

# FORECAST: A flexible software to forward model cosmological hydrodynamical simulations mimicking real observations

Flaminia Fortuni<sup>1</sup>, Emiliano Merlin<sup>1</sup>, Adriano Fontana<sup>1</sup>, Carlo Giocoli<sup>3,4</sup>, Erik Romelli<sup>2</sup>, Luca Graziani<sup>5</sup>, Paola Santini<sup>1</sup>, Marco Castellano<sup>1</sup>, Stéphane Charlot<sup>6</sup>, and Jacopo Chevallard<sup>7</sup>

<sup>1</sup> INAF – Osservatorio Astronomico di Roma, via Frascati 33, 00078 Monte Porzio Catone (Roma), Italy  
e-mail: [flaminia.fortuni@inaf.it](mailto:flaminia.fortuni@inaf.it)

<sup>2</sup> INAF – Osservatorio Astronomico di Trieste, Via Tiepolo 11, 34131 Trieste, Italy

<sup>3</sup> INFN – Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy

<sup>4</sup> INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Gobetti 93/3, 40129 Bologna, Italy

<sup>5</sup> Dipartimento di Fisica, Università di Roma “La Sapienza”, Piazzale Aldo Moro 5, 00185 Roma, Italy

<sup>6</sup> Sorbonne Université, CNRS, UMR 7095, Institut d’Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France

<sup>7</sup> Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

Received 21 April 2023 / Accepted 18 July 2023

## ABSTRACT

**Context.** Comparing theoretical predictions to real data is crucial to properly formulate galaxy formation theories. However, this is usually done naively considering the direct output of simulations and quantities inferred from observations, which can lead to severe inconsistencies.

**Aims.** We present FORECAST, a new flexible and adaptable software package that performs forward modeling of the output of any cosmological hydrodynamical simulations to create a wide range of realistic synthetic astronomical images, and thus providing a robust foundation for accurate comparison with observational data. With customizable options for filters, field-of-view size, and survey parameters, it allows users to tailor the synthetic images to their specific requirements.

**Methods.** FORECAST constructs a light cone centered on the observer’s position exploiting the output snapshots of a simulation and computes the observed flux of each simulated stellar element, modeled as a single stellar population, in any chosen set of passband filters, including  $k$  correction, intergalactic medium absorption, and dust attenuation. These fluxes are then used to create an image on a grid of pixels, to which observational features such as background noise and PSF blurring can be added. This allows simulated galaxies to be obtained with realistic morphologies and star formation histories.

**Results.** As a first application, we present a set of images obtained exploiting the ILLUSTRISTNG simulation, emulating the GOODS-South field as observed for the CANDELS survey. We produced images of  $\sim 200$  sq. arcmin, in 13 bands (eight *Hubble* Space Telescope optical and near-infrared bands from ACS *B435* to WFC3 *H160*, the VLT HAWK-I *Ks* band, and the four IRAC filters from *Spitzer*), with depths consistent with the real data. We analyzed the images with the same processing pipeline adopted for real data in CANDELS and ASTRODEEP publications, and we compared the results against both the input data used to create the images and the real data, generally finding good agreement with both, with some interesting exceptions which we discuss. As part of this work, we have released the FORECAST code and two datasets. The first is the CANDELS dataset analyzed in this study, and the second dataset emulates the JWST CEERS survey images in ten filters (eight NIRC*am* and two MIRI) in a field of view of 200 sq. arcmin between  $z = 0-20$ .

**Conclusions.** FORECAST is a flexible tool: it creates images that can then be processed and analyzed using standard photometric algorithms, allowing for a consistent comparison among observations and models, and for a direct estimation of the biases introduced by such techniques.

**Key words.** virtual observatory tools – galaxies: evolution

## 1. Introduction

In the past two decades, several imaging and spectroscopic surveys have revolutionized our understanding of galaxies across the electromagnetic spectrum (e.g., Colless 1999; Abazajian et al. 2003; Giavalisco et al. 2004; Lilly et al. 2007; Scoville et al. 2007; Driver et al. 2009; Grogin et al. 2011; Koekemoer et al. 2011; Brammer et al. 2012; Tomczak et al. 2014; Pentericci et al. 2018). These surveys have provided observations of thousands of galaxies, enabling their systematic study and classification at different epochs. Advancements in technological capabilities are pushing the boundaries of space exploration, allowing us to observe the Universe farther in space and further in time, reaching the dawn of the first lights. JWST is providing

exquisite data on the early stages of galaxy evolution, yielding unprecedented results that are challenging our understanding of galaxy formation and evolution (e.g., Morishita & Stiavelli 2023; Treu et al. 2022; Roberts-Borsani et al. 2022; Castellano et al. 2022; Naidu et al. 2022; Finkelstein et al. 2023; Curtis-Lake et al. 2023; Robertson et al. 2023; Yan et al. 2023; Donnan et al. 2023; Harikane et al. 2023). The future ahead holds even greater excitement as it introduces a new generation of telescopes, including *Euclid*, the *Nancy Grace Roman* Telescope, the European-Extremely Large Telescope, and the *Vera Rubin* Observatory.

High-quality data must be compared to precise theoretical predictions. Cosmological simulations, encompassing a wide range of approaches, have achieved a remarkable level of

sophistication, producing detailed characterization of the Universe across an extensive range of spatial and temporal scales. Hydrodynamical simulations, in particular, self-consistently simulate the evolution of both dark matter and baryons, providing insights into the complex nonlinear processes involved in the growth of cosmic structures, including the formation of galaxies, the interplay between gas dynamics and gravitational forces, and the emergence of large-scale cosmic filaments. Their successful reproduction of observable properties and scaling relations of real galaxies (e.g., Hernquist et al. 1996; Choi et al. 2010; Devriendt et al. 2010; Park et al. 2012; Genel et al. 2014; Pillepich et al. 2018a; Kaviraj et al. 2017; Nelson et al. 2018; Vogelsberger et al. 2018; Cui et al. 2021; Di Cesare et al. 2023) establishes them as effective guidance for interpreting observational data.

However, comparing simulation predictions to observed data requires establishing a coherent linkage between the physical and the observable domains. This can be achieved (i) by moving from the observational to the physical quantities, using the features of real imaging data to estimate a set of underlying physical parameters or models (indirect approach), or (ii) by going in the opposite direction, reproducing and mimicking observations from theory (forward approach). Converting photometric or spectroscopic data into physical quantities using the indirect approach is a common practice in astronomical research. When only imaging data are available, which is often the case for large-scale and/or high-redshift surveys, the physical properties of the sources can only be estimated by exploiting multiwavelength photometry and spectral energy distribution (SED) fitting techniques. These methods involve assumptions that can introduce biases in the estimated physical properties. These assumptions include the choice of an initial mass function (IMF), a stellar population synthesis (SPS) model, and (usually) simple parametric star formation histories. Dust attenuation is modeled as a function of the color excess parameter  $E(B - V)$ , scaling with the dust column density, and with the interstellar medium (ISM) opacity  $k(\lambda)$  which is related to the properties of dust grains (e.g., Calzetti et al. 1994, 2000). Moreover, fluxes provided by synthetic SED models ought to be corrected for absorption processes in the interstellar medium and intergalactic medium (IGM), which are both wavelength dependent (and the latter is also redshift dependent). These assumptions make the fitting model prone to biases due to the simplifications with respect to the complexity encoded in a real SED (see e.g., Marchesini et al. 2009; Mobasher et al. 2015).

To address these challenges, we present FORECAST, a tool for forward modeling cosmological hydrodynamical simulations into mock observed images between rest-frame ultraviolet and near-infrared bands. Unlike existing tools that are primarily galaxy-based (e.g., Behroozi et al. 2020; Drakos et al. 2022; Snyder et al. 2023), FORECAST adopts a particle-based approach, translating the physical properties of individual resolution elements (particles or cells) into observed fluxes. This approach enables the creation of simulated images with realistic galaxy morphologies, interactions, and star formation histories, improving upon standard image simulation software tools, which typically adopt analytical functional forms to render galactic light profiles; readers can refer to GALSIM (Rowe et al. 2015), SKYMAKER (Bertin 2009), and SKYLENS (Plazas et al. 2019), for example.

The use of forward modeling techniques based on numerical simulations is a well-established practice in the literature. They have been employed to assess the reliability of photometric methods (Price et al. 2017; Parsotan et al. 2021) and to evaluate the performance of SED fitting (Laigle et al. 2019), including

the utilization of fully Bayesian inference fitting codes for reconstructing nonparametric star formation histories (Ji & Giavalisco 2022). While many studies often rely on the implementation of phenomenological prescriptions or semi-analytical models to construct mock catalogs (Blaizot et al. 2005; Kitzbichler & White 2007; Merson et al. 2013; Bravo et al. 2020; Behroozi et al. 2020; Somerville et al. 2021; Drakos et al. 2022; Yung et al. 2023), there is a growing interest on utilizing hydrodynamical simulations to create mock observations with specific scientific purposes (Snyder et al. 2023; Cochrane et al. 2023; Barrientos Acevedo et al. 2023). When the sample is derived from hydrodynamical simulations, it often comprises a small number of galaxies (Guidi et al. 2016; Price et al. 2017; Parsotan et al. 2021), or larger samples restricted to a specific redshift range, aimed at simulating observations from specific instruments (Snyder et al. 2017, 2023; Laigle et al. 2019; Nanni et al. 2023). We point out that while some effort has already gone also into building tools that produce images from empirical or semi-analytical models (Overzier et al. 2013; Taghizadeh-Popp et al. 2015; Bernyk et al. 2016), and some final products also publicly available (Behroozi et al. 2020)<sup>1</sup>, the strength of FORECAST stems from its inherently flexible and adaptable framework, specifically designed to emulate real observations and replicate comprehensive photometric surveys by leveraging the predictions of any hydrodynamical cosmological simulation as input.

