Abstract

We present a multiband analysis of the six Hubble Frontier Field clusters and their parallel fields, producing catalogs with measurements of source photometry and photometric redshifts. We release these catalogs to the public along with maps of intracluster light and models for the brightest galaxies in each field. This rich data set covers a wavelength range from 0.2 to 8 μm, utilizing data from the Hubble Space Telescope, Keck Observatories, Very Large Telescope array, and Spitzer Space Telescope. We validate our products by injecting into our fields and recovering a population of synthetic objects with similar characteristics to those in real extragalactic surveys. The photometric catalogs contain a total of over 32,000 entries, with 50% completeness at a threshold of magAB ∼ 29.1 for unblended sources and magAB ∼ 29 for blended ones, in the IR-weighted detection band. Photometric redshifts were obtained by means of template fitting and have an average outlier fraction of 10.3% and scatter σ = 0.067 when compared to spectroscopic estimates. The software we devised, after being tested in the present work, will be applied to new data sets from ongoing and future surveys.

Unified Astronomy Thesaurus concepts: HST photometry (756); Galaxy clusters (584); Intracluster medium (858)

1. Introduction

To develop a comprehensive picture of the primordial building blocks of the universe, identification and study of properties of the youngest galaxies (<1 Gyr) soon after the big bang are essential. Given their enormous distance, these systems are best detected at near-IR wavelengths. However, because of high sky background at these wavelengths, such observations need to be done from space, i.e., using the Hubble Space Telescope (HST) or James Webb Space Telescope (JWST). This also requires multi-waveband observations spanning a range of wavelengths from optical to near-IR. Furthermore, the study of galaxies at these redshifts is often biased, as we mainly sample the intrinsically brighter populations.

To accomplish the above, very deep multi-waveband observations over a large area are needed to detect statistically representative samples of very high redshift galaxies. This is currently beyond the reach of the largest available telescopes. To achieve such depths, one could leverage the natural phenomenon of gravitational lensing by targeting rich clusters of galaxies (Schneider 1984; Blandford & Narayan 1986, and references therein; see Kneib & Natarajan 2011 for a review). This magnifies fluxes from high-redshift sources located behind massive galaxy clusters, helping to probe deeper into the universe. This will be highly beneficial even after the largest space telescopes (i.e., JWST), with their unprecedented sensitivity, are commissioned. It has been demonstrated that JWST will be able to shed light on the mechanisms that reionized the intergalactic medium at $6 < z < 10$ only by exploiting the cluster lensing phenomenon, to observe galaxies with $M < 10^9 M_\odot$ (Atek et al. 2018).

The Hubble Frontier Field program (HFF; Lotz et al. 2017) is a survey designed with these objectives in mind. With an allocation of 630 HST orbits, it performed deep observations of six very massive clusters and their parallel fields in optical and near-IR bands, with Advanced Camera for Surveys (ACS) and Wide Field Camera 3 (WFC3), respectively. In addition, each field has been extensively observed in the mid-IR regime (between 3 and 5 μm) with the Infrared Array Camera (IRAC) on board the Spitzer Space Telescope.

The six clusters span the redshift range $0.3 < z < 0.55$, and their magnification power allows detection of galaxies up to $z ∼ 9$, i.e., at the reionization epoch (see Barkana & Loeb 2001; Miralda-Escudé 2003). HFF followed the successful tradition of deep, pencil-beam HST observations as in Hubble Deep Field (Williams et al. 1996), Hubble Ultra-Deep Field (HUDF; Beckwith et al. 2006), and CANDELS (Koekemoer et al. 2011), as well as programs covering wider areas like the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) and Galaxy Evolution from Morphology and SED (GEMS; Rix et al. 2004). With respect to galaxy cluster observations, the Cluster Lensing And Supernovae survey with Hubble (CLASH; Postman et al. 2012) paved the way for the HFF.

The HFF clusters will be the reference fields for the exploration of the distant universe. At present, additional HST coverage is provided by the “Beyond Ultra-deep Frontier Fields and Legacy Observations” (BUFFALO; Steinhardt et al. 2020), which aims at covering the outskirts of the HFF clusters over the same wave bands and to the same depth. Most notably, sources in the HFF clusters will also be the JWST targets, with some of them already selected by both Guaranteed Time Observation teams4 and an Early Release Science Program (Treu et al. 2017). Eventually, the knowledge acquired from the HFF shall be applied to analyze data from an unprecedented number of galaxy clusters that the Euclid and Roman space telescopes will discover by surveying thousands of square degrees of the sky (Laureijs et al. 2011; Spergel et al. 2015).

Apart from providing access to the high-redshift universe, the HFF also allows the study of properties of dark matter (e.g.,

4 https://jwst-docs.stsci.edu/jwst-opportunities-and-policies/jwst-cycle-1-guaranteed-time-observations-call-for-proposals/jwst-gto-observationSpecifications
Jauzac et al. (2015, and references therein) and the role of “environment” in the evolution of galaxies at $z < 1$ (e.g., Gao et al. 2005; Paranjape et al. 2018). It will also provide a standard sample of cluster galaxies, to be compared with nearby systems. The HFF parallel fields provide similar data to the same depth that would minimize selection effects and biases in any environmental study of galaxy populations. The cornerstone of all these studies is a multi-wave-band photometric catalog, containing self-consistent photometry for every detected source. Producing such a catalog is particularly challenging, given the depth of the imaging data, analysis of images in the crowded fields, the wide wavelength baseline (which implies substantial changes in the point-spread function [PSF]) required, and contamination from the intracluster light (ICL). The task has already been fulfilled by two distinct teams (AstroDeep⁵ and DeepSpace⁶) that released their respective HFF catalogs to the astronomical community (Castellano et al. 2016; Di Criscienzo et al. 2017; Shipley et al. 2018; Bradač et al. 2019). Given the challenges mentioned above in identifying individual sources, in performing accurate photometric measurements for the HFF galaxies, and in simultaneously dealing with a multiple-parameter space, it is important to apply completely independent techniques to generate catalogs for the same clusters. This allows a better understanding of the final data products, selection effects, and the caveats in the data and the data processing pipelines. In particular, the different data reduction steps could introduce systematic effects in the final results, which need to be studied. Such problems and inaccuracies in source identification and photometry will also be reflected into selection of different populations of galaxies. For example, comparing two independent estimates of photometric redshifts can be instrumental to selecting the most robust galaxy candidates at $z > 8$ (McLeod et al. 2016). The existing catalogs produced from the above-mentioned studies contain serious discrepancies in the photometry of the same galaxies. An independent analysis will help us study the origins of these differences.

This study, we generate new galaxy catalogs for the six HFF clusters and their parallel fields, available at doi:10.5281/zenodo.5338978. The results include photometric catalogs, photometric redshifts, ICL maps, and surface brightness models to remove the brightest galaxies from our images in order to detect fainter galaxies to deeper levels. The parallel fields provide ancillary deep-field data and will also serve as control samples.

The main difference between the strategy here and the previous methods is that we provide models of the ICL maps, as well as perform forward modeling of photometry to characterize biases and uncertainties in flux measurements. Moreover, the pipelines developed for this study will be applied to the new BUFFALO data (A. Pagul et al. 2021, in preparation) to extend the HFF catalogs to wider areas, providing self-consistent data for the HFF and BUFFALO galaxies.

The paper is organized as follows. In Section 2 we introduce the data sets. Section 3 presents a detailed study of the data reduction, followed by source extraction in Section 4. In Section 5, we study the completeness and photometric uncertainties in HFF clusters, as well as include comparisons between our results and previous works. In Section 6 we present the photometric redshifts for the HFF measurements. In Section 7, we discuss the lensing measurements. Section 8 presents information about the released data products. Section 9 summarizes our results.

Throughout this paper we assume standard cosmology with $\Omega_M = 0.23$, $\Omega_{\Lambda} = 0.76$, and $H_0 = 73$ Km s$^{-1}$ Mpc$^{-1}$. Magnitudes are in the AB system.

## 2. The Data

### 2.1. HST Observations

The HFF observations are the deepest of galaxy clusters to date, and second only to the HUDF when also considering the blank fields. This results from a total of 840 orbits, performed through coordinated HST parallel observations in the following filters: F435W, F606W, and F814W in ACS and F105W, F125W, F140W, and F160W in WFC3. These seven bands reached a depth of $m_{AB} \sim 29$ mag for point sources within a 0.4″-diameter aperture (Lotz et al. 2017). Through a different program, UV observations in the F275W and F336W filters of two HFF clusters, A2744 and MACS J0717, were carried out, with details presented in Alavi et al. (2016).

