts, we assumed these parameters held across the full redshift range over which Lyα is observable with the Euclid Blue Grism.
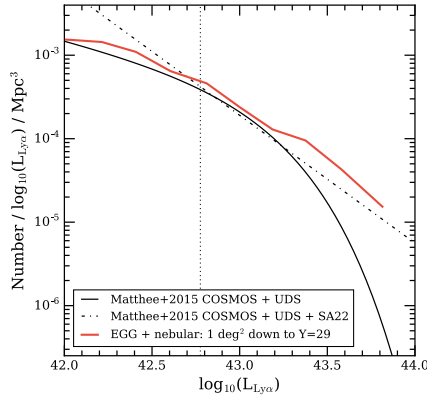

Figure 21 Lya luminosity functions derived from EGG, compared to observed lines (black lines)

- **[OIII] and [OII]:** The intrinsic [OIII] flux calculated from the intrinsic Hβ flux using the median [OIII] / Hβ ratio at z ~ 1 from Cullen et. al. 2016. The intrinsic [OII] flux was calculated from the median [OIII] / [OII] ratio at z ~ 2 from Nakajima et. al. 2014 (At higer redshift this ration can of course be different, but this assumption is fair in this context since [OII] and [OIII] can be detected by Euclid only at intermediate redshifts). Again both lines were corrected for dust assuming $A_{H\alpha}$ =1 plus a Calzetti attenuation law.

By combing these emission line prescriptions with EGG catalog data, we had a full description of the numbers counts and the UV - optical SEDs of galaxies over the redshift range accessible with both the blue and red Euclid grisms.

Finally the galaxy catalogues and spectra generated with EGG, plus our emission lines prescriptions, were used as input to the grism spectroscopic stimulater aXeSIM. The main inputs to aXeSIM are:

- Galaxy positions (x, y) and sizes in the image plane.

- Dispersion solutions for the Euclid grism spectra along with basic detector properties (readout noise, dark current, quantum efficiency, pixel size etc).
- Wavelength-dependant sensitivity files for the detector.
- A SED for each galaxy in the catalogue.
- Absolute magnitude in a given filter of each galaxy (to normalise the SEDs).
- Exposure time of observation

The official Euclid grism sensitivity and dispersion solution files are available for the red grism, however at present these are not available for the blue grism. For the blue grism we tuned the sensitivity such that and emission line with a flux of 6 x 10^{-17} erg/s/cm^2/A would be detected at 5σ for a 10 hour exposure.
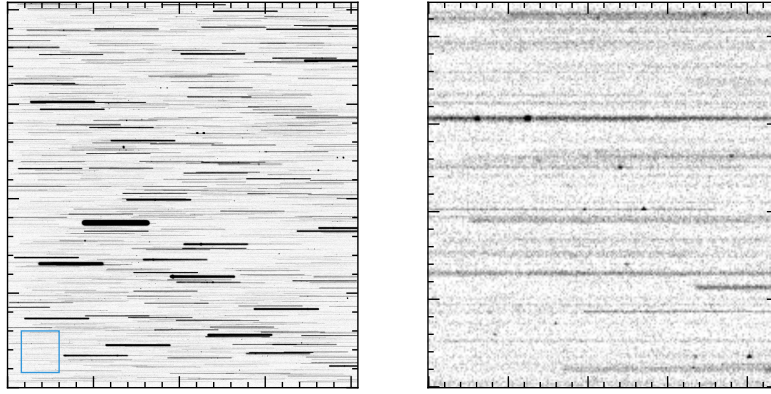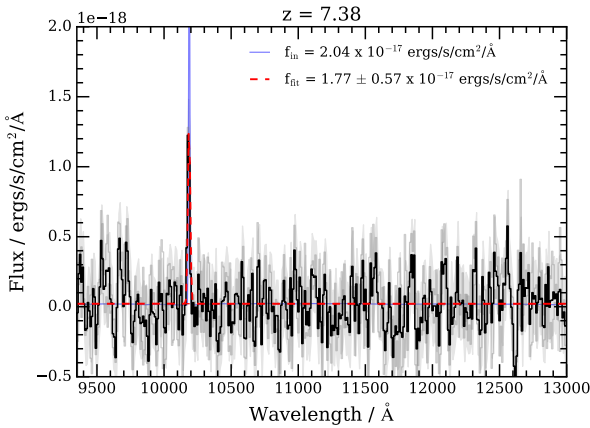

Figure 22 Euclid Grism Simulation with aXeSIM

Figure 23 Example of a z~7.4 simulated spectrum of a Lya emitter

aXeSIM outputs a grism image of the field-of-view (see Figure 22) along with the reduced 1D spectra of each object in the catalog.

An example spectrum of a Lyα emitter at z=7.38 is shown in Figure 23. By inspecting the output 1D spectra it is possible to quantify the completeness as a function of line flux for the Euclid Deep Survey, and with this completeness function, estimate the total number counts for each emission line.

3.2.3 Released Data Products

Finally, as a resource for the community, we have released on the AstroDeep website the data products of a full 0.5 deg^2 simulation in both the blue and red grisms. This includes the input galaxy catalogs from EGG along with the input spectra generated as described above, as well as the simulated grism images and output 1D spectra. These catalogs and images will allow

users to evaluate methods for optimally extracting spectra from Eucild images, and asses the impact of various problematic aspects of grism data reduction (i.e. spectra contamination, optimal extraction of faint lines etc.)

4. Science

ASTRODEEP has been much more than a coordinated effort of algorithm development and catalogue production. Its third pillar consists of the scientific discoveries that we have been able to achieve through the analysis of our data. A key feature of our approach is the recognition, based on extensive experience, that the development of new analysis methods, and data reduction must proceed in parallel with on-going attempts at scientific exploitation. This is essential for the efficient and timely recognition of the strengths and weaknesses of different techniques, and for the proper validation of science-ready data products for public release; the best-quality data are generally produced by those most motivated by their scientific exploitation.

In addition, competitive scientific activity is essential to guarantee a sound futurefor the post-doc that we hired and trained within ASTRODEEP, whom in many ways we consider our best legacy. For this reason, we have always left them free to dedicate a significant fraction of their time to research, using the ASTRODEEP data or other data, always in the context of the study of high-redshift galaxies.

Overall, 65 refereed papers have explicitly acknowledged the contribution of ASTRODEEP. These are listed chronologically below.

To relate these papers to the science areas listed in the Executive Summary, we have colour-coded them as follows. Papers focussed on algorithm development or catalogue production we have left in **black**. Papers based primarily on rest-frame UV data are coloured in **blue**, while papers studying dusty galaxies as revealed by far-IR/sub-mm data are coloured in **red**. Papers on the detailed astrophysics of galaxies at intermediate redshifts (i.e. studies of metallicity, morphology, or quiescent galaxies) are marked in **green**.

1.) A new method for classifying galaxy SEDs from multiwavelength photometry
Wild, V.; Almaini, O.; Cirasuolo, M.; Dunlop, J.S.; McLure, R.; Bowler, R.; Ferreira, J.; Bradshaw, E.; Chuter, R.; Hartley, W.
2014, MNRAS, 440, 1880

2.) The bright end of the galaxy luminosity function at z~7: before the onset of mass quenching?
Bowler, R.A.A.; Dunlop, J.S.; McLure, R.J.; Rogers, A.B.; McCracken, H.J.; Milvang-Jensen, B.; Furusawa, H.; Fynbo, J.P.U.; Taniguchi, Y.; Afonso, J.; Bremer ,M.N.; Le Fevre, O.
2014, MNRAS, 440, 2810

3.) The evolution of the dust temperatures of galaxies in the SFR-M∗ plane up to z ~ 2
Magnelli, B.; Lutz, D.; Saintonge, A.; Berta, S.; Santini, P.; Symeonidis, M.; Altieri, B.; Andreani, P.; Aussel, H.; Béthermin, M.; and 31 co-authors.
2014, A&A, 561, 86

4.) Constraints on the star-formation rate of z ~ 3 LBGs with measured metallicity in the CANDELS GOODS-South field
Castellano, M.; Sommariva, V.; Fontana, A.; Pentericci, L.; Santini, P.; Grazian, A.; Amorin, R.; Donley, J.L.; Dunlop, J.S.; Ferguson, H.C.; and 12 co-authors.