The mock images created with FORECAST can be processed and analyzed as real images. As a first application, in this paper we test FORECAST by forward modeling the ILLUSTRISTNG100 simulation (Weinberger et al. 2018; Pillepich et al. 2018b; Nelson et al. 2019) and creating a dataset that mimics the observational properties of the GOODS-South field, as observed by the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011), and using CANDELS (Guo et al. 2013) and ASTRODEEP-GS43 catalog (Merlin et al. 2021) for our comparisons. This simulated dataset is publicly available<sup>2</sup>, together with a JWST CEERS-like dataset and the FORECAST code.

The paper is organized as follows. In Sect. 2, we describe the methods implemented in FORECAST to forward model the simulated data. The synthetic dataset produced to test the code is described in Sect. 3. The results of the photometric analysis of our synthetic images are discussed in Sect. 3.2. In Sect. 4, we present the public release of the FORECAST code along with an additional dataset of synthetic images emulating the JWST CEERS survey. Finally, in Sect. 5 we summarize the main points of our work and discuss possible future work.

All the magnitudes are defined in the AB magnitude system (Oke 1974), with fluxes in units of  $\mu\text{Jy}$ , namely  $m_{\text{AB}} = -2.5 \log_{10}(f_{\nu}) + 23.9$ . We adopt the flat  $\Lambda\text{CDM}$  Cosmology constrained by Planck Collaboration XIII (2016), with  $\Omega_{\Lambda,0} = 0.6911$ ,  $\Omega_{m,0} = 0.3089$ ,  $\Omega_{b,0} = 0.0486$ ,  $\sigma_8 = 0.8159$ ,  $n_s = 0.9667$ ; and we use the Hubble constant in terms of  $h \equiv H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1} = 0.6774$ .

## 2. Description of the software

In this section, we provide a description of the algorithms included in FORECAST to build the synthetic images (Sect. 2.1) and to add realistic observational features (Sect. 2.2).

<sup>1</sup> <https://www.peterbehroozi.com/data.html>

<sup>2</sup> <http://www.astrodeep.eu/FORECAST>

**Table 1.** FORECAST input parameters.

Parameter	Units	Description
$L_{\text{box}}$	cMpc h <sup>-1</sup>	Side-length of simulation box
filters	–	List of $N$ photometric filters (whether provided by the software or by the user)
$L_{\text{fov}}$	deg	Side-length of the field of view
$N_{\text{pix}}$	–	Number of pixels per side-length of the mock image
$z_s$	–	Highest redshift in the light cone
SED resolution	–	Chosen resolution for SED in stellar population synthesis code (“lr” for low resolution, “hr” for high resolution)
IMF	–	Chosen IMF for SED in stellar population synthesis code (“chab” for <a href="#">Chabrier 2003</a> or “salp” for <a href="#">Salpeter 1955</a> )

### 2.1. The mock observatory

FORECAST uses a particle-based approach to reconstruct the observable photometric properties of galaxies within the field of view. The reconstruction is based on the properties of individual resolution elements in the simulation, represented by stellar particles. Each stellar particle corresponds to a single stellar population (SSP) and collectively forms the simulated objects, representing galaxies in the field of view.

In short, FORECAST reads the physical properties of the stellar particles from the output snapshots of a chosen simulation and translates them into observable quantities as follows. The flux of each particle is computed considering its rest-frame SEDs; then, the SED is  $k$ -corrected consistently with the redshift of the particle. The SED is convolved with chosen passband filter to obtain the theoretical observed flux in that band (Sect. 2.1.3). Gas elements (either particles or cells, depending on the simulation) are used as tracers for dust, which attenuates stellar particle fluxes in the blue and visible range (Sect. 2.1.4). The software does not implement the effects of dust emission.

Finally, the comoving three-dimensional coordinates of each SSP are first projected onto the two-dimensional field of view of an observed light cone (Sect. 2.1.2), and then to a pixel grid (Sect. 2.1.5). Instrumental effects such as PSF blurring and observational noise are added in post-processing (Sect. 2.2).

FORECAST has two available options for stellar population synthesis models: [Bruzual & Charlot \(2003\)](#), modeling stellar emission, and [Gutkin et al. \(2016\)](#), which additionally incorporates the rest-frame ultraviolet and optical nebular emission from HII regions around young stellar populations. We point out that we currently do not include Active Galactic Nuclei and individual Milky Way stars in the rendering of the simulated galaxies; this is left for future work.

#### 2.1.1. Input parameters

FORECAST is adaptable to the choices of the user by selecting a set of input parameters, described in Table 1. In particular, it is possible to choose the hydrodynamical simulation that provides the backbone of the light cone (box with side-length  $L_{\text{box}}$ ); the highest redshift to be included,  $z_s$ , which determines the maximum distance covered by the light cone,  $D_s$ ; the dimensions of the field of view,  $L_{\text{fov}}$ ; the resolution of the ideal simulated images, setting the number of pixels per image side-length  $N_{\text{pix}}$ .

#### 2.1.2. Light-cone construction

Simulation output data are organized in snapshots, which are photographs of the simulated cosmological volume at a specific

time of its evolution. FORECAST creates a light cone placing an observer at  $z = 0$  and rearranging the data from the output snapshots of the chosen simulation, projecting the positions of the simulated objects on a two-dimensional field of view.

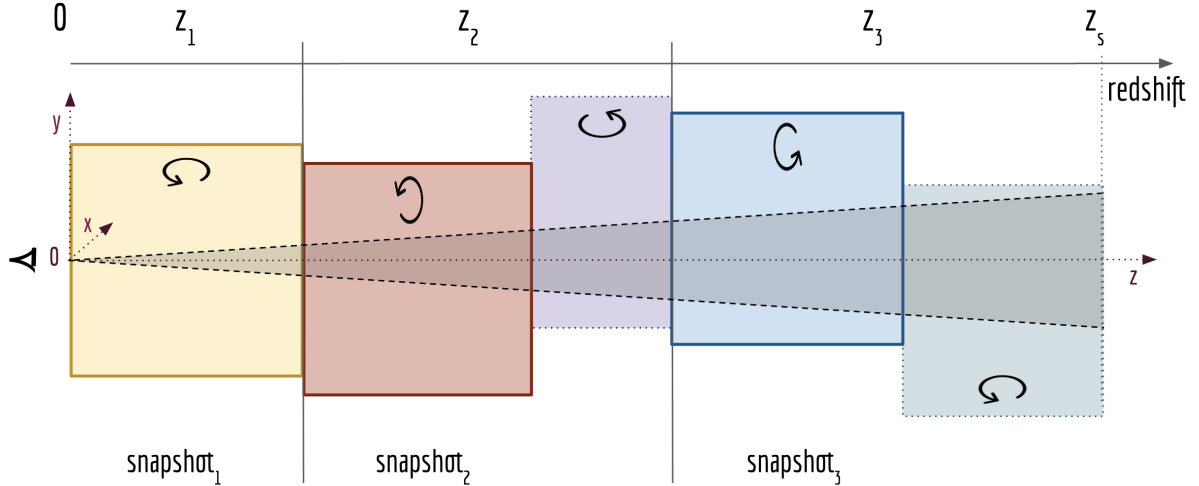
The light-cone construction procedure is inherited from the software MapSim by [Giocoli et al. \(2015\)](#). The snapshots used to build the light cone must include the following properties for the stellar resolution elements:

- comoving coordinates within the simulated volume,  $(x_*, y_*, z_*)$  in ckpc
- stellar mass,  $M_*$  in  $M_{\odot}$
- initial stellar mass,  $M_{i,*}$  in  $M_{\odot}$
- stellar metallicity  $Z_*$  as  $M_Z/M_{\text{TOT}}$
- age,  $t_{\text{SSP}}$  in yr
- redshift,  $z_*$
- subhalo membership ID

The initial stellar mass is the amount of mass owned by a stellar element when it is born, while the stellar mass accounts for mass returned through winds and supernovae to the ISM by evolved stars.

FORECAST constructs deep light cones stacking simulation boxes, using different snapshots to cover the entire chosen redshift interval (partitions of the light cone). When the length of the simulation box along the  $z$ -axis, which is the order of a few hundred Megaparsecs in typical cosmological hydrodynamical simulations, is smaller than the distance between two subsequent snapshots, FORECAST adds a replica of the previous or following snapshot, tailoring it to fill the gap. The software replicates the snapshot with the closest redshift to the redshift of the midpoint of the gap. This operation ensures a seamless construction of the complete light cone without any missing sections. Each snapshot is adjusted with rotated, inverted and shifted coordinates to pick structures at an evolutionary stage as close as possible to the one they would be if the entire redshift range was continuously sampled by simulation snapshots. Similar procedures are adopted by [Roncarelli et al. \(2006\)](#), [Croft et al. \(2001\)](#) who produced maps to study X-ray emission, and similarly by [Scaramella et al. \(1993\)](#), [da Silva et al. \(2000, 2001a,b\)](#) to study the Sunyaev-Zel’dovich effect.