The gravitational potential of the clusters’ halo, besides binding together the galaxies in the system, produces a lensing magnification that could detect background objects to an intrinsic brightness of 30–33 mag, i.e., 100–1000 times fainter than previous surveys. Details of the HFF survey design are provided in Lotz et al. (2017). In Table 1 we report the main characteristics of the six clusters, with a summary of the observations in Table 2. We use mosaics that have been reduced by the Frontier Fields team, with a pixel scale of 0.06 pixel$^{-1}$. These images have been reduced using

### Table 1

| Field       | Cluster Center (J2000) | Parallel Center (J2000) | $z_{\text{cl}}$ | $M_{\text{vir}}$ | $L_{X}$ |
|-------------|------------------------|-------------------------|----------------|-----------------|---------|
| A370        | 02:39:52.9, $-01:34:36.5$ | 02:40:13.4, $-01:37:32.8$ | 0.375 | $\sim 1 \times 10^{15}$ | $1.1 \times 10^{45}$ |
| MACS J0717.5+3745 | 07:17:34.0, +37:44:49.0 | 07:17:17.0, +37:49:47.3 | 0.545 | $\sim 2-3 \times 10^{15}$ | $3.3 \times 10^{45}$ |
| MACS J0416.1−2403 | 04:16:08.9, −24:04:28.7 | 04:16:33.1, −24:06:48.7 | 0.396 | $1.2 \times 10^{15}$ | $1.0 \times 10^{45}$ |
| AS 1063     | 22:48:44.4, −44:31:48.5 | 22:49:17.7, −44:32:43.8 | 0.348 | $1.4 \times 10^{15}$ | $1.8 \times 10^{45}$ |
| A2744       | 00:14:21.2, −30:23:50.1 | 00:13:53.6, −30:22:54.3 | 0.308 | $1.8 \times 10^{15}$ | $3.1 \times 10^{45}$ |
| MACS J1149.5+2223 | 11:49:36.3, +22:23:58.1 | 11:49:40.5, +22:18:02.3 | 0.543 | $2.5 \times 10^{15}$ | $1.8 \times 10^{45}$ |

⁵ http://www.astrodeep.eu/frontier-fields/
⁶ http://cosmos.phy.tufts.edu/~danilo/HFF/Home.html
⁷ https://archive.stsci.edu/prepds/frontier/
Table 2

| Band    | FWHM (arcsec) | \( \lambda_{\text{pivot}} \) (\( \AA \)) |
|---------|---------------|------------------------------------------|
| F275W  | 0.075         | 2707                                     |
| F336W  | 0.109         | 3355                                     |
| F435W  | 0.109         | 4329                                     |
| F606W  | 0.112         | 5922                                     |
| F814W  | 0.111         | 8045                                     |
| F105W  | 0.175         | 10551                                    |
| F125W  | 0.176         | 12486                                    |
| F140W  | 0.172         | 13923                                    |
| F160W  | 0.173         | 15369                                    |
| \( K_s \) | 0.364       | 21524                                    |
| I1     | 1.29          | 35634                                    |
| I2     | 1.42          | 45110                                    |
| I3     | 1.50          | 57993                                    |
| I4     | 1.84          | 79595                                    |

**Note.** Representative PSF FWHM values for the photometric bands in this study. These values were calculated for the A2744 cluster. Bands F275W and F336W are only available for A2744 and MACS J0717. IRAC channels 3 and 4 are only available for A2744, A370, AS 1063, and MACS J0717.

The HST science data products pipeline (Koekemoer et al. 2011). Data from other HST programs using the same filters have also been included by the HFF team, with all the exposures aligned relative to each other using TweakReg. Other steps in the reduction process include correction of standard imaging artifacts, bad pixel and cosmic-ray rejection, geometric distortion correction, and image stacking using AstroDrizzle (Gonzaga et al. 2012). Throughout this work we will focus on the analysis and detection of these previously reduced and calibrated data.

2.2. Ancillary Data

A number of independent observations of the HFF have generated complementary data to those available from HST. The Spitzer Space Telescope dedicated more than 1000 hr of Director’s Discretionary Time to obtain IRAC 3.6 \( \mu m \) (channel 1) and 4.6 \( \mu m \) (channel 2) imaging down to the depths of 26.5 and 26.0 mag in cluster and parallel fields, respectively. These observations are crucial for photometric redshift measurement and for identifying low-redshift interlopers and are beneficial in constraining galaxy properties since they provide a good proxy for galaxy stellar mass. In addition, there are legacy IRAC observations at 5.8 \( \mu m \) (channel 3) and 8.0 \( \mu m \) (channel 4), which are also included in our analysis. To produce the Frontier Fields mosaics, the following Spitzer Program IDs were used:

1. A2744: 83, 90275
2. MACS J0416.1−2403: 80168, 90258
3. MACS J0717.4+3745: 40652, 60034, 90009, 90259
4. MACS J1149.4+2223: 60034, 90009, 90260
5. AS 1063 (RXC J2248.7−4431): 83, 10170, 60034
6. A370: 137, 10171, 60034

Another follow-up program is the \( K_s \)-band Imaging of the Frontier Fields (KIFF; Brammer et al. 2016), carried out with the High Acuity Wide Field \( K \)-band Imager (HAWK-I) at the Very Large Telescope (VLT). This reached a depth of 26.0 mag (5\( \sigma \), point-like sources) for A2744, MACS-0416, AS 1063, and A370 clusters in the southern hemisphere. The \( K \)-band imaging campaign of the HFF also used the Multi-Object Spectrometer for Infrared Exploration (MOSFIRE) at Keck to observe MACS-0717 and MACS-1149 in the northern hemisphere to a 5\( \sigma \) depth of 25.5 and 25.1 mag, respectively.

Table 3 provides a summary of the available ancillary data.

3. Data Reduction

The workflow followed for the data processing in this work is presented in Figure 1. The filter throughput for the ACS and WFC instruments is shown in Figure 2 (top panel).

3.1. Modeling the Point-spread Function in Different Bands

Accurate knowledge of the PSF as a function of wavelength and optical system is crucial to perform consistent photometry within a “panchromatic” baseline. Knowledge of the PSF for HFF galaxies is needed to reconstruct their intrinsic morphology, as well as mapping the ICL (Section 3.2). To perform consistent multi-band photometry, taking into account band-to-band variations, we need to convolve images to the same PSF. This task, known as “PSF matching,” is performed via kernel convolution. Images in multiwavelength surveys are often affected by different diffraction levels, which make it difficult to obtain homogeneous measurements, e.g., of aperture photometry. Figure 3 illustrates the broad range of PSF sizes by showing postage stamps extracted from a few HFF images from \( \sim 0.3 \) to \( \sim 2 \mu m \). Besides the optical performance of the instrument itself, the final PSF model also depends on the specific observing strategy of the survey (depth, dithering, etc.) and, in the case of ground-based facilities, also on the seeing conditions (in the case of Figure 3, VLT/HAWK-I).

For HST data (divided between cluster and parallel fields), PSFs were estimated by stacking unsaturated stars. We use IRAF to visually inspect stellar light curves and determine whether an object is saturated. Then, we stack unsaturated stars using the IDL routine Star Finder (Diodati et al. 2000). The resulting PSF models used in PSF matching are \( 2\pi \times 2\pi \) postage stamps in the HST and \( K_s \) bands, and 28\( \times \)56 for Spitzer. These sizes are chosen in order to enclose diffraction spikes (as one can see in Figure 3).

Our procedure works well for the HST images and the \( K_s \) band, as the variation of the PSF across the image is small. However, stacking unsaturated stars does not produce a robust result in the mid-IR Spitzer channels. This is due to large variations of the PSF as a function of the position on the image, as well as the asymmetry of its shape. This makes IRAC PSFs depend on the orientation of the camera. Moreover, IRAC channel 1 and 2 pixels undersample the response of a point source. Thus, instead of Star Finder, we use a synthetic pixel response function (PRF) that combines the information on the PSF, the detector sampling, and the intrapixel sensitivity variation in response to a point-like source. A PRF model for a given position on the IRAC mosaic is generated by the code PRFMap (A. Faisst, 2021 private communication) by taking into account the single-epoch frames contributing to that mosaic. To do so,

---

8 Part of the DrizzlePac software suite: https://www.stsci.edu/scientific-community/software/drizzlepac.html.
9 Retrieved from the Spitzer Heritage Archive, https://sha.ipac.caltech.edu/applications/Spitzer/SHA/.
10 See the Spitzer/IRAC handbook at https://irsa.ipac.caltech.edu/data/Spitzer/docs/irac/calibrationfiles/psfrpf/.
11 More information in the the Spitzer/IRAC handbook at https://irsa.ipac.caltech.edu/data/Spitzer/docs/files/Spitzer/simfitreport52_final.pdf.
PRFMap stacks individual PRF models with the same orientation of the frames, resulting in a realistic, spatially dependent PSF model (an example for IRAC channel 1 is shown in Figure 3).