2014, A&A, 566, 19

5.) **The colour distribution of galaxies at redshift five**
Rogers, A.B.; McLure, R.J.; Dunlop, J.S.; Bowler, R.A.A.; Curtis-Lake, E.F.; Dayal, P.; Faber, S.M.; Ferguson, H.C.; Finkelstein, S.L.; Grogin, N.A.; and 4 co-authors.
2014, MNRAS, 440, 3714

6.) **A Submillimeter Galaxy Illuminating its Circumgalactic Medium: Lyman alpha Scattering in a Cold, Clumpy Outflow**
Geach, J.E.; Bower, R.G.; Alexander, D.M.; Blain, A.W.; Bremer, M.N.; Chapin, E.L.; Chapman, S.C.; Clements, D.L.; Coppin, K.E.K.; Dunlop, J.S.; and 10 co-authors.
2014, ApJ, 793, 22
7.) The Hawk-I UDS and GOODS Survey (HUGS): Survey design and deep K-band number counts
Fontana, A.; Dunlop, J.S.; Paris, D.; Targett, T.A.; Boutsia, K.; Castellano, M.; Galametz, A.; Grazian, A.; McLure, R.J.; Merlin, E.; and 33 co-authors.
2014, A&A, 570, 11

8.) **The Progenitors of Local Ultra-massive Galaxies Across Cosmic Time: From Dusty Star-bursting to Quiescent Stellar Populations**
Marchesini, D.; Muzzin, A.; Stefanon, M.; Franx, M.; Brammer, G.G.; Marsan, C.Z.; Vulcani, B.; Fynbo, J. P. U.; Milvang-Jensen, B.; Dunlop, J.S.; Buitrago, F.
2014, ApJ, 794, 65

9.) **A reassessment of the redshift distribution and physical properties of luminous (sub-) millimetre galaxies**
Koprowski, M.P.; Dunlop, J.S.; Michalowski, M.J.; Cirasuolo, M.; Bowler, R.A.A.
2014, MNRAS, 444, 117

10.) **The bulge-disc decomposed evolution of massive galaxies at 1 < z < 3 in CANDELS**
Bruce, V.A.; Dunlop, J.S.; McLure, R. J.; Cirasuolo, M.; Buitrago, F.; Bowler, R.A.A.; Targett, T.A.; Bell, E.F.; McIntosh, D.H.; Dekel, A.; and 8 co-authors
2014, MNRAS, 444, 1001

11.) **The decomposed bulge and disc size-mass relations of massive galaxies at 1 < z < 3 in CANDELS**
Bruce, V.A.; Dunlop, J.S.; McLure, R.J.; Cirasuolo, M.; Buitrago, F.; Bowler, R.A.A.; Targett, T.A.; Bell, E.F.; McIntosh, D.H.; Dekel, A.; and 8 co-authors.
2014, MNRAS, 444, 1660

12.) **Colour matters: the effects of lensing on the positional offsets between optical and submillimetre galaxies in Herschel-ATLAS**
Bourne, N.; Maddox, S.J.; Dunne, L.; Dye, S.; Eales, S.; Hoyos, C.; González-Nuevo, J.; Smith, D.J.B.; Valiante, E.; de Zotti, G.; and 2 co-authors
2014, MNRAS, 444, 1884

13.) **Determining the stellar masses of submillimetre galaxies: the critical importance of star formation histories**
Michalowski, M.J.; Hayward, C.C.; Dunlop, J.S.; Bruce, V.A.; Cirasuolo, M.; Cullen, F.; Hernquist, L.
2014, A&A, 571, 75

14.) A mass threshold in the number density of passive galaxies at z ~ 2

Sommariva, V.; Fontana, A.; Lamastra, A.; Santini, P.; Dunlop, J.S.; Nonino, M.; Castellano, M.; Ferguson, H.; McLure, R.J.; Galametz, A.; and 10 co-authors

2014, A&A, 571, 99.

15.) Accounting for Cosmic Variance in Studies of Gravitationally Lensed High-redshift Galaxies in the Hubble Frontier Field Clusters

Robertson, B.E.; Ellis, R.S.; Dunlop, J.S.; McLure, R.J.; Stark, D.P.; McLeod, D.

2014, ApJ, 796, L27

16.) The temperature dependence of the far-IR-radio correlation in the Herschel-ATLAS

Smith, D.J.B.; Jarvis, M.J.; Hardcastle, M.J.; Vaccari, M.; Bourne, N.; Dunne, L.; Ibar, E.; Maddox, N.; Prescott, M.; Vlahakis, C.; and 5 co-authors.

2014, MNRAS, 445, 2232

17.) Essential physics of early galaxy formation

Dayal, P.; Ferrara, A.; Dunlop, J.S.; Pacucci, F.

2014, MNRAS, 445, 2545

18.) The SCUBA-2 Cosmology Legacy Survey: ALMA Resolves the Rest-frame Far-infrared Emission of Sub-millimeter Galaxies

Simpson, J.M.; Smail, I.; Swinbank, A.M.; Almaini, O.; Blain, A.W.; Bremer, M.N.; Chapman, S.C.; Chen, C.-C.; Conselice, C.; Coppin, K.E.K.; and 19 co-authors.

2015, ApJ, 799, 81

19.) The SCUBA-2 Cosmology Legacy Survey: the submillimetre properties of Lyman-break galaxies at z = 3-5

Coppin, K.E.K.; Geach, J.E.; Almaini, O.; Arumugam, V.; Dunlop, J.S.; Hartley, W.G.; Ivison, R.J.; Simpson, C.J.; Smith, D.J.B.; Swinbank, A.M.; and 21 co-authors.

2015, MNRAS, 446, 1293

20.) Stellar masses from the CANDELS survey: the GOODS-South and UDS fields

Santini, P.; Ferguson, H.C.; Fontana, A.; Mobasher, B.; Barro, G.; Castellano, M.; Finkelstein, S.L.; Grazian, A.; Hsu, L.T.; Lee, B.; and 28 co-authors.

2015, ApJ, 801, 97

21.) Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope

Robertson, B.E.; Ellis, R.S.; Furlanetto, S.R.; Dunlop, J.S.

2015, ApJ, 802, L19

22.) The SCUBA-2 Cosmology Legacy Survey: ALMA Resolves the Bright-end of the Sub-millimeter Number Counts

Simpson, J. M.; Smail, I.; Swinbank, A.M.; Chapman, S.C.; Geach, J.E.; Ivison, R.J.; Thomson, A.P.; Aretxaga, I.; Blain, A.W.; Cowley, W.I.; and 13 co-authors

2015, ApJ, 807, 128

23.) GOODS-HERSCHEL: star formation, dust attenuation and the FIR-radio correlation on the Main Sequence of star-forming galaxies up to z~4
Pannella, M.; Elbaz, D.; Daddi, E.; Dickinson, M.E.; Hwang, H.S.; Schreiber, C.; Strazzullo, V.; Aussel, H.; Bethermin, M.; Buat, V.; and 18 co-authors
2015, ApJ, 807, 141

24.) The Herschel view of the dominant mode of galaxy growth from z=4 to the present day
Schreiber, C.; Pannella, M.; Elbaz, D.; Béthermin, M.; Inami, H.; Dickinson, M.E.; Magnelli, B.; Wang, T.; Aussel, H.; Daddi, E.; and 13 co-authors.
2015, A&A, 575, 74

25.) The SCUBA-2 Cosmology Legacy Survey: the submillimetre properties of Lyman-break galaxies at z = 3-5
Coppin, K.E.K.; Geach, J.E.; Almaini, O.; Arumugam, V.; Dunlop, J.S.; Hartley, W.G.; Ivison, R.J.; Simpson, C.J.; Smith, D.J.B.; Swinbank, A.M.; and 21 co-authors
2015, MNRAS, 446, 1293

26.) GOODS-Herschel: identification of the individual galaxies responsible for the 80-290 μm cosmic infrared background
Leiton, R.; Elbaz, D.; Okumura, K.; Hwang, H S.; Magdis, G.; Magnelli, B.; Valtchanov, I.; Dickinson, M.; Béthermin, M.; Schreiber, C.; and 7 co-authors
2015, A&A, 579, 93