The use of the same snapshot (required to fill potential gaps in the light cone) to reproduce adjacent but different volumes of the cone can cause the repeated appearance of the same structures aligned in radial direction (or in transverse direction, if the same volume is replicated at the same redshift to extend the field of view; this feature is not included). On the other hand, the use of different snapshots to reproduce different volumes of the cone at different cosmic times can also cause the recurrence of objects in the final image because each snapshot of a



**Fig. 1.** Illustration of the construction of the light cone. The cosmological volume between  $z=0$  and  $z_s$  is divided in bins of redshift, delimited by vertical black lines; the observer is located at the position  $O$ , on the left of the figure, centered with respect to the first simulation box at  $z = 0$ . The full light cone is realized by firstly stacking the comoving volume of the simulation at the proper redshift, represented by snapshots 1, 2, 3 located at redshifts  $z_1, z_2, z_3$  (boxes with thick contours), with rotated coordinates and shifted centers to avoid repetition of the structures along the  $z$ -axis. Then, the possible gaps between two contiguous snapshots (occurring because the distance between two subsequent snapshots is larger than the size of the simulation box along the  $z$ -axis) are filled up with replicas of existing snapshots (boxes with dashed contours), with volumes tailored to fill the gaps, and with rotated coordinates and shifted centers. The darker area is the light cone, growing with the comoving distance utmost to the transverse comoving size of the simulation box.

simulation includes the same sources at different epochs of their evolution. These periodicity effects caused by the repetition of the structures throughout the cone are mitigated by adopting the random combination of the following geometrical readjustments on each of the snapshot boxes used to construct the light cone: (i) the rotation of the positions of stellar particles of  $0, \pi/2, \pi$  or  $3\pi/2$  around each axis, (ii) the shift of their positions of random amplitude, in  $[0, L_{\text{box}}]$ , in  $(x, y, z)$  directions, imposing periodic boundary conditions, and (iii) the inversion of one randomly picked axis (see Blaizot et al. 2005). Figure 1 shows a sketch of the procedure adopted by FORECAST to construct the light cone, with the colored boxes representing the stacking of the snapshots along  $z$ , and the darker area outlining the maximum field of view up to  $z_s$ .

In a real light cone, the redshift of the sources varies continuously along the line of sight. However, the output of a simulation consists of a finite number of snapshots, each at a given precise redshift – so for example all the particles in the snapshot at  $z = 0$  have  $z = 0$ , even though the simulation box spans up to many comoving Megaparsecs. To cope with this, the actual redshift assigned to each particle is computed from its comoving distance from the observer (which is computed using its coordinates in the snapshot).

FORECAST recovers the subhalo membership of each particle, as previously assigned from the simulation procedure (e.g., hydrodynamical simulations usually adopt friends-of-friends group-finding algorithm, Davis et al. 1985; and SUBFIND algorithm for substructures identification, Springel et al. 2001; Dolag et al. 2009), in order to track the overall emission of the galaxy. The software then selects only the particles within the field of view (FoV), whose dimension is assigned by the user in the input file (see  $L_{\text{fov}}$  in Table 1). To this aim, it computes the distance between the particle and the observer located in the center of the box at  $z = 0$ , that is at  $(0.5, 0.5, 0.0) \cdot L_{\text{box}}$ , and it converts the comoving coordinates of each particle within the cone to angular positions. Particles with right ascension  $\alpha_* < L_{\text{fov}}$  and declination  $\delta_* < L_{\text{fov}}$ , and with comoving distance  $d_*$  within the range

of the considered partition ( $D_{\text{min}} \leq d_* \leq D_{\text{max}}$ ), are included in the cone. The dimension of the FoV cannot exceed the projected angular size given by the comoving box size placed at  $D_s$  from the observer. Because the light cone grows to a transverse comoving size equal maximum to the simulation box size, at low redshift only a small region of the simulation box is used. The discontinuities at the edge of the tiled partitions are a standard issue in cone construction (see e.g., Blaizot et al. 2005; Kitzbichler & White 2007; Bernyk et al. 2016). Moreover, some structures might be only partially included and cut on the edge of the field of view; since we work with particles rather than galaxies, we easily identified the partially built structures to be a few percentage ( $\sim 0.5\%$ ) within the field of view.

Since the dimension of the input files (snapshots) is typically large and it might be too demanding to have all of them simultaneously saved in the working space, FORECAST is designed to use each snapshot independently, allowing the user to make parallel runs.

### 2.1.3. Let there be light

For each stellar particle within the light cone, we then infer observational quantities starting from the knowledge of its intrinsic properties. The Bruzual & Charlot (2003) synthetic stellar population model (BC03) is linked to each stellar particle on the basis of its characteristics, namely the age and the metallicity, assuming a Chabrier (Chabrier 2003) or Salpeter (Salpeter 1955) initial mass function, at user choice. The stellar particle is assigned to the BC03 SED with age and metallicity closest to its nominal age and metallicity. The rest-frame, intrinsic spectral energy distribution  $L_{\lambda,*}(\lambda, t, Z)$  of the SSP is then converted into the observer-frame flux per unit wavelength  $F_{\lambda,*}(\lambda, t, Z)$ , taking into account the redshift  $z$  and therefore the luminosity distance  $d_{L,*}(z)$  of the considered particle from the observer:

$$F_{\lambda,*}(\lambda_{\text{obs}}, t, z, Z) = \frac{L_{\lambda,*}(\lambda_{em}, t, Z)}{4\pi(1+z)d_{L,*}^2(z)} e^{-\tau_{\text{igm}}(\lambda_{\text{obs}}, z)}, \quad (1)$$

where  $\tau_{\text{igm}}(\lambda_{\text{obs}}, z)$  is the optical depth of the intergalactic medium, computed from the IGM absorption model by [Inoue et al. \(2014\)](#). Finally, the apparent AB magnitude of the SSP at redshift  $z$ , corresponding to the integrated photon flux collected at  $z = 0$  from the chosen detector with a filter response  $R(\lambda)$  is computed following [Fukugita et al. \(1996\)](#). Firstly, FORECAST evaluates the apparent magnitude of a  $1 M_{\odot}$  SSP, namely  $m_{1 M_{\odot}, AB, *}$ ; it is then rescaled with the initial stellar mass of the particle  $M_{i, *}$  in order to follow the same stellar mass loss as in BC03. The final apparent magnitude of the SSP

$$m_{AB, *} = m_{1 M_{\odot}, AB, *} - 2.5 \cdot \log_{10}(M_{i, *} / M_{\odot}), \quad (2)$$

is finally reconverted into integrated observed flux in units of  $\mu\text{Jy}$ .

#### 2.1.4. Adding dust attenuation

Correctly taking into account dust extinction by the ISM in a simulation would require the knowledge of the chemical composition, structure, and size distribution of dust grains for each given physical state, and this is rarely included ab initio in simulations. A detailed inclusion of dust physics has only been achieved recently in galaxy formation simulations (e.g., [Bekki 2015](#); [Aoyama et al. 2018](#); [McKinnon et al. 2018](#); [Graziani et al. 2020](#)); more often models incorporate a full treatment of dust with radiative transfer codes (e.g., SKIRT by [Baes et al. 2003, 2011](#); DIRTY by [Gordon et al. 2001](#); Misselt et al. 2001; SUNRISE by [Jonsson 2006](#); [Jonsson et al. 2010](#); HYPERION by [Robitaille 2011](#)), that are able to handle absorption, scattering, and thermal emission by interstellar dust with different solution methods for the radiative transfer equation (e.g., probabilistic methods, numerical methods). However, these methods can be computationally very expensive if the number of particles and/or the number of interactions between particles are increased in the attempt to reduce stochastic fluctuations (e.g., Monte Carlo methods), or can lead to very complex numerical schemes when adopting numerical solutions in the attempt to minimize the introduced numerical errors (e.g., ray-tracing methods).

Since dust resolution elements are not self-consistently included in most large-scale hydrodynamical models and, in general, it is more likely to work on simulations that do not include dust, FORECAST explicitly models the effect of dust in post-processing. It manipulates the properties of gas resolution elements already incorporated in the simulation to turn their observed fluxes, derived considering only the stellar component or the stellar component combined with nebular emission (depending on the model chosen to generate SEDs, see Sect. 2.1), into dust-corrected fluxes. To improve the readability, we refer to the fluxes computed without the explicit dust attenuation contribution as “dust-free” fluxes, even if they include nebular lines in the modeled SEDs.

Following [Guiderdoni & Rocca-Volmerange \(1987\)](#); [Devriendt & Guiderdoni \(2000\)](#); [Nelson et al. \(2019\)](#); [Vogelsberger et al. \(2020\)](#), we adopt (i) a semi-analytic model to account for the effect of dust below the resolution limit; (ii) an explicit, geometry-dependent model to account for attenuation by dust in the resolved gas component, using the neutral fraction of gas elements as dust tracer. The fluxes that include these models in their computation are tagged as “dust-corrected” fluxes. We point out that dust emission, including both the predominantly impacting far-infrared and submillimeter wavelengths as well as the mid-infrared contributions from PAH emission ([Draine et al. 2021](#); [Liu et al. 2023](#)), is currently not included in FORECAST.