We use GalSim in order to calculate the FWHM for each PSF/PRF listed in Table 2. More specifically, we use the calculateFWHM function, which computes the maximum intensity of the PSF $I_0$, the centroid of the PSF intensity, the first pixel from the centroid at which the intensity is $I < I_0/2$, and the last pixel at which the intensity is $I > I_0/2$ and then linearly interpolates between these last two to estimate the value of the FWHM. For the Spitzer bands, we input into this function an averaged stack of all the normalized PRFs modeled in the field for a more conservative estimate.

### 3.2. Modeling the Intracluster Light

Given the richness of HFF clusters, those fields are particularly crowded with a significant probability of having cluster members aligned along the same line of sight of more

---

**Figure 1.** Workflow of the data processing as performed in the present work.

**Table 3**

Existing Multiwavelength HFF Coverage from Follow-up Programs, as Used in the Present Work

| Field          | Observatory          | Wavelengths   | Depth       | Reference                          |
|----------------|----------------------|---------------|-------------|------------------------------------|
| A370           | VLT/HAWK-I           | 2.2 μm        | ~26.18      | Brammer et al. (2016)              |
|                | Spitzer IRAC 1.2     | 3.6 μm, 4.5 μm| ~25.19, 25.09| (PI: T. Soifer and P. Capak)       |
|                | Spitzer IRAC 3,4     | 5.8 μm, 8.0 μm| ~23.94, 23.39| See Section 2.2                   |
| MACS J0717.5+3745 | Keck/MOSFIRE      | 2.2 μm        | ~25.31      | Brammer et al. (2016)              |
|                | Spitzer IRAC 1, 2, 3, 4 | 3.5 μm, 4.5 μm| ~25.04, 25.17| (PI: T. Soifer and P. Capak)       |
|                | Spitzer IRAC 3,4     | 5.8 μm, 8.0 μm| ~23.94, 23.39| See Section 2.2                   |
| MACS J0416.1−2403 | VLT/HAWK-I           | 2.2 μm        | ~26.25      | Brammer et al. (2016)              |
|                | Spitzer IRAC 1, 2     | 3.5 μm, 4.5 μm| ~25.31, 25.44| (PI: T. Soifer and P. Capak)       |
|                | Spitzer IRAC 3,4     | 5.8 μm, 8.0 μm| ~22.96, 22.64| See Section 2.2                   |
| AS 1063        | VLT/HAWK-I           | 2.2 μm        | ~26.31      | Brammer et al. (2016)              |
|                | Spitzer IRAC 1.2     | 3.6 μm, 4.5 μm| ~25.04, 25.04| (PI: T. Soifer and P. Capak)       |
|                | Spitzer IRAC 3,4     | 5.8 μm, 8.0 μm| ~22.78, 22.45| See Section 2.2                   |
| A2744          | VLT/HAWK-I           | 2.2 μm        | ~26.28      | Brammer et al. (2016)              |
|                | Spitzer IRAC 1.2     | 3.6 μm, 4.5 μm| ~25.32, 25.08| (PI: T. Soifer and P. Capak)       |
|                | Spitzer IRAC 3,4     | 5.8 μm, 8.0 μm| ~22.78, 22.45| See Section 2.2                   |
| MACS J1149.5+2223 | Keck/MOSFIRE      | 2.2 μm        | ~25.41      | Brammer et al. (2016)              |
|                | Spitzer IRAC 1, 2     | 3.5 μm, 4.5 μm| ~25.24, 25.01| (PI: T. Soifer and P. Capak)       |

**Note.** The table provides a summary of the available ancillary data, with their filter transmission curves shown in Figure 2 (bottom panel). The $5\sigma$ point-source depth was estimated by integrating the noise in a Gaussian PSF aperture with the values of FWHM given in Table 2.
distant background galaxies. The analysis of background systems is also impaired by the ICL, i.e., the residual emission from stars that are generally not bound to any cluster galaxy (Morishita et al. 2017; Montes & Trujillo 2019; Sampaio-Santos et al. 2021). Blending between ICL and bright cluster galaxies is also a source of uncertainty. In an effort to alleviate these effects, we attempt to model ICL and cluster members using GALFIT (Peng et al. 2010) and GALAPAGOS-2 (Barden et al. 2012; Häußler et al. 2013), respectively, following a similar procedure to that in Morishita et al. (2017). Because the severity of the ICL–galaxy blending decreases toward the bluer wavelengths, we perform these fits for all but the bluest HST bands used here (i.e., F275W and F336W).

In this section we focus on ICL modeling, while in Section 3.3 we will describe the procedure to fit the brightest cluster members. Subtracting their flux from the images will allow us to detect some of the faintest objects that otherwise would remain hidden by their overshadowing neighbors (see discussion in Section 4).

To model the ICL, we follow the methodology presented in Morishita et al. (2017). We first run Source Extractor (Bertin & Arnouts 1996; hereafter SE) on each image/band to get a first-pass estimate of the morphological parameters of detected galaxies and their segmentation map. In this way we can create a mask removing the detected pixels of sources fainter than 26 mag, as the fitting procedure (described below) is less reliable below that threshold. Then, we produce an 18” × 18” (300 × 300 pixel) patch for every remaining source, with the patch centered at the coordinate identified by SE as the centroid of the galaxy. By means of GALFIT, we simultaneously fit single Sérsic profiles to all the objects included in that patch. For patches whose GALFIT fits failed in the first instance, we randomly move the cutout center, creating five additional 18” × 18” patches on which we run the same algorithm.

It should be noted that the number of patches across the field depends on the projected density of sources. This is illustrated in Figure 4 (left panel) by showing a representative “coverage map” for A2744. In case a given pixel with coordinates (x, y) is only included in one cutout, the ICL emission (FICL) is defined as the local background measurement found by GALFIT (namely, the sky value parameter). If there are overlapping cutouts in (x, y), we use the inverse \( \chi^2 \)-weighted mean of their background measurements:

\[
F_{\text{ICL}}(x, y) = \frac{\sum_i s_i(x, y) / \chi^2_i(x, y)}{\sum_i 1 / \chi^2_i(x, y)},
\]

where \( s_i \) and \( \chi^2_i \) are the sky fit and \( \chi^2 \) values from GALFIT for the \( i \)th cutout. The overall ICL map, in the case of A2744, is shown in Figure 4 (middle panel).

The merged ICL map, represented by \( F_{\text{ICL}}(x, y) \), may still contain some sharp features owing to the finite nature of the cutouts used to build it. In order to avoid these features, we introduce an additional step to the method presented in Morishita et al. (2017) by smoothing the merged ICL map with a Gaussian kernel. The size of this kernel is chosen by analyzing the (radially averaged) power spectrum of the coverage map (see the left panel of Figure 4) defined as

\[
P(k) = \sum_{x, y} I_{\text{ICL}}(x, y) e^{-ikr},
\]

where \( r \) is the radial distance from the center of the image, \( r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \); \( x, y \) are the coordinates of each pixel; \( I_{\text{ICL}}(x, y) \) is the value of the coverage map at the pixel \( x, y \) at a distance \( r \) from the center; and \( x_0, y_0 \) are the coordinates of the center of the image. We find this scale to be at \( k \approx 0.08 \), i.e., about 72 pixels, or 4”32. In the right panel of Figure 4 we can see that this strategy successfully mitigates the impact of sharp features in the merged ICL map. As a sanity check, we generate a histogram of the sky background by masking sources using the corresponding segmentation map. We then compare this between the original image and the ICL-subtracted image. We find that the resulting histogram of background pixels is more symmetric, narrower, and centered at zero in the ICL-subtracted image (Figure 5).

The resulting map, shown in Figure 4 for the A2744 cluster, serves not only as a representative model of the ICL but also as a background correction for the entire HFF mosaic. Figure 4 also shows the difference between the original map (middle image) and the one after kernel filtering (rightmost image): the sharp flux gradients caused by finite box numbers and sizes are mitigated.