27.) Probing the galaxy-halo connection in UltraVISTA to z ~ 2
McCracken, H. J.; Wolk, M.; Colombi, S.; Kilbinger, M.; Ilbert, O.; Peirani, S.; Coupon, J.; Dunlop, J.; Milvang-Jensen, B.; Caputi, K.; and 3 co-authors.
2015, MNRAS, 449, 901

28.) New redshift z ~ 9 galaxies in the Hubble Frontier Fields: Implications for early evolution of the UV luminosity density
McLeod, D.J.; McLure, R.J.; Dunlop, J.S.; Robertson, B.E.; Ellis, R.S.; Targett, T.T.
2015, MNRAS, 450, 3032

29.) The galaxy luminosity function at z ~ 6 and evidence for rapid evolution in the bright end from z ~ 7 to 5
Bowler, R.A.A.; Dunlop, J.S.; McLure, R.J.; McCracken, H.J.; Milvang-Jensen, B.; Furusawa, H.; Taniguchi, Y.; Le Fevre, O.; Fynbo, J.P.U.; Jarvis, M.J.; Haussler, B.
2015, MNRAS, 452, 1817

30.) Cold dust emission from X-ray AGN in the SCUBA-2 Cosmology Legacy Survey: dependence on luminosity, obscuration and AGN activity
Banerji, M.; McMahon, R.G.; Willott, C.J.; Geach, J.E.; Harrison, C.M.; Alaghband-Zadeh, S.; Alexander, D.M.; Bourne, N.; Coppin, K.E.K.; Dunlop, J.S.; and 8 co-authors
2015, MNRAS, 454, 419

31.) Evolution of star formation in the UKIDSS Ultra Deep Survey Field - II. Star formation as a function of stellar mass between z = 1.46 and 0.63
Drake, A.B.; Simpson, C.; Baldry, I.K.; James, P.A.; Collins, C.A.; Ouchi, M.; Yuma, S.; Dunlop, J.S.; Smith, D.J.B.

2015, MNRAS, 454, 2015

32.) **The galaxy stellar mass function at 3.5<z<7.5 in the CANDELS/UDS, GOODS-South, and HUDF fields**
Grazian, A.; Fontana, A.; Santini, P.; Dunlop, J.S.; Ferguson, H C.; Castellano, M.; Amorin, R.; Ashby, M.L.N.; Barro, G.; Behroozi, P.; and 31 co-authors.
2015, A&A, 575, 96

33.) **T-PHOT: A new code for PSF-matched, prior-based, multiwavelength extragalactic deconfusion photometry**
Merlin, E.; Fontana, A.; Ferguson, H.C.; Dunlop, J.S.; Elbaz, D.; Bourne, N.; Bruce, V.A.; Buitrago, F.; Castellano, M.; Schreiber, C.; and 12 co-authors
2015, A&A, 582, 15

34.) **The imprint of rapid star formation quenching on the spectral energy distributions of galaxies**
Ciesla, L.; Boselli, A.; Elbaz, D.; Boissier, S.; Buat, V.; Charmandaris, V.; Schreiber, C.; Béthermin, M.; Baes, M.; Boquien, M.; and 5 c-oauthors
2016, A&A, 585, 43

35.) **The Lyman continuum escape fraction of galaxies at z = 3.3 in the VUDS-LBC/COSMOS field**
Grazian, A.; Giallongo, E.; Gerbasi, R.; Fiore, F.; Fontana, A.; Le Fèvre, O.; Pentericci, L.; Vanzella, E.; Zamorani, G.; Cassata, P.; and 45 co-authors
2016, A&A, 585, 48

36.) **Infrared color selection of massive galaxies at z > 3**
Wang, T.; Elbaz, D.; Schreiber, C.; Pannella, M.; Shu, X.; Willner, S.P.; Ashby, M.L.N.; Huang, J.-S.; Fontana, A.; Dekel, A.; and 22 co-authors
2016, ApJ, 816, 84

37.) **Identification of z >2 Herschel 500 micron sources using color-deconfusion**
Shu, X.W.; Elbaz, D.; Bourne, N.; Schreiber, C.; Wang, T.; Dunlop, J.S.; Fontana, A.; Leiton, R.; Pannella, M.; Okumura, K.; and 12 co-authors
2016, ApJS, 222, 4

38.) **The galaxy UV luminosity function at z ~ 2 - 4; new results on faint-end slope and the evolution of luminosity density**
Parsa, S.; Dunlop, J.S.; McLure, R.J.; Mortlock, A.
2016, MNRAS, 456, 3194

39.) **Non-parametric analysis of the rest-frame UV sizes and morphological disturbance amongst L$_*$ galaxies at 4 < z < 8**
Curtis-Lake, E.; McLure, R.J.; Dunlop, J.S.; Rogers, A.B.; Targett, T.; Dekel, A.; Ellis, R.S.; Faber, S.M.; Ferguson, H.C.; Grogin, N.A.; and 5 co-authors
2016, MNRAS, 457, 440

40.) **The ASTRODEEP Frontier Fields catalogues. I. Multiwavelength photometry of Abell-2744 and MACS-J0416**

Merlin, E.; Amorín, R.; Castellano, M.; Fontana, A.; Buitrago, F.; Dunlop, J.S.; Elbaz, D.; Boucaud, A.; Bourne, N.; Boutsia, K.; and 21 co-authors
2016, A&A, 590, 30

41.) The ASTRODEEP Frontier Fields catalogues. II. Photometric redshifts and rest frame properties in Abell-2744 and MACS-J0416
Castellano, M.; Amorín, R.; Merlin, E.; Fontana, A.; McLure, R.J.; Mármol-Queraltó, E.; Mortlock, A.; Parsa, S.; Dunlop, J.S.; Elbaz, D.; and 28 co-authors
2016, A&A, 590, 31

42.) Constraints on Photoionization Feedback from Number Counts of Ultra-faint High-redshift Galaxies in the Frontier Fields
Castellano, M.; Yue, B.; Ferrara, A.; Merlin, E.; Fontana, A.; Amorín, R.; Grazian, A.; Mármol-Queralto, E.; Michałowski, M.J.; Mortlock, A.; and 4 co-authors
2016, ApJ, 823, L40

43.) Chandra counterparts of CANDELS GOODS-S sources
Cappelluti, N.; Comastri, A.; Fontana, A.; Zamorani, G.; Amorin, R.; Castellano, M.; Merlin, E.; Santini, P.; Elbaz, D.; Schreiber, C.; and 11 co-authors
2016, ApJ, 823, 95

44.) The SCUBA-2 Cosmology Legacy Survey: galaxies in the deep 850-micron survey, and the star-forming `main sequence'
Koprowski, M.; Dunlop, J.S.; Michalowski, M.J.; Roseboom, I.; Geach, J.E.; Cirasuolo, M.; Aretxaga, I.; Bowler, R.A.A.; Banerji, M.; Bourne, N.; and 7 co-authors
2016, MNRAS, 458, 4321

45.) The z = 9-10 galaxy population in the Hubble Frontier Fields and CLASH surveys: the z = 9 luminosity function and further evidence for a smooth decline in ultraviolet luminosity density at z≥ 8
McLeod, D.J.; McLure, R.J.; Dunlop, J.S.
2016, MNRAS, 459, 3812

46.) Changing physical conditions in star-forming galaxies between redshifts 0<z<4: [O III]/H β evolution
Cullen, F.; Cirasuolo, M.; Kewley, L.J.; McLure, R.J.; Dunlop, J.S.; Bowler, R.A.A.
2016, MNRAS, 460, 3002

47.) The evolution of the equivalent width of the Ha emission line and specific star-formation rate in star-forming galaxies at 1<z<5
Marmol-Queralto, E.; McLure, R.J.; Cullen, F.; Dunlop, J.S.; Fontana, A.; McLeod, D.J.
2016, MNRAS, 460, 3587

48.) Planetary nebulae in the Small Magellanic Cloud
Ventura, P.; Stanghellini, L.; Di Criscienzo, M.; García-Hernández, D.A.; Dell'Agli, F.
2016, MNRAS, 460, 3940

49.) Evolved stars in the Local Group galaxies - I. AGB evolution and dust production in IC 1613

Dell'Agli, F.; Di Criscienzo, M.; Boyer, M.L.; García-Hernández, D.A.

2016, MNRAS, 460, 4230

50.) T-PHOT version 2.0: Improved algorithms for background subtraction, local convolution, kernel registration, and new options

Merlin, E.; Bourne, N.; Castellano, M.; Ferguson, H.C.; Wang, T.; Derriere, S.; Dunlop, J.S.; Elbaz, D.; Fontana, A.