For the unresolved dust component, we follow [Charlot & Fall \(2000\)](#). In their model, young stellar populations ionize the inner regions of dense birth clouds within the ISM; then, line photons emitted in the HII region, and ultraviolet (UV) and optical non-ionizing continuum from young stars are absorbed by dust in the outer HI region and the ISM. Nonetheless, the stellar UV continuum from stars that are no longer in their birth clouds results to be less attenuated than HII emission lines from newborn stars because, after the birth clouds disruption, it is attenuated only within the ISM. Therefore, the intrinsic luminosity of each SSP is obscured as  $L_{\lambda, *}^a(\lambda) = L_{\lambda, *}^{\text{int}}(\lambda) \cdot r_{\text{me}}^{\tau_{\lambda}^{\text{un}}}$ , with the unresolved dust optical depth  $\tau_{\lambda}^{\text{un}}$

$$\tau_{\lambda}^{\text{un}} = \begin{cases} \tau_1 (\lambda/5500 \text{ \AA})^{-0.7} & t_{\text{ssp}} \leq 10 \text{ Myr}, \\ \tau_2 (\lambda/5500 \text{ \AA})^{-0.7} & t_{\text{ssp}} > 10 \text{ Myr}. \end{cases} \quad (3)$$

It is assumed that birth clouds and ambient ISM have the same absorption curves, with different normalization coefficients. All the parameters are taken from the original work, in particular the normalization coefficient at  $5500 \text{ \AA}$ ,  $\tau_1 = 1.0$ , accounts for both photon lines and continuum radiation absorption; this value is then lowered to  $\tau_2 = 0.3$  after  $t_{\text{bc}} = 10 \text{ Myr}$ , when birth clouds typically dissipate in Milky Way ([Murray et al. 2010](#); [Murray 2011](#)).

We also include absorption due to resolved dust component by using the distribution of gas-resolution elements in and around each galaxy as tracers for dust (we do not consider intergalactic dust). FORECAST selects gas elements along the line of sight of each stellar particle belonging to a galaxy; by considering the nominal properties of the selected gas along all line of sights, it computes the gas properties averaged over the whole galaxies to determine the dust optical depth. This optical depth is then applied to the full galaxy SED to compute the attenuated galaxy fluxes in the chosen filters. Finally, the software derives the dust attenuation as the ratio between the full galaxy dust-free and the dust-corrected fluxes in each band. This derived dust attenuation is then applied to the dust-free fluxes of each stellar particle in the galaxy, since the software works on a particle-basis.

In more detail, the dust-corrected (d-c) luminosity of the galaxy is computed from the dust-free (d-f) luminosity, considering the internal dust model by [Calzetti et al. \(1994\)](#), according to which dust and ionized gas are uniformly mixed

$$L_{\text{d-c}}(\lambda) = L_{\text{d-f}} \frac{1 - e^{-\tau_{\lambda}^r}}{\tau_{\lambda}^r}, \quad (4)$$

where  $\tau_{\lambda}^r$  is the dust optical depth that accounts for absorption and the effect of scattering. Introducing the albedo of grains  $\omega_{\lambda}$  ([Draine & Lee 1984](#)), the scattering anisotropy weight parameter  $h_{\lambda}$ , the scattering optical depth  $\tau_{\lambda}^s$ , and the absorption optical depth  $\tau_{\lambda}^a$  for resolved dust elements, we derive the total resolved dust optical depth

$$\tau_{\lambda}^r = \tau_{\lambda}^a \cdot \tau_{\lambda}^s = \tau_{\lambda}^a \cdot \left[ h_{\lambda} \sqrt{1 - \omega_{\lambda}} + (1 - h_{\lambda})(1 - \omega_{\lambda}) \right]. \quad (5)$$

Scattering parameters are taken from [Calzetti et al. \(1994\)](#) in  $\lambda$  belonging to  $[1000, 7000] \text{ \AA}$  range. The resolved dust absorption optical depth ([Nelson et al. 2019](#)) depends on the metallicity  $Z_{\text{g}}$  and on the neutral hydrogen column density  $N_{\text{HI}}$  of the gaseous elements as follows:

$$\tau_{\lambda}^a = \left( \frac{A_{\lambda}}{A_{\text{V}}} \right)_{\odot} (1 + z)^{-0.5} \left( \frac{Z_{\text{g}}}{Z_{\odot}} \right)^{\gamma} \left( \frac{N_{\text{HI}}}{N_{\text{HI}, 0}} \right). \quad (6)$$

The first term is the extinction law in the solar neighborhood taken from Table 2 in Cardelli et al. (1989). The second and third terms express the dependency of the dust-to-gas ratio on redshift and metallicity, as studied by Dunne et al. (2011); Rémy-Ruyer et al. (2014); McKinnon et al. (2016). According to Guiderdoni & Rocca-Volmerange (1987),  $\gamma$  exponent is a broken power law:  $\gamma = 1.35$  for  $\lambda < 2000 \text{ \AA}$ , and  $\gamma = 1.6$  for  $\lambda > 2000 \text{ \AA}$ ; the normalization parameters are the solar metallicity  $Z_{\odot} = 0.02$  (Anders & Grevesse 1989) and the neutral hydrogen column density  $N_{\text{HI},0} = 2.1 \times 10^{21} \text{ cm}^{-2}$ .  $N_{\text{HI}}$  and  $Z_g$  are estimated as HI mass-weighted averages of the properties of the gas elements that lie along the  $z$ -axis line of sight of each stellar particle forming a galaxy;  $\tau_{\lambda}^r$  and  $\tau_{\lambda}^a$  are computed for the whole galaxy. The gas properties are taken from the output files of the hydrodynamical simulation.

To account for dust in the light cone, FORECAST requires the following properties to be available for each gas particle or cell in the snapshot files:

- comoving coordinates of the geometrical center within the snapshot,  $(x_g, y_g, z_g)$  in ckpc
- gas mass,  $M_g$  in  $M_{\odot}$
- gas comoving volume,  $V_g$  in  $\text{ckpc}^3$
- gas metallicity,  $Z_g$  in  $M_Z/M_{\text{TOT}}$  (with  $M_Z$  the total mass all metal elements)
- neutral hydrogen column density within each gas cell,  $N_{\text{HI}}$  in  $\text{ckpc}^{-2}$
- redshift,  $z_g$
- subhalo membership ID

Gas cells are the resolution element in Eulerian hydrodynamical codes or moving-mesh codes (e.g., Teyssier 2002; Bryan et al. 2014; Hopkins 2015; Springel 2010), and their volume can be approximated as a cube or a sphere. In case of Smoothed Particles Hydrodynamics codes (e.g., Springel 2005; Wadsley et al. 2004), where resolution elements are represented by particles, one can consider the smoothing length of gas particles as their spatial extension. FORECAST approximates the volume of a gas resolution element as a cube, thus it computes its linear dimension as  $L_g = (V_g)^{1/3}$ .

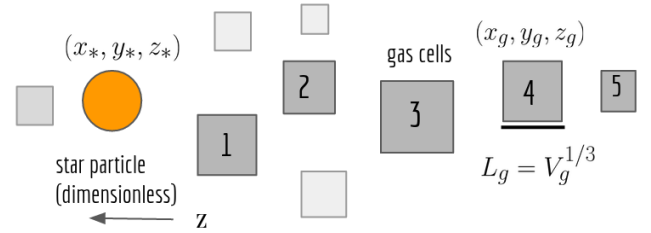
Since we assume Eq. (6) to relate the properties of the gas with dust absorption, we require  $N_{\text{HI}}$ , whose availability in output may depend on the specific hydrodynamical simulation used to construct the mock images. As an example, for ILLUSTRISTNG the informations about  $N_{\text{HI}}$  are available only at some redshift; thus we calibrated the relation between the temperature of the gas cells and the ratio between the neutral hydrogen mass and the total mass within gas cells on TNG data to evaluate the neutral hydrogen column density, as we show in Sect. 3.1.1.

For each stellar particle, FORECAST selects the gas resolution elements that have its same subhalo membership, satisfying the following conditions along its  $z$ -axis line of sight

$$\begin{cases} x_g - 0.5 \cdot L_g \leq x_* \leq x_g + 0.5 \cdot L_g, \\ y_g - 0.5 \cdot L_g \leq y_* \leq y_g + 0.5 \cdot L_g, \\ z_g < z_* . \end{cases} \quad (7)$$

The selection procedure is illustrated in Fig. 2, in case gas resolution elements are cells.

To maintain the computing time reasonable (each snapshot might include up to billions of star and gas elements), we compute the dust optical depth on the whole galactic spectrum by averaging the properties of the gas selected over the entire galaxy. Then, we derive and assign the attenuation in each band to the fluxes of each SSP constituting the galaxy.



**Fig. 2.** Selection procedure of gas cells (gray squares) along the line of sight of a stellar particle (yellow circle). The stellar particle, with comoving coordinates  $(x_*, y_*, z_*)$ , is a dimensionless point, while gas cells have a linear extension  $L_g$  in  $(x, y, z)$ , assuming cubic cells with volume  $V_g$ . The  $z$ -axis grows from right to left, meaning that darker gray gas cells are in front of the star particle and along its line of sight (LoS), compared to the observer located at the origin  $n$  of the  $z$ -axis. Lighter gray gas cells are not included in the selection, because they are behind or out of the LoS of the star particle.

The neutral hydrogen column density and the metallicity of the gas within each galaxy are independently computed as the result of neutral hydrogen mass-weighted quantities of the selected  $i$ th gas cells in front of each SSP belonging to that galaxy as follows:

$$\begin{aligned} \langle Z_g \rangle_{\text{gal}} &= \frac{\sum_i Z_{i,g} \cdot M_{i,\text{HI}}}{\sum_i M_{i,\text{HI}}}, \\ \langle N_{\text{HI}} \rangle_{\text{gal}} &= \frac{\sum_i N_{i,\text{HI}} \cdot M_{i,\text{HI}}}{\sum_i M_{i,\text{HI}}}. \end{aligned} \quad (8)$$

These quantities, which are computed on galaxy basis, are used to estimate the resolved dust optical depth in Eq. (6), which is applied to each galaxy SED; then, we determine the galaxy dust-corrected observed flux per unit of wavelength, which is then evaluated within the chosen filter response  $R(\lambda)$  to obtain the galaxy dust-corrected integrated flux.