To model the ICL in the \( K_s \) and Spitzer images, we use a local background routine built into the T-phot software (Merlin et al. 2016b), which calculates a background template for each object and merges them into a single image. As done in the optical and other near-IR bands, we smooth the image with a representative kernel and subtract the result from the HFF mosaic. The smoothed ICL mosaics will be released as two-dimensional arrays in the same units (ADU per second) as the HFF mosaics.
**Figure 3.** Representative examples of PSF for the instruments used in this study, corresponding to a $0.006 \text{ pixel}^{-1}$ scale normalized with the $z$Scale algorithm. From left to right, panels show F336W (WFC3), F606W (ACS), F125W (WFC3), $K_s$ (VLT), and IRAC channel 1 (Spitzer). See Section 3.1 for more details.

**Figure 4.** We illustrate the steps to generate the ICL map (as described in Section 3.2). Here the example is presented for the F160 band in A2744. Top left panel: “coverage map” showing the number of fit cutouts that overlap with each pixel. Top middle panel: resulting ICL map for the same band and cluster created by combining the fit background value of each cutout. Overlapping stamps are stacked using Equation (1). Top right panel: final ICL map after smoothing the map in the top middle panel with a representative Gaussian kernel. More details about kernel creation are provided in Section 3.2. Bottom left panel: original A2744 F160W science image. Bottom right panel: F160W image after smoothed ICL (top right panel) subtraction.
3.3. Removal of the Bright Galaxies in Clusters

After correcting the cluster images for ICL and sky background, we model the brightest galaxy members of each cluster (i.e., galaxies with MAGAUTO_{F160W} < 19). These objects are selected via a first-pass SExtractor run in the F160W band. The aim is to remove them and their diffuse light so that one could push the observations deeper. We use the publicly available code GALAPAGOS-M\(^{12}\) (Barden et al. 2012; Häußler et al. 2013), which is a software that automates source detection and bulge–disk Sérsic modeling and takes advantage of multiwavelength information in its parameter fits.

The advantage of using GALAPAGOS-M is in its ability to perform single- and multicomponent (bulge/disk decomposition) fits and to input galaxies using information from multiple bands simultaneously. In order to robustly measure color gradients in large galaxies, we restrict the degrees of freedom in the GALAPAGOS-M fits by imposing a wavelength dependency (with a quadratic function) for the half-light radius and Sérsic index. Compared to fitting morphology independently in each band, this approach is stabler and tightens the constraints of the morphological parameters measured for each galaxy (see Häußler et al. 2013). However, a multicomponent model is susceptible to overfitting the images; therefore, we use the residual flux fraction (RFF, as in Hoyos et al. 2012) to assess the number of components that most effectively models the light profile of bright cluster members; such residual flux is obtained by subtracting the model from the input image. The residual flux fraction is a measure of signal excess in the residual image, not due to background fluctuations, and is defined as

\[
\text{RFF} = \frac{\sum |I_{x,y} - I_{x,y}^{\text{model}}| - 0.8 \times \sum \sigma_{BKG,x,y}}{\sum I_{x,y}},
\]

where \(I_{x,y}\) and \(I_{x,y}^{\text{model}}\) are, respectively, the observed and model fluxes for a given pixel with coordinates \((x,y)\), while \(\sigma_{BKG,x,y}\) is the background rms in the same location. We sum over the pixels associated with a given galaxy. The 0.8 factor is included to ensure that the mean RFF of an image is null when it is exclusively affected by Gaussian noise with constant variance. For a given object to fit, we favor either a single- or two-component model depending on which of them results in the smallest RFF. Generally, studies of galaxy morphology require the RFF relative difference between two models, i.e., \((\text{RFF}_{1} - \text{RFF}_{2})/\text{RFF}_{1}\), to be larger than 1.0–1.6 (e.g., Hoyos et al. 2012). This is a conservative threshold for the selection of the multicomponent fit (RFF\(_2\)) to prevent overfitting. Here we are more concerned about effectively removing the flux of a given galaxy (to get a smaller residual) than providing a realistic description of its morphology. For this reason we select the multicomponent solution when RFF\(_2 < \text{RFF}_1\).

After modeling process and subtracting both ICL and bright cluster members, we apply a median filter to the cleaned image. This is a well-known smoothing technique that replaces the value of a given pixel by the median of its neighboring pixels (see Merlin et al. 2016b). We use a filter with a box size of 1° per side. We apply this filtering only to pixels with a flux within 1σ of the background level in order to reduce the effects of oversubtraction in the residual. The resulting improvement can be seen in the bottom right panel of Figure 6. Note that this process does not affect the outskirts of the cluster.

4. Source Extraction

4.1. HST Detections

After performing the modeling described in the previous section, we homogenize the various HST bands by matching their respective PSFs. Images in a panchromatic data set, especially if they sample a wide wavelength range like extragalactic surveys, are usually affected by different PSF shapes (see Figure 3). This means that the fraction of the flux that falls inside an aperture, as well as the resulting noise, varies as a function of bandwidth and pivotal wavelength.

PSF matching is a convolution procedure to solve these problems by reshaping multiwavelength images into a common reference system having certain resolution and diffraction properties (i.e., the same “target” PSF; see Tomaney & Crotts 1996; Alard & Lupton 1998). It has become a standard technique to measure consistent fluxes over a range of wavelengths (e.g., Galametz et al. 2011) and from different instruments under varied observing conditions (e.g., Laigle et al. 2016) to effectively characterize the spectral energy distribution (SED) of the observed galaxies.

Specifically, we use the PSF models derived in Section 3.1. In our case, the target PSF is the one in the F160W band, as it is the reddest HST band. To convert the other HST images, we use a convolution kernel obtained by taking (in Fourier space) the ratio between their original and target PSFs. Note that in some cases the target PSF is marginally smaller than the original. Despite this, we are able to relate target and original PSF since we do not require

\(^{12}\) https://github.com/MegaMorph/galapagos
the convolution kernel to be positive-definite. We show the growth curves for the PSFs extracted from the HST data, normalized by the growth curve of the F160W PSF before and after homogenization, in Figure 7. We note that after PSF homogenization the normalized growth curves agree, on average, within 2% for all bands except F275W, which agrees at the 8% level. PSF matching is not applied to ground-based images or Spitzer: their resolution is appreciably lower than the HST bands, and in that case the photometric extraction is performed with a different method (see Section 4.2).

We then proceed to perform source photometry for each HST image using SE. The strong clustering of the objects in the HFF images poses the additional problem of detection completeness because the objects that are close pairs along the line of sight are not always successfully deblended by the software. To improve the extraction process, we create a co-added IR image from the weighted mean, using their inverse variance as weights, of four WFC3 bands (namely, F105W, F125W, F140W, and F160W). Such a stacking enhances the signal-to-noise ratio of faint sources, being effectively deeper than any individual band. We use the stack of infrared images as a detection image, running SE in dual mode.

Moreover, given the variety of scientific goals, it is important to precisely measure both the bright galaxies in the foreground and the fainter galaxies in clusters. However, source extraction with SE is a trade-off between rigorous source deblending of galaxies close to each other on the angular plane and spurious shredding of structure from a single galaxy. For this reason, previous works (Caldwell et al. 2008; Gray et al. 2009; Galametz et al. 2013, and references therein) proposed a “dual-mode,” or high dynamic range (HDR), approach with a cold-mode and a hot-mode SE run, where the cold mode detects and performs photometry on the objects contained within visually reliable Kron radii and the hot mode more aggressively deblends and performs photometry on the smallest and faintest objects. Given the aggressive deblending of the hot mode, it can break up larger objects into individual pieces, which we discard when considering the cold catalog. We modify the HDR approach described in Galametz et al. (2013) and discard galaxies from the hot mode that fall within a 0.2 × Kron radius (Kron 1980) of a cold-mode-detected object. We show the resulting magnitude distributions of this process in Figure 8.

To alleviate the possibility of fake source detection around the residuals, we utilize the Kron radius of the modeled object and discard objects that fall within it. Furthermore, we remove sources corresponding to diffraction spikes and those

Figure 6. Summary of the various steps in bright cluster + ICL modeling (in this case for cluster A2744). Top panels show the original image (left) and the galaxy/ICL models (right). Bottom panels show the residual image before and after median filtering (left and right panels, respectively). The color bar denotes the pixel intensity in counts s⁻¹. See Sections 3.2–3.3 for more details.
susceptible to edge effects on the perimeter of the images. This procedure is performed on both the cluster and parallel fields.