2016, A&A, 595, 97

51.) The evolution of post-starburst galaxies from z=2 to 0.5

Wild, V.; Almaini, O.; Dunlop, J.; Simpson, C.; Rowlands, K.; Bowler, R.; Maltby, D.; McLure, R.

2016, MNRAS, 463, 832

52.) VLA and ALMA Imaging of Intense Galaxy-wide Star Formation in z~2 Galaxies

Rujopakarn, W.; Dunlop, J.S.; Rieke, G.H.; Ivison, R.J.; Cibinel, A.; Nyland, K.; Jagannathan, P.; Silverman, J.D.; Alexander, D.M.; Biggs, A.D.; and 23 co-authors

2016, ApJ, 833, 12

53.) T-PHOT: PSF-matched, prior-based, multiwavelength extragalactic deconfusion photometry

Merlin, E.; Fontana, A.; Ferguson, H C.; Dunlop, J.S.; Elbaz, D.; Bourne, N.; Bruce, V.A.; Buitrago, F.; Castellano, M.; Schreiber, C.; and 12 co-authors

2016, ascl.soft09001M

54.) The dust-to-stellar mass ratio as a valuable tool to probe the evolution of local and distant star-forming galaxies

Calura, F.; Pozzi, F.; Cresci, G.; Santini, P.; Gruppioni, C.; Pozzetti, L.; Gilli, R.; Matteucci, F.; Maiolino, R.

2017, MNRAS, 465, 54

55.) Characterizing the evolving K -band luminosity function using the UltraVISTA, CANDELS and HUDF surveys

Mortlock, A.; McLure, R.J.; Bowler, R.A.A.; McLeod, D.J.; Mármol-Queraltó, E.; Parsa, S.; Dunlop, J.S.; Bruce, V.

2017, MNRAS, 465, 672

56.) The Hubble Space Telescope UV Legacy Survey of Galactic globular clusters - XI. The horizontal branch in NGC 6388 and NGC 6441

Tailo, M.; D'Antona, F.; Milone, A.P.; Bellini, A.; Ventura, P.; Di Criscienzo, M.; Cassisi, S.; Piotto, G.; Salaris, M.; Brown, T.M.; and 5 co-authors

2017, MNRAS, 465, 1046

57.) A deep ALMA image of the Hubble Ultra Deep Field

Dunlop, J.S.; McLure, R.J.; Biggs, A.D.; Geach, J.E.; Michalowski, M.J.; Ivison, R.J.; Rujopakarn, W.; van Kampen, E.; Pope, A.; Scott, D.; and 25 co-authors

2017, MNRAS, 466, 861

58.) Unveiling the nature of bright z ~ 7 galaxies with the Hubble Space Telescope

Bowler, R.A.A.; Dunlop, J.S.; McLure, R.J.; McLeod, D.J.
2017, MNRAS, 466, 3612

59.) EGG: hatching a mock Universe from empirical prescriptions
Schreiber, C.; Elbaz, D.; Pannella, M.; Merlin, E.; Castellano, M.; Fontana, A.; Bourne, N.; Boutsia, K.; Cullen, F.; Dunlop, J.; and 7 co-authors
2017, arXiv:1606.05354

60.) The ALMA Redshift 4 Survey (AR4S): I. The massive end of the z=4 main sequence of galaxies
Schreiber, C.; Pannella, M.; Leiton, R.; Elbaz, D.; Wang, T.; Okumura, K.; Labbé, I.
2017, arXiv:1606.06252

61.) Evolution of cosmic star formation in the SCUBA-2 Cosmology Legacy Survey
Bourne, N.; Dunlop, J S.; Merlin, E.; Parsa, S.; Schreiber, C.; Castellano, M.; Conselice, C.J.; Coppin, K.E.K.; Farrah, D.; Fontana, A.; and 12 co-authors
2016, arXiv:1607.04283

62.) No evidence for Population III stars or a Direct Collapse Black Hole in the z = 6.6 Lyman-alpha emitter 'CR7'
Bowler, R.A.A.; McLure, R.J.; Dunlop, J.S.; McLeod, D.J.; Stanway, E.R.; Eldridge, J.J.; Jarvis, M.J.
2016, arXiv:1609.00727

63.) The SCUBA-2 Cosmology Legacy Survey: the nature of bright submm galaxies from 2 deg² of 850-micron imaging
Michałowski, M.J.; Dunlop, J.S.; Koprowski, M.P.; Cirasuolo, M.; Geach, J.E.; Bowler, RA.A.; Mortlock, A.; Caputi, K.I.; Aretxaga, I.; Arumugam, V.; and 8 co-authors
2016, arXiv:1610.02409

64.) AGN-Host Connection at $0.5 < z < 2.5$: A rapid evolution of AGN fraction in red galaxies during the last 10 Gyr
Wang, Tao; Elbaz, D.; Alexander, D. M.; Xue, Y. Q.; Gabor, J. M.; Juneau, S.; Schreiber, C.; Zheng, X-Z.; Wuyts, S.; Shi, Y.; and 6 coauthors
2016, arXiv:1612.03179

65.) Characterizing elusive, faint dusty star-forming galaxies: a lensed, optically undetected ALMA galaxy at z~3.3
Santini, P.; Castellano, M.; Fontana, A.; Merlin, E.; Maiolino, R.; , Mason, C.; Mignano, A.; , Pilo, S.; Amorin, R.; Berta, S.; Bourne, N.; Calura, F.; Daddi, E.;Elbaz, D.; and 8 co-authors
2016, arXiv:1610.08515

Potential impact and main dissemination activities and exploitation results

1 Dissemination activities

1.1 Science conferences

One of the main goals of ASTRODEEP is to broaden the collaboration between the main teams that are active in Europe and world-wide in the exploitation of the deepest images of the Universe.

For this purpose, well-focused scientific meetings are crucial to establish connections, foster new ideas, and spread the latest concepts, news and discoveries among the scientists.

In this context, ASTRODEEP has organized a series of scientific meetings focused on gathering together the participants of the CANDELS and GOODS-HERSCHEL surveys to discuss well-defined scientific issues and goals.

These conferences were explicitly planned as specific actions in our project (Deliverables 7.3 and 7.6)

These conferences have been held in January 2015 and January 2016. We have organized another, similar one in January 2017, after the formal end of AstroDeep, with the same goal and format. They were organized inheriting the successful format of two previous CANDELS collaboration meetings organized in the same location. They have been organized by the Sexten Center for Astrophysics (SCfA), located in Sexten (Italy).

The main aim of SCfA is organising and hosting every year small and medium-size workshops and schools, in order to offer to scientists, working in the fields of Astrophysics, Cosmology and Physics, the opportunity to meet in an informal environment and to carry out collaborative work.

Each Conference lasted 4 full days, and about 40 talks were presented at each meeting. We explicitly decided to allow for long presentations (between 20 and 30 min each) in order to allow a proper presentation of the data and ample time for lively discussions. The choice of the format (a restricted number of participants, ample time for presentations, all participants located in the same location with ample time for informal meeting and discussions) proved crucial/invaluable for establishing excellent and creating relationships, and stimulating many fruitful discussions among the scientists involved.

We note that no ASTRODEEP funds have in practice been dedicated to the organization of the conferences, except for the bare cost of participation of the ASTRODEEP members.

At each conference we made sure to take advantage of the excellent opportunity (at least 5-8 talks each time) to present the technical and scientific results of ASTRODEEP to the community of interested scientists, who were carefully selected as representative of the main research groups worldwide. Most presentations by ASTRODEEP have been given by the younger participants of ASTRODEEP, thus also giving them the opportunity to present and re-affirm the originality and scientific value of their research.

The number of speakers not connected with ASTRODEEP has been increasing with time, as a

consequence of the shift in focus of the conferences (from ASTRODEEP-focused meeting to public conferences). They were 17/34 at the 2015 meeting, 23/38 at the 2016 one, and 30/38 at the last (2017) one.

The success of this initiative is confirmed by the fact that it has been decided to make the Sexten galaxy evolution workshop a regular annual event, even after the end of ASTRODEEP. We have already organized the third workshop in the series, again in Sesto, from 18 to 21 January 2017.