Galaxy dust-free integrated flux and dust-corrected integrated flux are finally used to estimate the average resolved dust attenuation for each galaxy

$$A_{\text{gal}}^r = -2.5 \log_{10} \left( \frac{F_{\text{gal,d-c}}}{F_{\text{gal,d-f}}} \right) = -2.5 \log_{10} \left( \frac{1 - e^{-\tau_{\lambda}^r}}{\tau_{\lambda}^r} \right). \quad (9)$$

Since we use a particle-based approach and we are interested in obtaining the observed integrated flux of each SSP in the light cone, this attenuation is then applied to the dust-free apparent AB magnitude of each stellar particle, to get their dust-corrected apparent AB magnitudes

$$m_{\text{AB},*}^{\text{d-c}} = m_{\text{AB},*}^{\text{d-f}} + A_{\text{gal}}^r. \quad (10)$$

To map the light of the SSPs onto the mock image, we convert the dust-corrected apparent AB magnitude of each SSP into observed flux, in  $\mu\text{Jy}$  units. We point out that to streamline computational time, our approach assumes uniform dust attenuation for the SSPs belonging to the same galaxy. Nevertheless, we are working on enhancing both the computational efficiency and the realism of the dust attenuation model to achieve more accurate results.

We accurately checked that the adopted dust model produces the expected changes in the observed fluxes: accounting for dust, the rest-frame UV and optical magnitudes are the most affected, increasing up to one order of magnitude with respect to their dust-free counterpart, especially in shorter wavelengths, while IR magnitudes remain unchanged.

### 2.1.5. Mapping

The fluxes of the stellar particles that build up the simulated galaxies must finally be mapped on a bi-dimensional pixel grid. First, the angular coordinates of the particles in each partition (portion of the light cone) are projected on a bi-dimensional plane located in the central point of the considered partition, and then they are translated in pixel coordinates.

FORECAST first converts the right ascension and declination coordinates of the particles in the considered partition into pixel coordinates, accounting for the FoV dimension in pixels ( $N_{\text{pix}}$  is defined in Table 1). The total flux of each pixel is obtained as the sum of the fluxes of all particles having coordinates within it. Since, at this stage, the synthetic image is ideal (i.e., the number of photons hitting the mirror of the telescope is a smooth function from the theoretical intensity of any source; there are no diffraction effects due to the limited surface of the optics of the telescope, or other sources of uncertainties and errors), we choose not to apply any kernel convolution matrix that distributes the flux of a particle on adjacent pixels; instrumental effects, including noise and PSF smoothing, are attached in post-processing on the final image in a chosen filter (see Sect. 2.2). In the real world, each single point source (i.e., SSP) gets smeared by the PSF of the telescope, and therefore the most accurate way to simulate this effect would be to stack PSF stamps (one per SSP), shifted, and rebinned in order to have their centers at the exact, sub-pixel position of the corresponding particle. However, we have implemented this algorithm and checked that the final result is virtually identical to another one in which we first sum the flux of all the SSPs falling into a pixel, and then simply PSF-smooth the light of that pixel. We therefore choose to adopt the latter method, which is simpler and requires a much smaller expense of computational time.

The final image is obtained by stacking the images corresponding to each projected plane of the partitions that build up the light cone.

### 2.2. Adding realism: Noise and PSF

Here we describe the two final steps necessary to simulate a real mosaic, which are the inclusion of observational noise and the convolution with a Point Spread Function of the sources of interest. These two steps are independently performed with a `python` script and can be executed multiple times on the ideal images produced by FORECAST to simulate them with different depths.

When FORECAST assembles its output in a chosen band, the simulated image is noiseless and its resolution is only limited by the pixel scale. First of all, as mentioned in Sect. 3.1, we convolve it with the PSF of the instrument and filter that is being simulated; this spreads the flux coming from a single pixel over an extended region.

Then, we add a noise background that limits the depth of the image. To this aim, we create an RMS map as a flat image with a constant value, that is the chosen standard deviation of the background noise pixels, which can be obtained from the desired limiting magnitude of the image as

$$\sigma = \frac{10^{-0.4(m_{\text{lim},n} - \text{ZP})}}{n \sqrt{\pi} r_{\text{ap}}}, \quad (11)$$

where  $n$  is the signal-to-noise ratio to which the limiting magnitude corresponds.  $r_{\text{ap}}$  is the radius of the aperture used to compute the limiting magnitude; it does not have a fixed value, as it depends on how the depth of an image is defined (e.g., if

**Table 2.** Parameters for image post-processing.

Parameter	Units	Description
INPUT imm	–	FORECAST image
PSF	–	Point spread function
ZP	–	Image zero-point
FWHM	Arcsec	Full width half maximum
TEXP	s	Integrated exposure time
RMSimg	–	RMS image
RMSerr	–	RMS error
PS	arcsec pixel <sup>-1</sup>	Pixel scale
APER	arcsec	Diameter aperture
BKGD	–	Background

a band has limiting magnitude  $mag_{\text{lim}}$  at  $N\sigma$  in 2", the radius will be 1" in pixels, and the noise level will be set to match the expected depth). ZP is the zero-point of the image (in this case, ZP = 23.9 with the pixels in  $\mu\text{Jy}$ ).

Then, we add to the RMS map the contribution of the photon noise. This is done by replacing the value of each pixel,  $\sigma_{\text{sky}}$ , with

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{sky}}^2 + \frac{f_{\text{source}}}{t_{\text{exp}}}}, \quad (12)$$

where  $t_{\text{exp}}$  is the total exposure time of the image, and  $f_{\text{source}}$  is the flux coming from luminous sources falling in that pixel (this formula can be derived from first principles and is discussed in Merlin et al. 2023).

Finally, we create the noise image as a random realization with the  $\sigma$  of the distribution for each pixel provided by the RMS map, and we sum it to the original noiseless image containing the simulated galaxies to obtain the final mock image. The RMS map can also be used to compute the errors on any photometric measurement performed on the scientific image.

We included an option to further slightly smooth the scientific image, in order to simulate the result of possible noise correlation of pixels introduced by mosaicing, rebinning, and stacking. The apparent limiting magnitude of the scientific image (i.e., the noise standard deviation when the background is subtracted) can be chosen to be different from the one given by the RMS map; if this option is chosen, the image is iteratively smoothed with a Gaussian kernel until the desired apparent depth is obtained.

All the images are finally normalized to  $\mu\text{Jy}$  units by default, meaning that the magnitude associated to each pixel flux value is simply  $m = -2.5 \log(f) + 23.9$ . We include the possibility of creating the images with any different normalization, giving a different zero-point in input. The values that can be configured to perform the post-processing of the image are summarized in Table 2.

## 3. Testing FORECAST: Emulation of the CANDELS GOODS-South Field

Before exploiting the tool to make forecasts for the next-generation surveys, we tested its capabilities emulating a well-known dataset, to compare the new mock data with existing photometric catalogs, investigating which are the most relevant tensions with observed data, and how the procedure can be improved for future work. These data products, together with the

JWST CEERS dataset (see Sect. 4), are publicly available for scientific analysis.

In Sect. 3.1, we illustrate the procedure adopted to build the images and catalog analyzed in this work. We describe the cosmological hydrodynamical simulation employed to build the images in Sect. 3.1.1. We then describe the mock light-cone setup adopted for the present analysis and give an overview of the simulated images in Sect. 3.1.2.

### 3.1. Constructing the images

In order to examine the performance of the software tool, we tested the synthetic images against a thoroughly investigated observational counterpart. To this aim, the size and filter set of the simulated FoV and the extension in redshift of the mock light cone have been chosen to emulate the Great Observatories Origins Deep Survey Field South (GOODS-South, GS), making use of the state-of-the-art hydrodynamical simulation ILLUSTRISTNG. The GS field, located at RA = 3 h 32 m 30.39 s and Dec =  $-27^{\circ} 48' 11.28''$  with a covered region of  $10' \times 16'$ , has been targeted for deep, multiwavelength observations from ground and from space as part of several survey programs (e.g., Grogin et al. 2011; Brammer et al. 2012; Curtis-Lake et al. 2023; Robertson et al. 2023).

We elect to produce thirteen images corresponding to 13 band-passes covering from the rest-frame optical to the near-infrared (NIR) wavelength range: HST ACS B435, V606, I814, Z850, and WFC3 Y105, J125, JH140, and H160; plus a ground-based VLT HAWK-I  $K_s$  band, and four IRAC channels, namely CH1, CH2, CH3, and CH4. We post-processed the FORECAST ideal images with the real PSF models adopted in the CANDLES image analysis described in Guo et al. (2013), see Sect. 2.2 for further details. Since our goal was to test the tool over a fairly wide range of wavelengths, we only emulated 13 out of the 17 bands of the reference catalog.