We present the magnitude histograms in the F814W band for each cluster and their parallel fields (Figure 9). This shows the overabundance of bright galaxies in the clusters when compared to the parallel fields. We note that even after ICL + bright galaxy subtraction the presence of such bright objects diminishes our ability to detect faint objects. We also show the ratio of number counts between the cluster and parallel fields in the F814W and F160W bands (Figure 10). The observed trend here is consistent across all clusters, demonstrating the importance of having parallel fields with comparable depths.

4.2. Photometry in Lower-resolution Images

In addition to HST, we also rely on imaging observations from Keck, VLT, and Spitzer. Those facilities have lower angular resolutions than the HST instruments, increasing the blending between sources. In order to maximize the information extracted from each image, we use the prior-based code T-PHOT (Merlin et al. 2016b). This code uses the high-resolution HST images and their corresponding SE catalogs as priors to perform dual-mode photometric extraction in the lower-resolution images. This allows measurement of blended sources that are not directly detectable in $K_s$ or IRAC bands. On the other hand, the method misses the so-called “HST dark” sources, i.e., galaxies so faint in optical and near-IR that they are not identified in the prior image, even though they are sufficiently bright to be visible in the low-resolution image. HST dark sources may constitute a hidden population of galaxies providing nonnegligible contribution, e.g., to the cosmic star formation rate density budget (see Wang et al. 2019). However, the identification of HST dark sources is subsequent to the creation of an HST-based catalog and shall be addressed in future work.

Concerning the Spitzer images, before running T-PHOT we perform a series of corrections related to the weight maps (defined as the inverse variance per pixel, i.e., $1/\sigma^2$) using the pull diagnostic (Gross 2018), described as follows. Assuming that the distribution is Gaussian, in principle, the ratio between the variable and its standard deviation should produce a Gaussian with a standard deviation $(\sigma)$ equal to 1. In the limit of large statistics (owing to the central limit theorem), we can approximate the signal on the background pixels to follow a Gaussian distribution. For this purpose, we compute the distribution of the ratio between the background level and its standard deviation, $\sigma_{\text{bkg}}$. Such a distribution is expected to have a standard deviation $\sigma_{\text{pull}} \approx 1$ if $\sigma_{\text{bkg}}$ is properly estimated. A value of $\sigma_{\text{pull}} < 1$ means that the uncertainties are likely underestimated, and $\sigma_{\text{pull}} > 1$ means that the uncertainties are likely overestimated. In order to have conservative values for

Figure 7. Growth curves for each HST-derived PSF normalized by the F160W growth curve before (top) and after (bottom) homogenization.

Figure 8. Magnitude distributions for the objects detected using the cold (blue) and hot (orange) modes for the original image vs. ICL+BCG-subtracted image (top) and ICL-bcg-subtracted image (bottom); matched (black) refers to total number of galaxies detected and kept in the catalog per magnitude bin. We see a significant increase in the number of detected galaxies in the cluster-subtracted image, as well as a slight shift in mean magnitude toward fainter objects.
the error maps, we multiply these by $\sigma_{\text{pull}}$ when the latter is greater than 1. For Spitzer channels 1, 2, 3, and 4 we only correct the error maps in the largest contiguous region where the cluster and parallel fields are located. We check that this is a good approximation since the resulting corrected background distributions show a single peak. The corrected errors, $\sigma_{\text{pull}}$, are discussed in Appendix A.

As the high-resolution prior, we use the science image in the F160W band and the segmentation map created from the weighted IR stack. Previous studies (De Santis et al. 2007; Galametz et al. 2013) found that SE tends to underestimate the isophotal area of objects in a single-band detection image; such a systematic effect is inversely proportional to the flux. A solution to this issue (Galametz et al. 2013; Merlin et al. 2016a) is to dilate the segmentation areas of individual sources. We verify how prominent this effect is by comparing one of the HFF clusters (A370) to the deeper BUFFALO survey, given that the isophotal area increases with increasing signal-to-noise ratio. We choose to compare the HFF IR stack, given its depth, to the BUFFALO F160W band. We find that when using the IR stack the isophotal area of each source is slightly larger, with minimal dependence on the area. The difference is small enough that we found it unnecessary to dilate the segmentation map before feeding it into T-PHOT.

We run T-PHOT for the $K_s$ and IRAC bands. An example of the TPHOT run on the IRAC Channel 1 band of the Abell 2744 cluster is shown in Figure 11. Because the IRAC PSF varies across the field of view, we take advantage of the T-PHOT “multikernel” option, which allows for the inclusion of a separate kernel for each object. These kernels are generated from the grid of PRFs produced by PRFMap, as described in Section 3.1. We emphasize that the output provided by T-PHOT (namely, the parameter FitQty) is an estimate of the total flux emitted by the given source.

### 5. Validation of the Photometric Catalog

#### 5.1. Quality and Completeness via Simulations

To measure the completeness, we add randomly distributed simulated objects in the observed image and then try to recover them. The number of the simulated objects should be high enough to cover the full field of view and avoid statistical fluctuations, while their surface density should be low enough...
not to further increase the already large fraction of overlapping sources in the images. This way, we still check the performance of the deblender without introducing an unrealistic number of new blends.

In order to measure completeness, we injected 100 point-like sources in the original images, with magnitudes randomly assigned from a uniform distribution between $23<\text{mag}_{\text{AB}}<31$. The magnitudes of these objects were selected to be the same in all bands, which allows us to compute the IR-weighted magnitude straightforwardly. This was done to simplify interpretation of the results. The processing of these images was performed using the same steps presented in previous sections. We generated two sets of simulations (hereafter called “flavors”):

1. One set where we assign the position of the injected objects randomly within the image. We will refer to these simulations as flavor 1.
2. A set where we assign the centroid position of the injected objects randomly within the empty parts of the image, i.e., where the segmentation map has no detected pixels. We will refer to these simulations as flavor 2.

The reason behind the generation of two sets of simulations is that in the first flavor the fraction of blended objects is overestimated, since we are randomly placing new sources in a field that is already overcrowded. This leads to a pessimistic characterization of our pipeline’s performance. However, using the second flavor, we explicitly avoid overlap with other sources. In the latter case the quality assessment will be overoptimistic.

Ten independent realizations were generated for each of these flavors. The injected sources were independently generated for each cluster as 15″ side noise-free square cutouts. We then formed a collage of the same size of the original image and added this collage to the original image of each cluster. Therefore, only the original noise in the HFF mosaics is affecting the present test. This is a good approximation for faint sources, where the background dominates over the shot noise due to the source itself.

After the injection of simulated sources, these images are analyzed by the same software used for the real catalog (see previous sections). We compare the output to the list of input (simulated) sources, as well as their intrinsic versus recovered properties. We considered a source successfully recovered if there is an entry in the output catalogs whose centroid is within 0.6″ (10 pixels) from the input location and the offset between intrinsic and extracted magnitude is <0.5 mag. These thresholds have been chosen to minimize the number of by-chance matches between simulated and real sources, allowing at the same time for some difference between the input and output values.

Results from this test are summarized in Figure 12. For flavor 2 simulations, we find that completeness is above 60% for magnitudes brighter than 29 and above 80% for most of the clusters up to magnitude ~28. This confirms the estimated depth of our images and catalogs. When we do not avoid overlap between the objects (i.e., using flavor 1 simulations), the overall completeness decreases noticeably with respect to the simulation where we avoid blending. Thus, we expect the actual completeness of our catalog to be somewhere in between the pessimistic scenario portrayed by our flavor 1 simulations and the optimistic case of our flavor 2 simulations.

5.2. Photometric Uncertainties

Given the complexity of the process to detect objects and measure their flux in different bands, an analytic derivation of their associated uncertainties (in flux, position, and other derived quantities) becomes almost intractable. Forward modeling will enable us to characterize such uncertainties. To this purpose, we inject another set of simulated objects into the original HFF mosaics. This time the simulated sources are realistic galaxy models from the GREAT3 challenge (Miyatake et al. 2014), created by means of GalSim (Rowe et al. 2015). GalSim is a modular, open-source code to render any type of astronomical images. We opted for the GREAT3 library because it includes models with realistic morphology, obtained by fitting deep COSMOS images with multiple Sérsic profiles (see Mandelbaum et al. 2014, 2015). Given that the galaxies from the GREAT3 challenge available in the GalSim catalog only have information for the F140W band, we identify these galaxies in the Laigle et al. (2016) catalog using their ID values to obtain the magnitudes in the other bands used in our study. In this way we have realistic colors for our injected galaxies.