Further details can be found on the web pages of the conferences, at these links:

http://www.sexten-cfa.eu/en/conferences/2015/details/52-the-spectral-energy-distribution-of-high-redshift-galaxies-lessons-learned-and-open-questions.html

http://www.sexten-cfa.eu/en/conferences/2016/details/62-HST_Spitzer_Herschel.html

http://www.sexten-cfa.eu/en/conferences/2017/details/76-star-formation.html

1.2 The CDS interface: porting data access to the next generation

One of the main goals of the ASTRODEEP collaboration is to efficiently distribute the output of our work to the community. For this reason, CDS is one of the partners in our project. CDS is the established leader in Europe for data preservation and dissemination, and guarantees the long-term accessibility of all datasets and catalogues produced by the ASTRODEEP collaboration. In addition to the database, we have used specific ASTRODEEP resources and know-how to design and develop a dedicated web-portal with ad-hoc tools that will greatly enhance data access and mining capabilities for external users. The tool is more extensively described elsewhere (Science Report n. 3 and Deliverable 7.5) but we summarize here the basic features.

The portal is designed to provide a full access to the data (images, catalogues and high-level products) developed and released by ASTRODEEP, offering the option to browse the images and catalogues, to select specific classes of objects, to visualize individual objects and Spectral Energy Distributions (SEDs) in a full multi-wavelength approach.
It has been designed combining the know-how of the ASTRODEEP scientists (who figured out the requirements and the layout of the interface) with the technical skill of the CDS team, who realized the tool with the most advanced techniques.
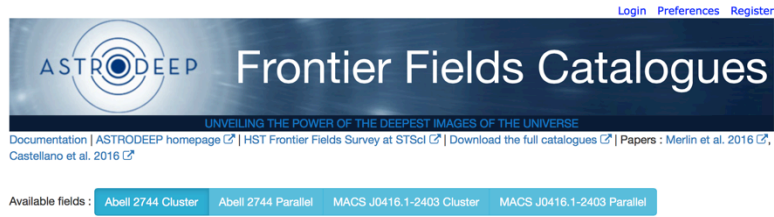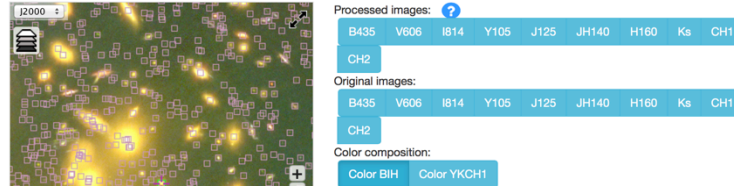
Figure 24 *A partial view of the CDS AstroDeep portal home page*

As can be seen in Figure 24, the external user can choose which band of which field he/she would like to visualize and they are displayed in the left panel. The positions of all detected sources are over-plotted on the displayed image. By selecting a single object, the most important information, such as magnitude in the *H*-band (MAG_H160) and best photometric redshift fit (ZBEST), are displayed next to the image along with the SED plot. Simultaneously the source is highlighted in the catalogue below, where more information is available. An example of this procedure is shown in Figure 25 and Figure 26.
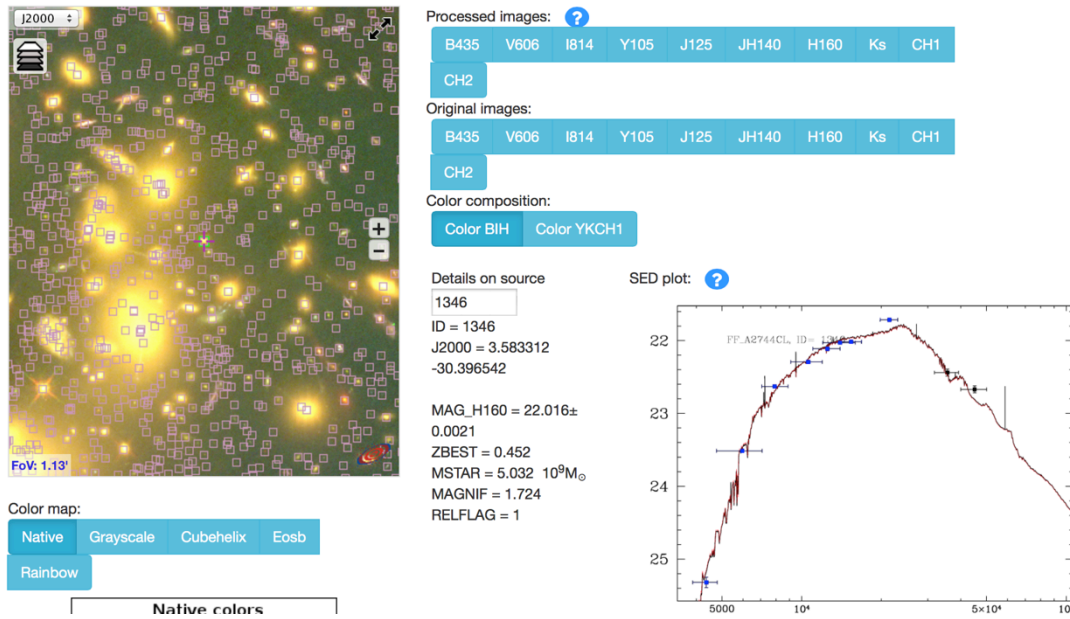
Figure 25 *When the user selects a source, in this case 1346, key information is displayed next to the image along with the SED plot*

ID ▲	RA	DEC	MAG_B435	MAG_V606	MAG_I814	MAG_Y105	MAG_J125	MAG_JH140	MAG_
1345	3.588135	-30.396181	99.0	99.0	99.0	99.0	99.0	99.0	99.0
1346	3.583312	-30.396542	25.3197	23.5141	22.6295	22.2928	22.111	22.0277	22.0157
1347	3.604188	-30.39719	25.4742	25.1148	24.933	24.5141	23.9682	23.7417	23.6969
1348	3.585216	-30.396502	-30.2137	29.5906	29.5475	28.9932	28.8876	28.6895	28.5254
1349	3.573801	-30.396477	-30.1137	28.6862	28.7755	28.5781	28.6591	28.1982	28.3528
1350	3.57682	-30.396534	28.7981	28.1468	28.2917	27.2953	27.1002	26.8696	26.9282
1351	3.607189	-30.39723	26.1346	24.5744	23.951	23.731	23.613	23.5614	23.5878
1352	3.592872	-30.396357	25.1403	23.2448	22.3654	21.9634	21.737	21.6345	21.5912
1353	3.572193	-30.396547	-29.616	28.1888	27.1722	27.1676	26.7692	26.5596	26.3893
1354	3.587674	-30.396469	99.0	99.0	99.0	99.0	99.0	99.0	99.0

Showing 1,344 to 1,355 of 3,587 entries
Download this XML table (VOTable format)

Share http://astrodeep.u-strasbg.fr/ff/   Copy URL to clipboard

Figure 26 *The same source is highlighted in the catalogue below the image, where more information is available.*

It is possible, for the external user, to add his/her own catalogue, in VOTable format, on the image in Aladin Lite using the link "Upload your catalogues". Such catalogues must first be uploaded to the user's personal storage space at the CDS. For correct visualization, RA and DEC columns have to be selected before sending the file to Aladin Lite. With the "View your catalogues" button, the user can get a list of all saved files and using the checkbox the file is then displayed in Aladin Lite. The colour that appears near the checkbox is the colour in which the sources are displayed in Aladin Lite. For more options in the analysis, data can be sent via Samp to external tools, like Aladin or TOPCAT.

1.3 Outreach activities

We describe here the activities that we performed to disseminate our results among the general public. They cover different aspects: exhibitions and presentations at outreach events; articles on non-professional journals, images and movies obtained from our data and tools that can convey the basic message of our work and can be used in talks and lectures.

1.3.1 "Fete de la Science"

The "Fête de la Science" is a French national event held each year, intended to promote interest in research and scientific activities towards kids and the general public. The Strasbourg observatory took part in a "Village des sciences", with some astronomy-related exhibitions (9-11/10/2015, 14-16/10/2016). One of the exhibitions was set up to raise awareness on the small fraction of the sky covered by the very deep fields such as those studied in ASTRODEEP, compared to the richness of their contents. A poster representing the Moon at the right apparent angular size was displayed side-by-side with the footprint of the HUGS K_s-band survey to compare their sky coverage. A high-resolution image of the HUGS field was also printed, so the public could realize how many galaxies are detected in such very deep surveys.