#### 3.1.1. IllustristNG simulation

The ILLUSTRISTNG (or TNG) Project (Weinberger et al. 2018; Pillepich et al. 2018b; Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Springel et al. 2017) is a suite of cosmological magneto-hydrodynamical simulations, following the evolution of cosmological volumes between  $z = 20$  to  $z = 0$ . The moving-mesh code AREPO (Springel 2010) solves coupled equations for gravity and magneto-hydrodynamics: Poisson's equations for full Newtonian gravity are treated with a hybrid TreePM scheme (Xu 1995; Bagla 2002), while an unstructured and moving Voronoi mesh is adopted to solve equations of hydrodynamics. Processes such as the cosmic gas accretion into halos, tidal and ram-pressure stripping, and dynamical friction, as well as the hierarchical growth of halos and galaxies, and galaxy mergers naturally emerge as the solution of the equations of gravity and hydrodynamics in an expanding Universe with gravitationally collapsing structures.

Its galaxy formation model (see Weinberger et al. 2018; Pillepich et al. 2018b) is built upon the original ILLUSTRIS model (Vogelsberger et al. 2013; Genel et al. 2014; Sijacki et al. 2015) and accounts for all the processes occurring below the resolution scale of the simulation. It includes gas density-threshold star formation, adopting a Chabrier IMF (Chabrier 2003), and evolution of stellar populations represented by star particles; chemical enrichment of the ISM with the tracking of nine chemical elements (H, He, C, N, O, Ne, Mg, Si, Fe); gas heating and cooling; feedback from supernovae through galactic winds;

seeding and growth of supermassive black holes, and energy- and momentum-driven feedback into the surrounding gas. TNG cosmology is consistent with recent observational constraints from Planck Collaboration XIII (2016), cosmological constant  $\Omega_{\Lambda,0} = 0.6911$ , matter density  $\Omega_{m,0} = 0.3089$ , baryon density  $\Omega_{b,0} = 0.0486$ , power spectrum normalization  $\sigma_8 = 0.8159$ , and spectral index  $n_s = 0.9667$ ; the Hubble constant is  $H = 100h$  Mpc, with  $h = 0.6774$ .

The ILLUSTRISTNG Project consists of three physical simulation boxes with periodic cubic volumes of roughly 50, 100, and 300 comoving Mpc (cMpc) side-length, named TNG50, TNG100, and TNG300 respectively, each of them reproduced at high, medium and low-resolution level ( $-1$ ,  $-2$ ,  $-3$  suffix). Gas cell masses (i.e., the mass resolution of a simulation) and gas cell sizes (i.e., the spatial resolution of a simulation) both form continuous distributions. Stars inherit the gas mass from which they form, so they also have a variable mass resolution, and they continuously decrease in mass due to stellar evolution (Pillepich et al. 2018b).

The ILLUSTRISTNG dataset has been publicly released with a complete user guide in Nelson et al. (2019). In each output snapshot of the simulation, overdensities of dark matter are identified using the Friends-of-Friends algorithm (Davis et al. 1985); self-bound subhalos, which are primarily constructed with dark matter particles, and have their baryons (i.e., gas and stars) associated with the geometrically nearest DM particle, are later identified using the SUBFIND algorithm for substructures identification (Springel et al. 2001; Dolag et al. 2009). They also generate two distinct merger trees at the subhalo level, with SUBLINK (Rodriguez-Gomez et al. 2015) and LHALOTREE (Springel et al. 2005) algorithms.

This work makes use of the TNG100-1 simulation boxes ( $L = 75 h^{-1}$  Mpc = 110.7 Mpc), but the tool is configured to use any realization of the simulation.

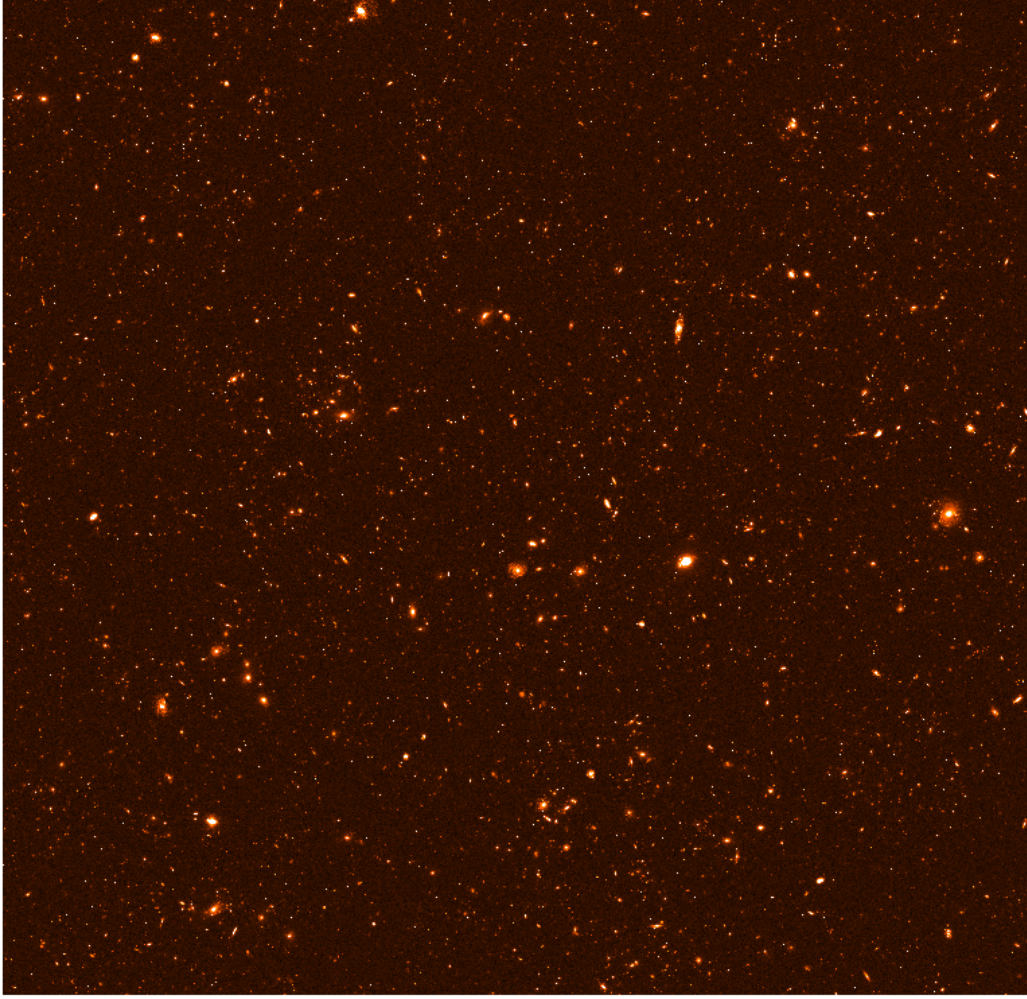
#### 3.1.2. Emulation setup and output

We created a light cone from  $z = 0.1$  (we excluded lower redshift snapshot to avoid excessive contamination from large local sources) up to  $z_s = 7.2$ , corresponding to a comoving distance  $D_s = 6034.14$  cMpc  $h^{-1}$  (with  $h = 0.6774$ ). The light cone is built with 122 partitions, with 89 output snapshots. We adopt a Chabrier IMF for the modeled SSPs, consistently with the TNG choice.

The mock survey is designed to provide a coverage of galaxies from optical to near-IR wavelengths, emulating a squared field of view of 200 sq. arcmin (comparable to the GOODS-South field area) realized on a grid of 200 million pixels, resulting in a pixel scale of 0.06 arcsec, which is the typical for HST observations (for simplicity, the VLT image and the *Spitzer* images were directly created on the same pixel scale, rather than going through a rebinning process).

We simulated the thirteen broad-band images adopting resolution and limit magnitude from Merlin et al. (2021), shown in Table 3.

We show the final simulated F160W (or H160) image, post-processed with observational features, in Fig. 3. The complex morphologies and interactions of galaxies as observed in the real sky beautifully show up in the simulated image. We show three examples of galaxies or small groups located in the final FoV in four of the thirteen simulated bands in Fig. 4. These are all located in the low-redshift Universe, between  $z = 0.35$ – $0.45$ . It is possible to appreciate the morphology and brightness changes across the spectrum, with sources appearing more luminous in



**Fig. 3.** Final simulated image in  $H160$  band, after post-processing with PSF and noise, color-coded by fluxes in units of  $\mu\text{J}$ .

**Table 3.** Summary of the instrumental PSF and depths adopted for the image simulations in this work from [Merlin et al. \(2021\)](#).