Figure 11. Original IRAC channel 1 image in A2744 (left panel) compared to the residual map (right panel) after subtracting T-PHOT galaxy models (see Section 4.2).
across the entire photometric baseline, with the assumption that their morphology does not vary significantly from band to band. We also ensure that the centroid positions are the same across different bands. We convolve each galaxy profile with the corresponding PSF. The $15 \times 15$ square arcsecond cutout containing the injected objects (and no additional noise) is then added to the original image. We then repeat the photometric extraction process in a similar fashion to what has been done with our data. In particular, we inject 80 objects randomly placed in each image and generate 12 different realizations (images) of those objects in all the HFF bands.

Additionally, we also process the cutouts that only contain the injected objects using the same pipeline; we use the outcomes as “ground truth” to be compared to the catalog resulting from the injection-recovery procedure. To match the two samples, we use the same strategy as in Section 5.1. The photometric residuals for each match are then measured. The results are shown in Figures 13 and 14, indicating an unbiased estimate of input magnitudes across different HST bands included in this study. The positions of these objects were randomly assigned within the footprint (open circles—flavor 1, i.e., not blended) or randomly assigned in empty parts of the segmentation map (crosses—flavor 2). We introduce a small horizontal offset of the different markers and only include the error bars for the flavor 2 case in order to improve readability. We also add a horizontal solid line to mark 100% completeness and two dashed lines at 80% and 60% completeness as visual help. Detections are made in IR-weighted images.

bands, due to the fact that we are injecting galaxies with realistic SEDs from COSMOS. Additionally, we see a different behavior when we compare different bands. In particular, the uncertainties for the recovered fluxes in the F160W band seem to be smaller than in most of the other bands. The reason behind these differences is the larger depth of the images in F160W compared to the other bands. It is particularly interesting to notice larger uncertainties derived for F275W compared to the other HST bands. This is a combination of two factors: first, the signal-to-noise ratio for F275W is the lowest of all HST bands; second, this is also the bluest band, where we expect the largest morphological differences with respect to our segmentation map.

We only perform the simulations for the cluster fields. The reason is that we expect these fields to outperform the parallel fields, as the latter are not affected by the ICL and are less crowded. As a consequence, our measurements for the clusters will set an upper limit for the photometric biases and uncertainties in the parallel fields. As shown in Figure 15, the photometric bias appears to be uncorrelated with position when comparing cluster core offsets with those in the outer regions. Since the source extraction steps in both cluster fields and parallel fields are the same, the validation tests of the photometric measurements in the cluster fields are also applicable on the parallel fields.

Our forward-modeling approach allows us to estimate the accuracy of the flux errors as reported by $\sigma_{\text{SE}}$. Using the injected galaxies, we compute the forward-modeling estimation of the uncertainty in the flux, which we call $\delta F_{\text{FM}}$, as the rms of the difference between the measured flux for the injected sources and the input flux, $\delta F_{\text{FM}} = \sqrt{\langle (F_{\text{meas}} - F_{\text{input}})^2 \rangle}$, where $F_{\text{meas}}$ is the measured flux and $F_{\text{input}}$ is the injected flux. This estimation is performed in linearly spaced magnitude bins (with a bin width of 1.6 mag). We then compare $\delta F_{\text{FM}}$ to the mean uncertainties reported by $\sigma_{\text{SE}}$ in each of the bins, i.e., $\langle \delta F_{\text{SE}} \rangle$. This is done by studying the ratio $r_{\text{corr}} = \frac{\delta F_{\text{FM}}}{\delta F_{\text{SE}}}$ as a function of flux. We fit the resulting ratio to an exponential model $r_{\text{corr}} = A F^K$ and use this fit to correct the reported uncertainties for each detected object. The ratio and resulting fit are shown in

Figure 12. Completeness as a function of magnitude for injected point-like sources between 24 and 30 mag in the six analyzed clusters (top row: A2744, A370, AS 1063; bottom row: MACS-1149, MACS-0717, MACS-0416). The positions of these objects were randomly assigned within the footprint (open circles—flavor 1, i.e., not blended) or randomly assigned in empty parts of the segmentation map (crosses—flavor 2). We introduce a small horizontal offset of the different markers and only include the error bars for the flavor 2 case in order to improve readability. We also add a horizontal solid line to mark 100% completeness and two dashed lines at 80% and 60% completeness as visual help. Detections are made in IR-weighted images.
Figure 16 for F160W, $K_s$, and IRAC channel 1. The best-fit values for $a$ and $b$ can be found in Table 4. By binning the data, we lose information about the effects of blending and correlations between flux uncertainties for individual galaxies, but we get a more robust estimation of the overall uncertainty of the ensemble.

5.3. Comparison with Previous Work

There are other independent studies of the Frontier Fields. Given the complexities of extracting photometry in these deep and crowded fields, it poses a great challenge to determine photometric properties of the objects detected with the desired accuracy. Therefore, we believe that there is merit in both exploring the limits of our photometric pipeline and comparing with previous results. Figure 17 shows how our photometry compares with that of the previous teams, ASTRODEEP (Merlin et al. 2016a; Castellano et al. 2016; Di Criscienzo et al. 2017; Bradač et al. 2019) and DeepSpace (Shipley et al. 2018), for intersecting bands. For the HST bands, we note that there is general agreement in the photometry between the three
methods within 0.05 mag up to $m_{AB} \sim 23$ mag. Beyond this magnitude, our photometry agrees best with the ASTRODEEP team, which could be the result of our reduction procedure being more similar to theirs. Furthermore, we see that the three data sets are statistically compatible up to magnitude $\sim 25.5$ in most bands. However, in the fainter end, there are statistically significant differences between the DeepSpace and ASTRODEEP data sets and the data presented in this work. These differences are probably due to the difference in modeling, as the measure of fluxes/magnitudes depends on the concrete choices for apertures, etc. However, we check that the colors are consistent among these data sets for the HST and Ks bands, which can be inferred by the consistent behavior found in most bands. Thus, for applications where colors are most important (e.g., photometric redshifts), any of these data sets should give similar results (as evidenced by our photometric redshifts).

6. Photometric Redshifts

Having the self-consistent SEDs for individual galaxies in each of the HFF clusters and their associated parallel fields, we now measure the photometric redshifts using LePhare (Arnouts et al. 1999; Ilbert et al. 2006). The code fits galaxy templates to the observed SEDs to derive a redshift likelihood function ($L$) for each source. Photometric redshift estimates ($z_{phot}$) are then defined as the median of $L(z)$. The template library adopted here is the same as that used in Laigle et al. (2016) since it was shown that it works efficiently across a wide redshift range (up to $z \sim 6$). The library is built by spiral and elliptical galaxy models from Polletta et al. (2007), along
with other ones derived from the stellar population synthesis model of Bruzual & Charlot (2003). The latter is used to reproduce both young quiescent and starburst galaxies. For each model we produce different templates by changing their dust attenuation between $0 < E(B-V) < 0.5$ and assuming either Prevot et al. (1984) or Calzetti et al. (2000) extinction laws depending on the galaxy type (see Laigle et al. 2016, for more details). When Calzetti et al. (2000) is chosen, we produce three alternate versions by adding different parameterization of the 2175 Å bump (see Ilbert et al. 2009). Absorption by the intervening intergalactic medium is also implemented as prescribed in Madau (1995). We also add the main nebular emission lines in rest-frame optical to the templates corresponding to the star-forming SEDs, modeling line fluxes and their ratios as in Saito et al. (2020).