A few pictures for are attached for illustrations in Figure 27. The audience was estimated to be around 2500-3000 people.

Figure 27 ASTRODEEP at the 'Fete de la Science'
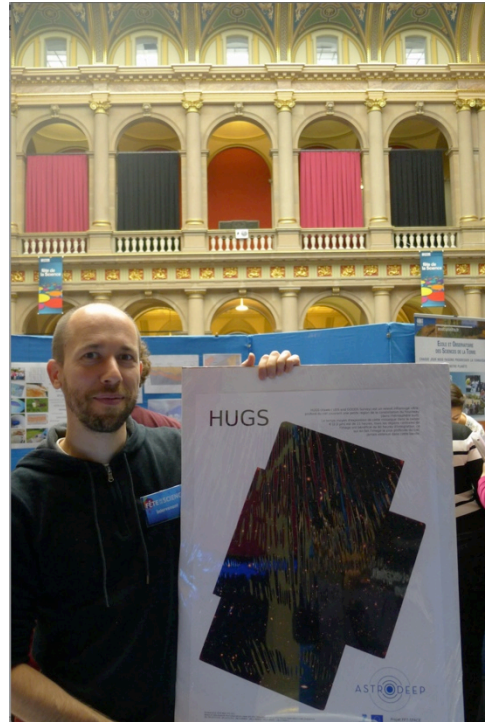
1.3.2 Articles.

We have prepared a general text that can be used to describe the main goal and results of ASTRODEEP. The text has been used for an article published in the Italian newspaper "Il Sole 24 Ore", on a special issue dedicated to advanced research. The article is added in the appendix.

1.3.3 Images

We have prepared a number of nice-looking images describing the images obtained and processed in ASTRODEEP, and how we cope with the main technical obstacle – the different resolution at the various wavelengths.

These images and movies are shown here (only a few frames are obviously shown for movies) and are available or used on the ASTRODEEP website.
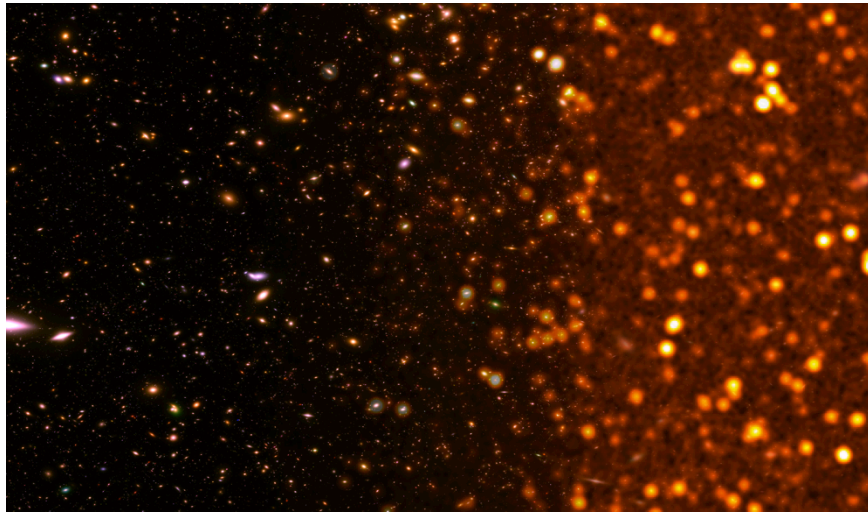
Figure 28 Image created with EGG, used for the EGG presentation. It shows a deep field observed at various resolution and wavelength
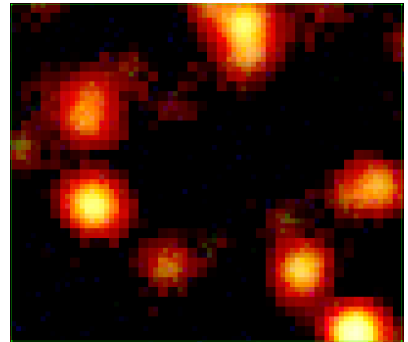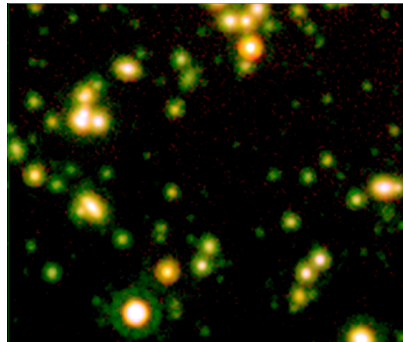

Figure 29 A region of the GOODS-S field observed with HST (left), Spitzer (center) and Herschel (right)
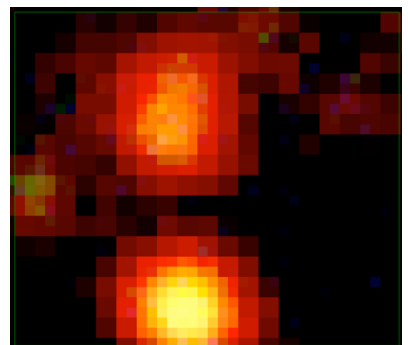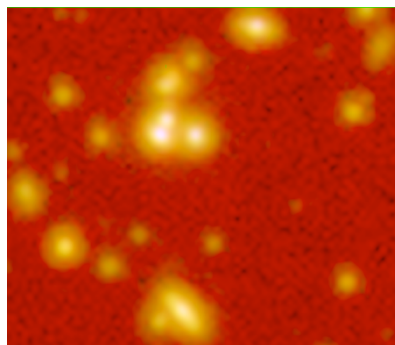

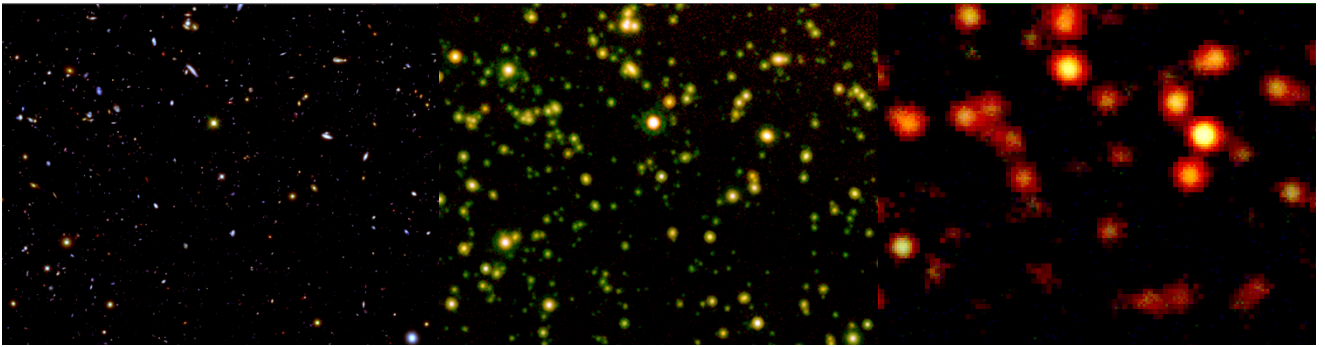**Figure 30 As above, a zoom-in of the previous image, with different color rendition**

Figure 31 A wider region in GOODS-S observed with HST (left), Spitzer (center) and Herschel (right).

1.3.4 Movies

We have also realized a number of movies to show the impact of different resolution and depth on the detectability and accuracy of photometry in the various instruments. These movies have been made with EGG, to obtain full control of the depth and resolutions that are displayed.