Instrument	Filter	PSF(arcsec)	$5\sigma$ depth AB
HST ACS	F435W	0.08	28.83 <sup>(a)</sup>
	F606W	0.08	29.24 <sup>(a)</sup>
	F814W	0.09	29.35 <sup>(a)</sup>
	F850LP	0.09	28.54 <sup>(a)</sup>
HST WFC3	F105W	0.15	28.70 <sup>(a)</sup>
	F125W	0.16	28.85 <sup>(a)</sup>
	F140W	0.17	27.64 <sup>(a)</sup>
	F160W	0.17	28.72 <sup>(a)</sup>
VLT HAWK-I	$K_s$	0.43	26.26 <sup>(b)</sup>
<i>Spitzer</i> IRAC	CH1	1.66	25.63 <sup>(b)</sup>
	CH2	1.72	25.51 <sup>(b)</sup>
	CH3	1.88	23.28 <sup>(b)</sup>
	CH4	1.98	23.16 <sup>(b)</sup>

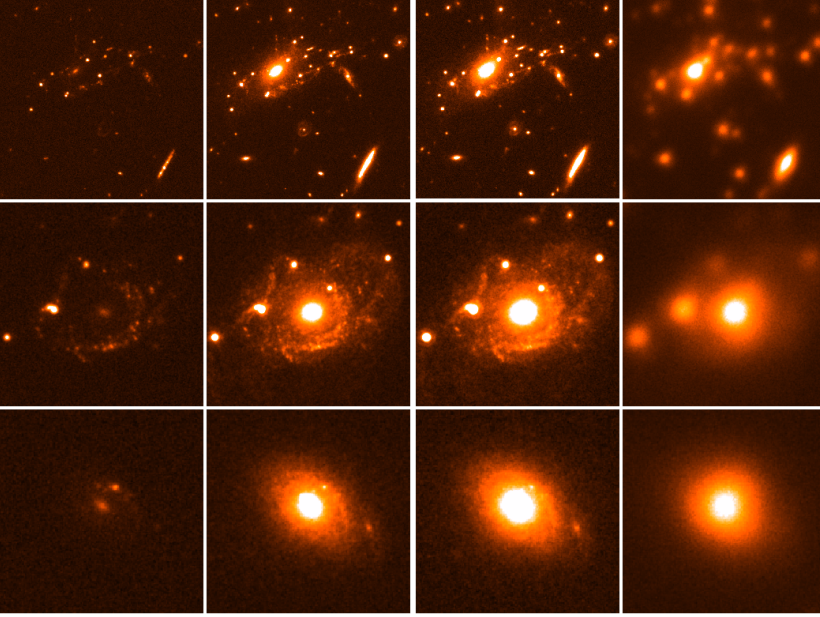
**Notes.** <sup>(a)</sup>Median aperture magnitudes within a fixed radius of  $0.17''$ ; the given values are averages of the varying depths in the CANDELS GOODS-South field (including CANDELS-deep field and the HUDF depths). <sup>(b)</sup>Median total magnitude at  $5\sigma$ ; the given values are averages of the varying depths in the CANDELS GOODS-South field.

infrared bands because of the typical SED shape of galaxies in which the star formation activity is not prominent. Bluer bands show more clearly the signs of recent star formation activity as concentrated blobs of high luminosity. It is possible to fully characterize these regions by checking the true ages and metallicities of the corresponding SSPs.

### 3.2. Photometric analysis

In order to validate the accuracy of the simulated field, we proceeded using a processing pipeline that is very similar to the one typically used to extract the photometric information, and then the scientific properties, from real imaging data. We have only bypassed the reduction processing steps of a typical raw imaging dataset (e.g., flat-fielding, bias and background subtraction, mosaicing, etc.), assuming they have been performed in an ideal way. We also did not extract new PSF models from the images, exploiting the ones used to build the simulated images.

Therefore, we start our analysis from a simulated image that is comparable to the final mosaic on which the CANDELS team performed the photometric measurements for the final catalogs. Specifically, for our analysis and comparisons, we used our multiwavelength photometric catalog ASTRODEEP-GS43, which is an upgrade of the CANDELS catalog by [Guo et al. \(2013\)](#), providing photometric fluxes in 43 passbands, plus



**Fig. 4.** Three examples of small areas of the simulated field of view containing a small group of galaxies or single objects (top to bottom), in four simulated bands (left to right:  $B435$ ,  $Y105$ ,  $H160$ , IRAC CH1, in  $\mu\text{J}$  units). They are low-redshift sources, with  $z = 0.35\text{--}0.45$ . The sizes of the areas are 0.401, 0.062 and 0.027 sq. arcsec, from top to bottom.

physical properties and estimations of the photometric redshift for  $\sim 35\,000$  sources located in the GS field.

The final output of FORECAST used for the analysis are (i) the synthetic images in all the simulated filters (see Table 3 for resolution and depth informations), post-processed as described in Sect. 2.2, and (ii) the input galaxy catalog (see Table C.1 for details on the fields contained in the catalog), that we call Input Universe (IU). This catalog, which is galaxy-based, is built from the properties of the stellar particles listed in the ILLUSTRISTNG output and included in the light cone by FORECAST. The additive properties (e.g., mass, flux) are computed as the sum from all SSPs belonging to a given galactic subhalo in the simulation; metallicity and age are weighted with the stellar mass of the SSPs, and redshift is computed as the mean value of the redshifts of all the membership SSPs. The coordinates of the center of the galaxies are computed as the flux-weighted sum (in  $H160$ ) of the coordinates of all the particles of the subhalo. For extended objects, which appear as separate clumps of light in the image but are identified as single objects in the ILLUSTRISTNG simulation, we refine the center computation with a  $3\sigma$ -clipping procedure; that is, we only consider the SSPs for which the distance from the previously determined center  $R(x_i, y_i, x_{c,0}, y_{c,0}) < 3\sigma(R)$ , thus excluding scattered, isolated particles which might bias the estimate of the center coordinates.

### 3.2.1. Detection and photometry

We performed the photometric analysis on the simulated images following the procedure adopted by Guo et al. (2013) and Merlin et al. (2021). We first detected sources on the  $H$  band image with SExtractor (Bertschinger & Gelb 1991) adopting Hot+Cold detection for a finer deblending of the sources. We then remeasured  $H$  fluxes with A-PHOT (Merlin et al. 2019), which yields a less biased estimate of the total flux (see Merlin et al. 2022). Fluxes in the remaining HST bands were obtained by correcting the total  $H$  flux by the color aperture term of the considered band, that is  $f_{\text{band,tot}} = f_{H,\text{tot}} \times (f_{\text{band,segm}} / f_{H,\text{segm}})$ , with the fluxes again measured with A-PHOT, after PSF-matching all the images to the  $H160$  resolution as described in Guo et al. (2013). For the  $K$  band and the four IRAC bands, which have lower resolution than the detection  $H$  image, we used T-PHOT

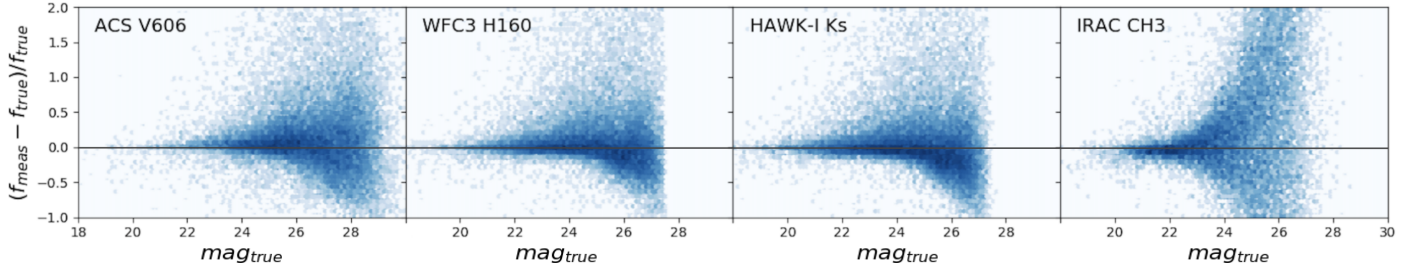
(Merlin et al. 2015, 2016) to perform template-fitting photometry; we also measured  $H$  band fluxes on the image PSF-matched with the lower-resolution ones, again to estimate a robust color term. Finally, we assembled a catalog with the total fluxes of all detected sources in all bands. All the uncertainties associated with the flux measurements have been computed using the RMS maps created as described in Sect. 2.2.

### 3.2.2. Multiband photometry

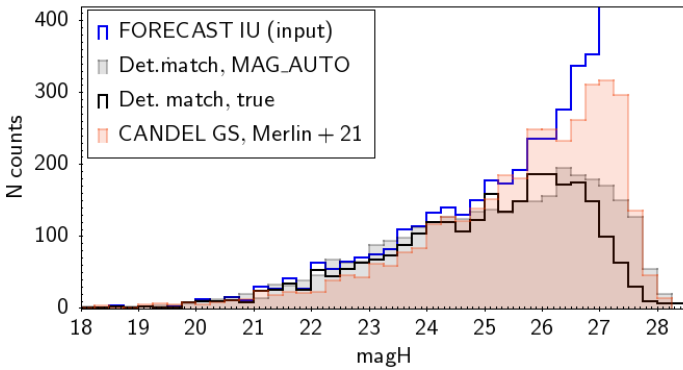
We assessed the accuracy of the fluxes measured with the standard photometric approach by comparing them with the input, true fluxes. The latter can be easily obtained as the sum of the dust and IGM-attenuated flux from all SSPs belonging to a given galactic subhalo in the simulation. Figure 5 shows the comparison between the input fluxes and the fluxes measured in  $V606$ ,  $H160$ ,  $K_s$  and IRAC CH3 bands, after a spatial cross-correlation of the SExtractor detections with the sources in the input IU catalog (only considering galaxies with  $H_{\text{true}} < 27.5$ ) using TOPCAT (Taylor 2005) to find the closest neighbors within a searching radius of 3 FWHM. We found 32 413 matched objects (99.6% of the total; the missing ones are spurious detections). We also checked the other nine simulated bands. The overall agreement is good. The most prominent features are the presence of some bright outliers, most likely due to contamination from neighboring sources, and most of all, a declining trend at low magnitudes; we note that the trend is present in the detection band  $H160$ , on which the total flux used to scale the colors in all bands is computed, and it is propagated to all other bands, while the colors term are estimated robustly.