Before running LePhare, a modification is required to the HST bands: the fraction of the flux lost by SE isophotal measurements (FLUX_ISO) must be taken into account. FLUX_ISO provides more accurate colors than FLUX_AUTO, which is another SE photometric measurement specifically designed to recover total flux by means of an adaptive aperture (see Kron 1980); therefore, the former are preferred for $z_{\text{phot}}$ computation. However, our ancillary photometry is extracted with T-PHOT, which does not provide an equivalent to FLUX_ISO. Therefore, we include in our baseline $K_s$ and IRAC total fluxes, while HST bands fluxes are rescaled by a factor

$$f_{\text{tot}} = \frac{\sum_i w_i (\text{FLUX}_{\text{AUTO}}/\text{FLUX}_{\text{ISO}})_i}{\sum_i w_i},$$

Figure 17. Comparison of our extracted photometry with that of the existing data produced by the ASTRODEEP and DEEPSPACE teams. Shown are the $\Delta$mag where data exist for cluster-only photometry in each band. The black dashed lines represent magnitude offsets of $\pm0.05$. The solid lines represent the running median, and the shaded region represents the interquartile range of those measurements.
i.e., the weighted mean of the AUTO-to-ISO flux ratio summed over the observed HST bands. Weights are defined as
\[ w = \sqrt{\sigma_{\text{AUTO}}^2 + \sigma_{\text{ISO}}^2}, \]
\[ \text{i.e., square root of the sum in quadrature of the corresponding SE errors.} \]

In order to empirically correct for systematic effects, we perform a calibration that relies on the spectroscopic redshifts \( z_{\text{spec}} \) available in each field. The sample includes spectroscopic data from several programs retrieved from the NASA/IPAC Extragalactic Database.\(^{13}\) The spectroscopic campaigns contributing to the majority of the data set are described in Owers et al. (2011), Ebeling et al. (2014), Richard et al. (2014), Balestra et al. (2016), Treu et al. (2015), Schmidt et al. (2014), Lagattuta et al. (2017), and Mahler et al. (2018). We also include more recent observations carried out by Lagattuta et al. (2019) in A370, which were not included in previous photometric redshift analysis in the literature. The calibration procedure is as follows. First, we run LePhare on the HFF galaxies with known spectroscopic redshifts after fixing their redshifts to their spectroscopic \( z_{\text{spec}} \) value. We only consider galaxies with a reliable spectroscopic redshift, i.e., with a “quality flag” < 3.\(^{14}\) We do not apply any magnitude cut. Once we obtain the best-fit solution from LePhare (i.e., the fitting model resulting in the smallest \( \chi^2 \)), we measure in each photometric band the difference between the observed flux and the prediction by the best-fit template. To find the systematic offset (in a given band), we compute the average difference in the spectroscopic sample. Since results are similar in the six fields, we eventually consider only the offsets measured for the A370 cluster because it has the largest number of spectroscopic sources (320 of them reliable spectroscopic redshifts). These offsets, when applied to the photometric baseline, will compensate for a possible bias in the template library and/or for calibration issues in data reduction. In the HST bands the corrections are between 2% and 5%, except for the F425W filter, which is 8%; a similar value is found for \( K_s \), while in the IRAC channels 1 and 2 the correction is a factor of 1.11. We note that similar offsets (namely, 0.1 mag in both IRAC channels 1 and 2) have been found by LePhare also in another extragalactic survey (Mehta et al. 2018). All offsets are quoted in Table 5, with the exception of IRAC channels 3 and 4 since the systematics here would be mainly driven by the absence of dust reemission in the templates (see Mehta et al. 2018). To take such a limitation into account, we increased the error bars in these two channels by adding 0.5 mag in quadrature, underweighing in practice their contribution to the fit.

We fit the SED of 1423 spectroscopic galaxies with \( 16 < z_{\text{phot}} < 26 \) across the six clusters to assess \( z_{\text{phot}} \) quality. The calibration offsets (Table 5) are taken into account. Similar to the previous work (e.g., Brammer et al. 2008), the scatter is defined as the normalized median absolute deviation (NMAD; Hoaglin et al. 1983), i.e.,
\[ \sigma_{\text{NMAD}} = 1.48 \times \text{median} \left\{ \frac{|\Delta z - \text{median}(\Delta z)|}{1 + z_{\text{spec}}} \right\}, \] (4)

13 The NASA/IPAC Extragalactic Database (NED) is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology (http://ned.ipac.caltech.edu).

14 It is common practice that the persons responsible for the \( z_{\text{spec}} \) measurements assign a quality flag, ranging from 1 to 4, following the scheme proposed for the first time in Le Fèvre et al. (2004), where a flag equal to 4 corresponds to the highest confidence level.

Figure 18. Assessing the quality of photometric redshifts estimated through SED fitting. Top panel: \( z_{\text{phot}} \) vs. \( z_{\text{spec}} \) comparison. Red circles are 1423 spectroscopic redshifts with \( 16 < F160W < 26 \), the solid line shows the 1:1 relationship, and the dashed lines enclose the \( z_{\text{phot}} = z_{\text{spec}} \pm 0.15(1 + z_{\text{spec}}) \) threshold used to identify outliers (i.e., catastrophic errors). NMAD scatter \( \sigma \) and outlier fraction \( \eta \) are reported in the upper left corner. Bottom panel: \( \Delta z = z_{\text{phot}} - z_{\text{spec}} \) scatter (red circles are spectroscopic objects), with the median bias indicated by a solid line; dashed lines represent the threshold for catastrophic errors as in the top panel. See Figure 19 for detailed performance in each individual cluster.

Table 5

| Band       | Multiplicative Factor |
|------------|-----------------------|
| F275W      | 1.04                  |
| F336W      | 1.02                  |
| F435W      | 1.07                  |
| F606W      | 0.99                  |
| F814W      | 0.96                  |
| F105W      | 0.98                  |
| F125W      | 1.01                  |
| F140W      | 1.02                  |
| F160W      | 1.01                  |
| K_s        | 1.00                  |
| I1         | 1.11                  |
| I2         | 1.11                  |

Note. The offsets found in the HST bands are consistent with the median magnitude residual values found in Figure 13, showing the robustness of our forward-modeling procedure.

with \( \Delta z = z_{\text{phot}} - z_{\text{spec}} \). The incidence of catastrophic errors \( \eta \) is defined as the fraction of redshift outliers having \( |\Delta z| > 0.15(1 + z_{\text{spec}}) \). For pathological PDF(z), where the median is not a good approximation of the main peak of the PDF, we used the redshift corresponding to the minimum \( \chi^2 \) solution (as in Davidson et al. 2017). For the entire sample \( \sigma = 0.067 \) and \( \eta = 10.3\% \) (Figure 18, top panel). After excluding the outliers, we also
recompute the mean of $\Delta z$ to assess the so-called redshift bias (e.g., Masters et al. 2017), which is 0.012 in our case. We repeat the same procedure for the parallel fields, in which a total of only 62 spectroscopic redshifts are available. In that case we estimate $\sigma = 0.044$ with no outliers. Individual cluster field redshift comparisons are shown in Figure 19. In this figure, the comparison is made using the highest-quality spectroscopic redshifts. Figure 20 also shows the photometric redshift distribution for objects in each cluster SED fit with a reduced $\chi^2 < 10$.

7. Demagnification

Leveraging the fact that these HFF clusters exhibit some of the strongest lensing currently observed, we calculate and include lens model magnification factors for all relevant objects in our final catalog (i.e., for those objects with redshifts larger than the clusters’ mean redshifts). We use the lens models provided by Bradač (Hoag et al. 2016, and references therein), Caminha (Caminha et al. 2017), CATS (Jauzac et al. 2014, and references therein), DIEGO (Diego et al. 2015, and references therein),
crowded field with bright cluster members and ICL that could bias the measurements. We have successfully removed the contamination from these sources to further reduce the image into a “blank field” on which we can detect the faintest objects in the field without biasing the flux measurements of objects on the perimeter of the field of view. We perform an error analysis using GALSIM by injecting COSMOS-like galaxies to both estimate the errors and analyze the validity of our pipeline. We note that there is no evident photometric bias close to the cluster core as compared to the outskirts.

We estimate the photometric redshifts for objects in all six clusters and their parallel fields as measured by LePhare. When compared to the available spectroscopic redshifts for the HFF clusters, this gives a combined outlier fraction of 10.3% and a redshift bias of 0.012 after excluding the outliers.

We find some differences in our photometry measurements with those from the literature and note the importance of cross-checks between data sets in such crowded and complex environments. We demonstrate the utility of performing source injection, where knowledge of ground truth is accessible to calibrate measurements and characterize biases and uncertainties.

We have scripted a pipeline to analyze each of the Frontier Fields in an efficient, streamlined, and reproducible manner. We plan to apply a similar version of the pipeline to the next-generation survey of the Frontier Fields, namely, the BUFFALO survey (Steinhardt et al. 2020), which expands these same fields in area by a factor of 3–4 and pushes the 5σ depth ∼1 mag fainter.