To maximize their impact, we have decided to adapt such movies to the forthcoming set of instruments from *JWST*, especially to compare it with *Spitzer*. For each of these movies we show here the first and last frame of the movie – the others show a resolution intermediate between the two, providing a smooth transition from the first to the last.
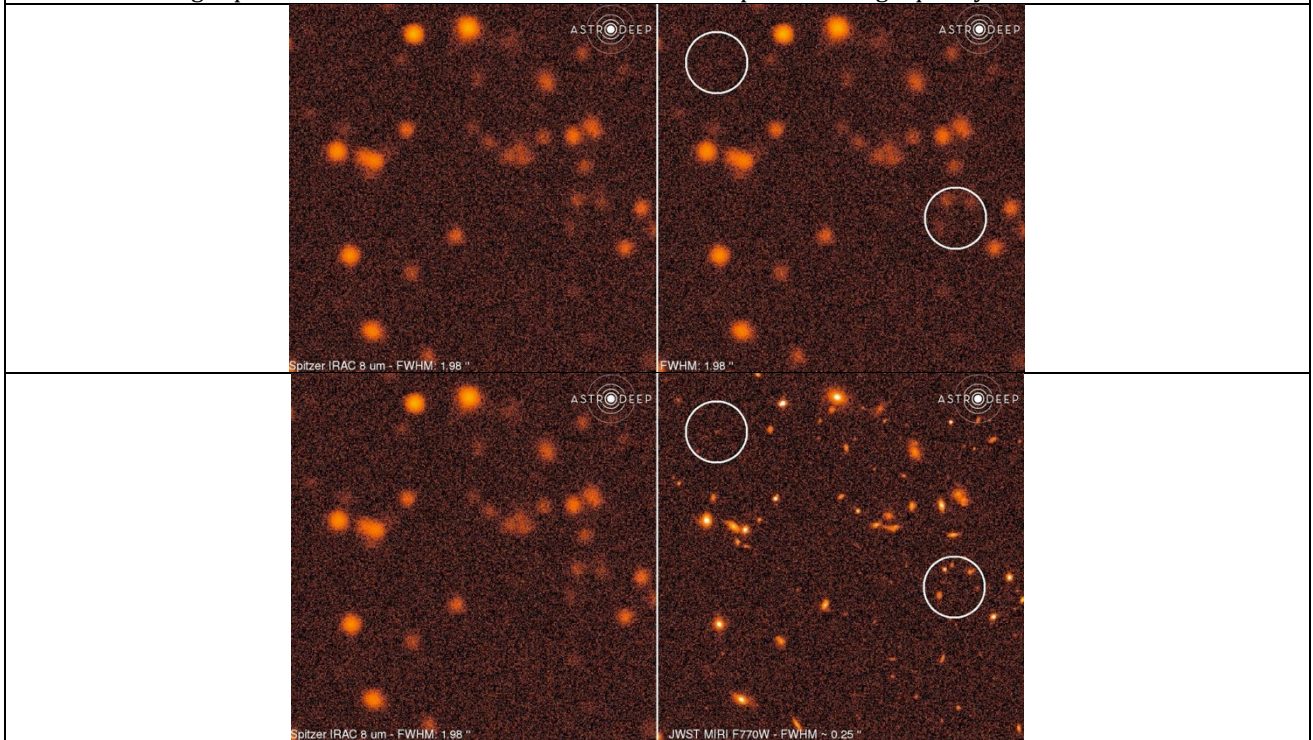
Simulation of the F444W filter in JWST – Comparison with Spitzer

The movie shows an extragalactic field with depth similar to GOODS-S as observed with Spitzer (on the left) and with JWST (right, end of the animation). The movie starts with the two images identical (upper row) and ends with the image on the right showing the expected JWST simulation. The transition from Spitzer to JWST is shown in the right panel and shows a smooth increase of the depth and image quality
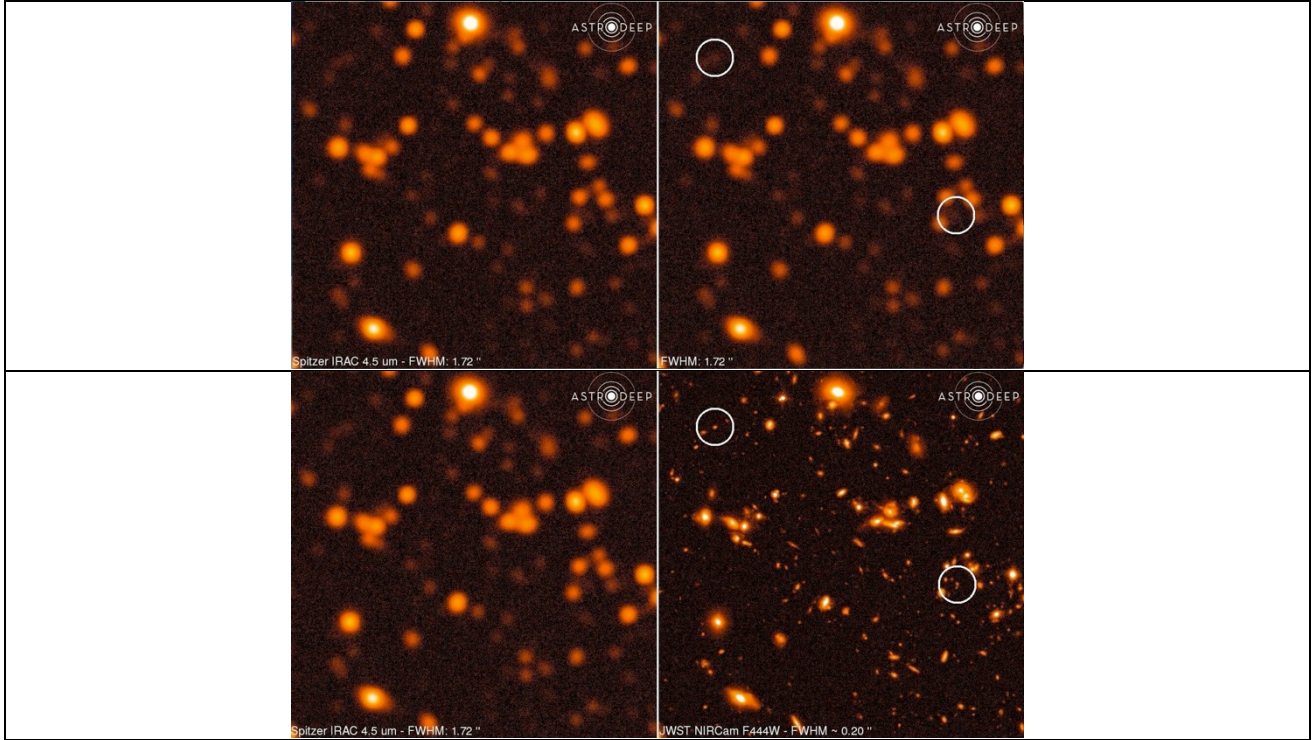

Simulation of the F770W filter in JWST – Comparison with Spitzer

The movie shows an extragalactic field with depth similar to GOODS-S as observed with Spitzer (on the left) and with JWST (right, end of the animation). The movie starts with the two images identical (upper row) and ends with the image on the right showing the expected JWST simulation. The transition from Spitzer to JWST is shown in the right panel and shows a smooth increase of the depth and image quality