A.P. would like to thank A. Alavi, B. Häußler, E. Merlin, and M. Nowinski for useful discussion and comments; A.P. would also like to thank A. Faisst for providing the code to generate Spitzer PRFs that are used in this work. A.P. acknowledges support from the NASA MUREP Institutional Opportunity (MIRO) through grant NNX15AP99A. F.J.S. acknowledges that this document was prepared using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under contract No. DE-AC02-07CH11359. I.D. acknowledges the support received from the European Union’s Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 896225. This work utilizes gravitational lensing models produced by PIs Bradač Natarajan & Kneib (CATS), Merten & Zitrin, Sharon, Williams, Keeton, Bernstein and Diego, and the GLAFIC group. This lens modeling was partially funded by the HST Frontier Fields program conducted by STScI. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. The lens models were obtained from the Mikulski Archive for Space Telescopes (MAST). This work has made use of the CANDIDE Cluster at the Institut d’Astrophysique de Paris and was made possible by grants from the PNCG and the DIM-ACAV. The Cosmic Dawn Center is funded by the Danish National Research Foundation under grant No. 140.

### Appendix A

#### Error Correction

In this section we show correction factors applied at various stages in the pipeline. The values for the error map correction from pull-plots calculated in Section 4.2 are shown in Table 6.
The procedure to obtain these correction factors is described in Section 4.2. The different columns refer to different IRAC channels.

### Appendix B

#### Catalog Columns

Each cluster’s photometric catalog has the following columns, with the descriptions taken from each software’s respective documentation (Bertin & Arnouts 1996; Arnouts et al. 1999; Merlin et al. 2016b):

1. **IDENT:** unique galaxy identifier
2. **ALPHA_J2000:** J2000 R.A. of the isophotal image centroid
3. **DELTA_J2000:** J2000 decl. of the isophotal image centroid
4. **FLUX_TOT_BAND:** weighted mean of the AUTO-to-ISO flux ratio
5. **FLUXERR_TOT_BAND:** rms error estimate for the weighted mean of the AUTO-to-ISO flux ratio
6. **FLUX_AUTO_BAND:** Kron-like automated aperture flux
7. **FLUXERR_AUTO_BAND:** rms error estimate for Kron-like automated aperture flux
8. **FLUX_ISO_BAND:** isophotal flux
9. **FLUXERR_ISO_BAND:** rms error estimate for isophotal flux
10. **CLASS_STAR:** star/galaxy classifier
11. **FLAGS:** source extraction flags
12. **MAGNIFICATION_{MODEL}:** magnification factors, \( \mu \), for each lens model. Demagnified fluxes for a given model, \( \text{MODEL} \), can be obtained by dividing the preferred flux estimate (e.g., ISO, AUTO) by the value given by MAGNIFICATION_{MODEL}.

There are two photometric redshift catalogs: one where the redshifts were calculated with all bands presented in this work, and one where only HST bands are used. The column order is as follows:

1. **IDENT:** unique galaxy identifier
2. **Z_BEST:** Zphot Best
3. **Z_BEST68_LOW:** Zphot min from \( \Delta \chi^2 = 1.0(68\%) \)
4. **Z_BEST68_HIGH:** Zphot max from \( \Delta \chi^2 = 1.0(68\%) \)
5. **Z_ML:** Zphot from Median of ML distribution
6. **Z_ML68_LOW:** Zphot min at 68% of ML distribution
7. **Z_ML68_HIGH:** Zphot max at 68% of ML distribution
8. **CHI_BEST:** lowest galaxy \( \chi^2 \) for galaxy
9. **MOD_BEST:** galaxy model for best \( \chi^2 \)
10. **EXTLAW_BEST:** extinction law
11. **EBV_BEST:** \( E(B-V) \)
12. **AGE_BEST:** age (yr)
13. **PDZ_BEST:** probability per Z bins
14. **Z_SEC:** galaxy secondary Zphot solution from \( F(z) \) function
15. **CHI_SEC:** \( \chi^2 \) evaluated at secondary Zphot from \( F(z) \) function
16. **MOD_SEC:** galaxy model for CHI_SEC
17. **EBV_SEC:** \( E(B-V) \)
18. **PDZ_SEC:** probability per Z bins for the secondary model
19. **MOD_STAR:** star solution
20. **CHI_STAR:** \( \chi^2 \) for star solution
21. **SCALE_BEST:** scaling factor
22. **NBAND_USED:** number of bands used for each object
23. **CONTEXT:** bands to be considered in fit
24. **ZS_SPEC:** spectroscopic redshift if available
25. **MAG_OBS_N:** observed magnitude of band \( N \)
26. **MAG_MOD_N:** predicted model magnitude
27. **AGE_MED:** age (yr)
28. **MASS_MED:** rescaled mass \( (M_{\odot}) \)
29. **SFR_MED:** rescaled SFR \( (M_{\odot}/yr) \)
30. **ZSPEC_FLAG:** quality flag for Zspec
31. **ZSPEC_REF:** reference for Zspec
32. **FIELD:** cluster name
33. **RA:** J2000 R.A. of the isophotal image centroid
34. **DEC:** J2000 decl. of the isophotal image centroid

#### ORCID iDs

A. Pagul at https://orcid.org/0000-0002-6015-8614
F. J. Sánchez at https://orcid.org/0000-0003-3136-9532
I. Davidzon at https://orcid.org/0000-0002-2951-7519

#### References

Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325
Alavi, A., Siana, B., Richard, J., et al. 2016, ApJ, 832, 56
Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540
Atek, H., Richard, J., Kneib, J.-P., & Schaerer, D. 2018, MNRAS, 479, 5184
Balestra, I., Mercorio, A., Sartoris, B., et al. 2016, ApJS, 224, 33
Barden, M., Häußler, B., Peng, C. Y., McIntosh, D. H., & Guo, Y. 2012, MNRAS, 422, 449
Barkana, R., & Loeb, A. 2001, PhR, 349, 125
Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blandford, R., & Narayan, R. 1986, ApJ, 310, 568
Bradač, M., Huang, K.-H., Fontana, A., et al. 2019, MNRAS, 489, 99
Brammer, G. B., Marchesini, D., Labbé, I., et al. 2016, ApJS, 226, 6
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Caldwell, J. A. R., McIntosh, D. H., Rix, H.-W., et al. 2008, ApJS, 174, 136
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Caminha, G. B., Grillo, C., Rosati, P., et al. 2017, A&A, 600, A90
Castellano, M., Amorín, R., Merlin, E., et al. 2016, A&A, 590, A31
Davidzon, I., Ilbert, O., Laglie, C., et al. 2017, A&A, 605, A70
De Santis, C., Grazian, A., Fontana, A., & Santini, P. 2007, NewA, 12, 271
Di Criscienzo, M., Merlin, E., Castellano, M., et al. 2017, A&A, 607, A30
Diego, J. M., Broadhurst, T., Benítez, N., et al. 2015, MNRAS, 446, 683
Dolaiti, E., Bendinelli, O., Bonacini, D., et al. 2000, Proc. SPIE, 4007, 879
Ebeling, H., Ma, C.-J., & Barrett, E. 2014, ApJS, 211, 21
Galametz, M., Madden, S. C., Galliano, F., et al. 2011, A&A, 532, 56
Häußler, B., Bamford, S. P., Vika, M., et al. 2013, MNRAS, 430, 449
Hoyos, C., Aragón-Salamanca, A., Gray, M. E., et al. 2012, MNRAS, 419, 2703
Gonzaga, S., Hack, W., Fruchter, A., & Mack, J. 2012, The DrizzlePac Handbook (Baltimore, MD: Space Telescope Science Institute)
Gray, M. E., Wolf, C., Barden, M., et al. 2009, MNRAS, 393, 1275
Groth, E. 2018, in Proc. 2017 European School of High-Energy Physics, ed. M. Mulders & G. Zanderighi (Geneva: CERN), 199
Häußler, B., Bamford, S. P., Vika, M., et al. 2013, MNRAS, 430, 330
Hoag, A., Huang, K. H., Treu, T., et al. 2016, ApJ, 831, 182
Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding Robust and Exploratory Data Analysis (New York: Wiley)
Hoyos, C., Aragón-Salamanca, A., Gray, M. E., et al. 2012, MNRAS, 419, 2703

---

**Table 6**

| Cluster | \( I \) | \( I_2 \) | \( I_3 \) | \( I_4 \) |
|--------|--------|--------|--------|--------|
| A370   | 3.12   | 2.68   | 1.03   | 1.02   |
| MACS J0717.5+3745 | 2.98   | 2.74   | 0.90   | 0.83   |
| MACS J0416.1−2403 | 3.04   | 2.88   | 0.91   | 0.86   |
| AS 1063 | 2.96   | 2.73   | …     | …     |
| A2744  | 5.83   | 4.44   | 0.88   | 0.85   |
| MACS J1149.5+2223 | 2.82   | 2.56   | …     | …     |

Note. The procedure to obtain these correction factors is described in Section 4.2. The different columns refer to different IRAC channels.