Full tricromy JWST – Comparison with Spitzer

This movie shows a similar case but objects are represented with colors representatives of the colors between the I, H and F444W band. The movie shows an extragalactic field with depth similar to GOODS-S as observed with Spitzer (beginning of the animation) and with JWST (rnd of the animation).
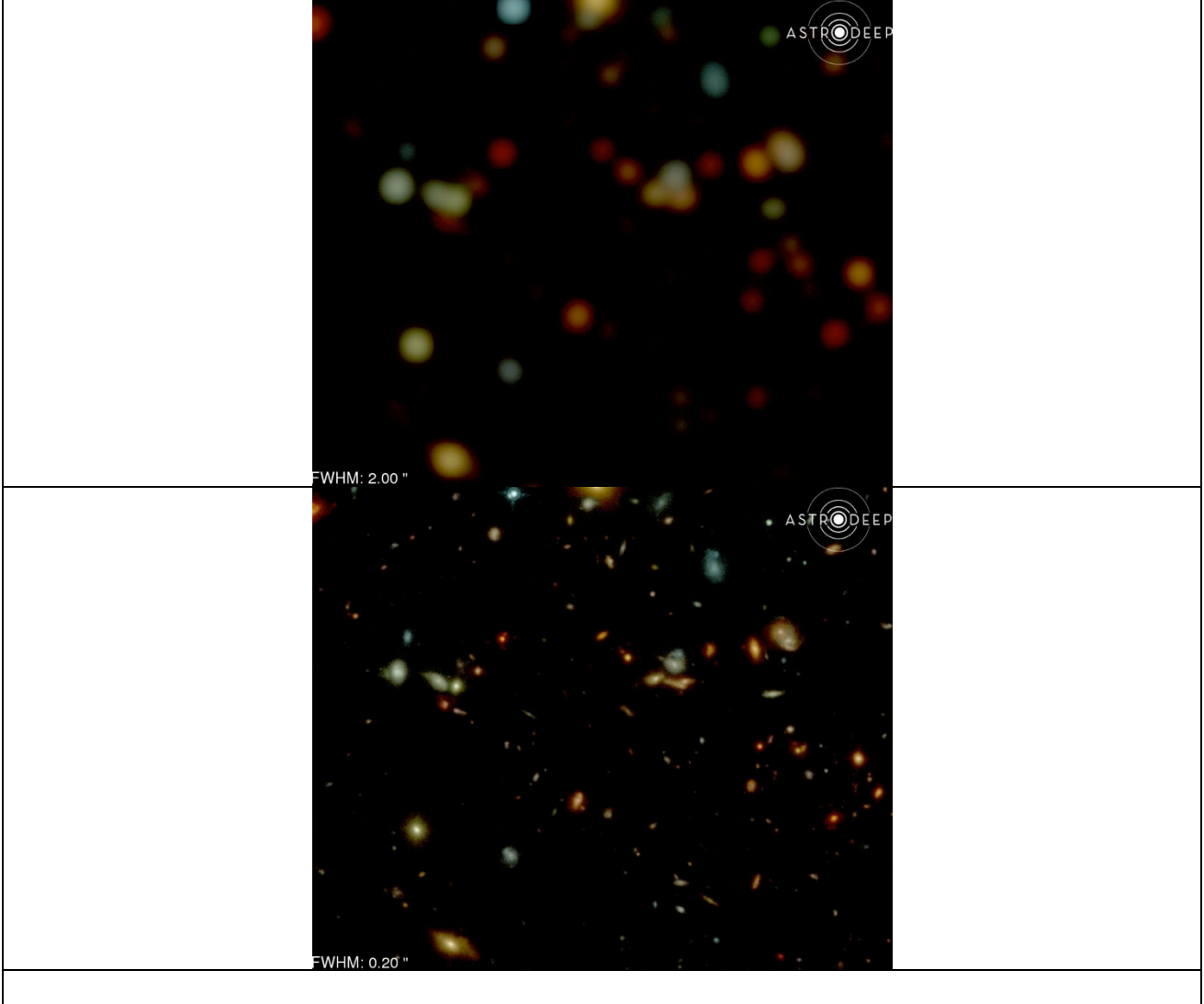
2 Potential impact and heritage

In this section we would like to provide a broader perspective on the potential impact of ASTRODEEP on the scientific community and its long-term heritage, beyond the specific results and dissemination activities.

The explicit aim of ASTRODEEP was to make "Europe the world leader in the exploitation of the deepest multi-frequency data from the major space and ground-based observatories". We can fairly claim that this goal has been successfully achieved.

The algorithms and tools that we have developed are clearly now the reference tools for the modern processing of deep extragalactic data, designed to extract most of the information from data that are typically observed under different image resolutions. Such tools are now being progressively adopted by several groups outside ASTRODEEP, including outstanding US

projects like GLASS. TPHOT is clearly the most successful case, with ~100 downloads and 31 citations after one year of release, currently used by several dozen of teams outside AstroDeep for applications that range from the preparation of JWST Granted Time proposals to big surveys with the South Pole Telescope.

We believe that the tools that we have assembled and delivered will be used, for years to come, and will be heavily applied to the new data sets that will be available from space and ground-based observatories.

The impact on the astronomical literature is also significant. On the technical side, the tools and the catalogues we have delivered so far are being used by more and more teams worldwide, and consequently the papers that presented and described these tools have been rapidly garnering citations. As we are still publishing the papers that present the latest catalogues and data, and considering that the peak in citation rate for a catalogue paper is typically reached some years after its release, we are confident that the impact on the literature of the already well-cited ASTRODEEP papers will continue to significantly increase over the next years, well beyond the completion of the project.

ASTRODEEP has not only been a project dedicated to data and algorithms. The scientific application of our data has been a major focus of our activity, and our young researchers have been stimulated to perform original and inventive work on many aspects of extragalactic astronomy. The team – and especially the young researchers hired under ASTRODEEP – has published 65 papers in high impact journals which directly acknowledge the support of ASTRODEEP. These papers span many topics, from the evolution of red high-redshift galaxies to the luminosity function at the highest redshifts, from galaxy morphology to the contribution of dusty galaxies to the global star-formation rates, and so on. Undoubtedly, these studies have had a major impact on the scientific literature and can be considered an important long-term heritage of ASTRODEEP.

Our team has demonstrated the capability to deliver data products of the highest quality within the planned timescale. This has certainly strengthened the credibility of our teams in participating in major observational campaigns in the field of extragalactic astronomy. In recent years members of our team have been awarded large observational projects with ALMA and VLT, and we believe that ASTRODEEP has contributed to the success of our proposals.

Perhaps the best evidence of this long-lasting legacy and impact is the major contribution that we are planning to offer to the forthcoming *JWST* and *Euclid* missions.

Concerning the latter, the algorithms developed with ASTRODEEP (and in particular T-PHOT) are currently being investigated and will likely be adopted for the data processing of the *Euclid* mission. Even more crucially, the expertise developed within ASTRODEEP has been directly conveyed into the *Euclid* framework, as two young researchers trained in ASTRODEEP (E. Merlin and M. Castellano) have been hired within the *Euclid* consortium to develop its photometric pipeline. In addition, ASTRODEEP post-doc F. Cullen has been invited into *Euclid,* to contribute to the case for the Blue Grism, and the *Euclid* Deep Survey. This work is founded on the ASTRODEEP simulated galaxy surveys created with EGG, coupled with the expertise developed by Cullen on emission-line galaxies and Grism data reduction.

In the case of *JWST*, our team is preparing to lead an ERS proposal for the first round of *JWST* proposals, which will likely evolve into a larger and more ambitious survey plan for Cycle-1

and Cycle-2. We shall propose an ambitious imaging survey that will deliver data on which the ASTRODEEP tools and know-how will be optimally used. Members of our team have also been invited to assume crucial responsibilities in the execution of other surveys that will be proposed within the ERS scheme. Given that participation to the ERS programs requires the proposers to guarantee the prompt processing and public release of the collected data, it is clear that all these proposals will heavily rely on the technical/scientific skill and credibility of ASTRODEEP to be competitive in the selection. Whatever the fate of these proposals, the ERS programme will undoubtedly deliver public data of interest, we look forward to continuing the collaboration among the ASTRODEEP partners into the *JWST* era..

Finally, we would like to mention in particular one of the main outcomes of ASTRODEEP – at least, the one we are particularly proud of. ASTRODEEP has proved a wonderful environment for nurturing the scientific and professional growth of the young and talented scientists that we have hired. ASTRODEEP has dedicated most of its budget to recruit as many as 14 young astronomers[1], sustaining them in the early phases of their career. Others, although not directly funded by ASTRODEEP, have interacted and collaborated with the team, sometimes providing the project with crucial contributions[2]. The international impact and the good salaries that European contracts can guarantee have certainly been instrumental in attracting some of the most talented young researchers in the field. We made every effort to assign them crucial responsibilities and grant them the full visibility that their original and dedicated work deserved, with many of the published ASTRODEEP papers led by one of these young researchers. Three of them have acquired a permanent position by now; many others have long-term research contracts, often in the context of long-term important projects like *JWST* and *Euclid*. We are confident that many of them will be able to eventually obtain a permanent position and will maintain the lively collaborations between them for the years to come.

3.0 Address of public website and relevant contact details

http://www.astrodeep.eu

http://www.astrodeep.org

Adriano Fontana,
INAF_Osservatorio Astronomico di Roma
00078 Monte Porzio Catone
Rome ITALY
tel: +39 06 94286456
mail: adriano.fontana@oa-roma.inaf.it

---

[1] Ricardo Amorin, Konstantina Boutsia, Nathan Bourne, Victoria Bruce, Fernando Buitrago, Marco Castellano, Laure Ciesla, Fergus Cullen, Marcella De Crescienzo, Emiliano Merlin, Paola Santini, Xinwen Shu, Tao Wang, Pascal Wassong
[2] Diego Paris, Corentin Schreiber.