### 3.2.3. H160 number counts

We then checked the number counts as a function of the magnitude of the detected objects in the simulated  $H$  image. Figure 6 shows a comparison of the counts between the simulation (both the IU and the detections) and a sample of the CANDELS GS “Deep” area. For this comparison, the GS sources were selected taking a crop of  $4000 \times 4000$  pixels ( $\sim 16$  sq. arcmin) in the deep region (centered on RA 53.0899 and Dec  $-27.8050$ ), thus avoiding the Hubble Deep Field and the shallower “Wide” region. We



**Fig. 5.** Comparison between input and measured fluxes of the detected galaxies matched with the IU, in the V606, H160, K<sub>s</sub>, and IRAC CH3 simulated band. The panels show the relative errors in flux measurement,  $(f_{\text{meas}} - f_{\text{true}})/f_{\text{true}}$ , as a function of the input true magnitude; so values above zero mean over-estimation of fluxes with respect to the true values, while values below zero mean under-estimation. The color coding gives the density of sources in every point of the plots.



**Fig. 6.** Number counts in the  $H$  detection band, relatively to an area of  $\sim 16$  sq. arcmin in order to be comparable with a region of the same area and with homogeneous depth from CANDELS GOODS-South DEEP (red shaded histogram). Blue line: IU counts; gray shadow: all detections, using SExtractor MAG\_AUTO as total flux estimate; black solid line: IU (true) flux of the detected sources with an IU match.

considered a region of the simulated  $H$  image with an equal area, and not containing large local objects, to compare the results fairly. After a cross-correlation between the coordinates of the sources in the IU and those in the catalogs of the detection, we found 3130 sources in the considered area out of 35,459 detected sources in the full FoV.

The comparison shows good agreement between the IU (blue line) and CANDELS (red shaded area) down to magnitude  $H \sim 26$ , after which the CANDELS counts start to deviate from IU, peaking at  $H \sim 27$  and falling at fainter magnitudes because of the incompleteness of the observed catalog. We also show two cases for the counts of the sources detected on the simulated image: the SExtractor measured MAG\_AUTO of all detected sources that have a match in the IU catalog (gray shaded area); and the true  $H$  magnitude of the same sources (black solid line). The distribution of the detected sources with true magnitudes is brighter with respect to the same distribution with measured values, as expected, since some flux in the faintest wings of the light profiles is always lost if a finite aperture is used to measure it. With the present configuration, the counts of the detected sources in the simulated image depart from IU counts at  $H \sim 25$  and peak at  $H \sim 26.5$ , slightly brighter than the peak at 27.0 of CANDELS counts. What is more interesting, though, is that the detections in the simulated image are well below the CANDELS counts already at  $H \sim 26$ .

We tried to pinpoint the origin of this discrepancy in the number counts at  $26 < H < 27$ . To check whether these inconsistencies could be due to an unfortunately sub-dense realization

**Table 4.** Number of galaxies with  $26 < H < 27$  in reference (Ref.) IU and EGG/GALSIM IU; their respective number of detected sources (Det.); and the detected sources matched with their respective IUs (match).

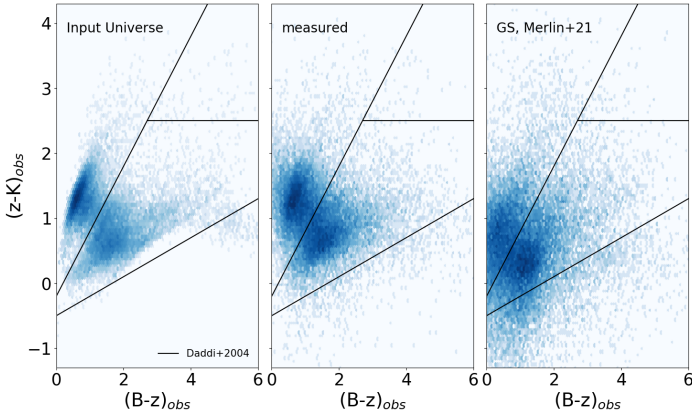
Sample $26 < H < 27$	$N_{\text{gal Ref.}}$	$N_{\text{gal GALSIM}}$
IU	12 829	14 883
Detected	11 318	14 514
Matched	11 135	14 299

of the simulation (since the reference light cone is constructed sampling a random area of each snapshot, see Sect. 2.1.2), we built and analyzed a tailored light cone selecting regions of the TNG snapshots having stellar mass density close to the average of the full simulated volume. However, the resulting counts were substantially similar, both in the IU and in the detections.

Then, to check whether the problem could originate from the input catalog or from the detection procedure, we created a new realization of our  $H$  image, including only sources with true magnitude in the relevant range, to avoid contamination and blending with brighter sources. We compared this tailored simulated image (which we tag as “reference IU”) with a similar image simulated with GALSIM (Rowe et al. 2015), which we fed with a mock galaxy catalog produced with the EGG software (Schreiber et al. 2017), again only including galaxies of the same input magnitude range. EGG catalogs are based on empirical relations calibrated on the observed CANDELS data, extrapolated to faint magnitudes; we created a catalog with a faint limiting magnitude ( $H = 31$ ) to ensure input completeness in our considered magnitude range. After having applied our post-processing pipeline, we run SExtractor with the very same parameters adopted for the detection on the reference image.

As shown in Table 4, the reference IU is less populated than the EGG IU, counting  $\sim 13.8\%$  fewer galaxies. On top of this,  $\sim 1500$  true sources are not detected in the FORECAST image, against only  $\sim 370$  missed on the GALSIM image.

We investigated the nature of unmatched undetected sources in both samples, distinguishing between sources that are not detected because they are blended with, or obscured by, other objects, and sources that may have very low surface brightness. Looking at the undetected sources in the reference IU (1694 sources), we found that 38% (638) is composed of blended sources, which are objects falling within larger and brighter galaxies in the full image; the remaining 62% (1061) consists of isolated galaxies, which can be either sources with low surface brightness or objects fragmented in multiple conglomerates of light which are individually too faint to be detected, the latter



**Fig. 7.**  $BzK$  diagrams for (left to right) (i) the IU fluxes of the detected sources, (ii) the measured colors of the same sources (matched with the IU), (iii) a sample of CANDELS sources from the deep region of the GS mosaic, with photometry from Merlin et al. (2021).

being a common kind of object generated by hydrodynamical models. In the EGG IU, the unmatched undetected sources are 51% (298) blended with close galaxies, and the remaining 49% (286) composed of isolated galaxies.

We conclude that the deficiency of detections in the FORECAST simulated image at  $26 < H < 27$  is due to two main factors: firstly, the IU is ab initio less dense than one generated using empirical prescriptions, possibly implying that the ILLUSTRISNG universe contains less faint galaxies than expected; then, a fraction of objects is undetected because of blendings and superpositions (which is reasonable given that it is very difficult to identify faint objects obscured by brighter ones in the real sky), and because of their fragmented morphologies, which do not show up in GALSIM galaxies given the analytic light profiles from which they are generated. We will investigate further this issue in future work.

### 3.2.4. $BzK$ diagram

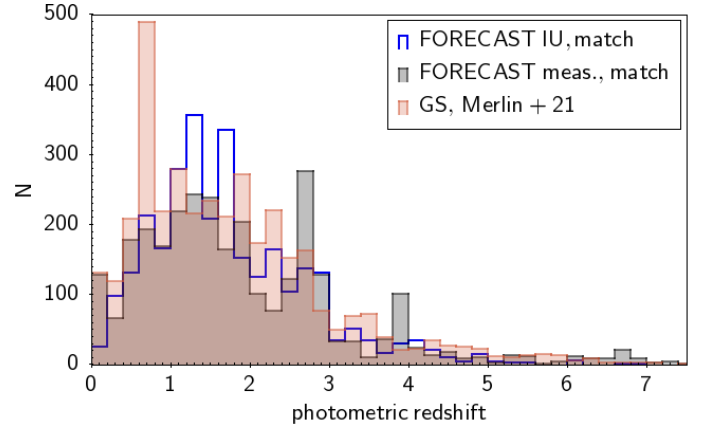
The  $BzK$  diagram (Daddi et al. 2004) is a widely used diagnostic color-color plot, useful to separate star-forming and quiescent galaxies using the observed  $B - z$  and  $z - K$  colors of  $1.4 \leq z \leq 2.5$  sources. The criterion is empirical, based on the spectroscopic redshifts from the K20 survey (Cimatti et al. 2002) and other publicly available data sets; however, synthetic stellar populations of both kinds (i.e., star-forming and passive) have been shown to indeed occupy the corresponding areas in this plot, when redshifted to  $1.4 < z < 2.5$ .

We plot the  $BzK$  colors of the simulated objects in Fig. 7, where the true colors (i.e., the ones obtained using IU fluxes) and the colors for a sample of the deep GS mosaic from Merlin et al. (2021) are also shown. The overall arrangement of the three distributions is indeed consistent, and we checked that star-forming sources at  $1.4 < z_{\text{true}} < 2.5$  are reasonably well isolated in the upper left region of the diagram.

### 3.3. Estimated physical properties

We finally checked the accuracy in the estimates of the redshift and the stellar mass of the detected galaxies.

To this aim, we performed a SED-fitting procedure with the code ZPHOT (Fontana et al. 2000), adopted in many studies (e.g., Castellano et al. 2016; Santini et al. 2015; Merlin et al. 2021). We



**Fig. 8.** Results of the SED-fitting photometric redshifts estimate using ZPHOT. The plot shows the distribution of the redshifts of the sources with a IU match, both estimated with ZPHOT (black shade) and from the input catalog (blue line); also shown is the distribution of a sample of sources from GOODS-South DEEP, taken from the ASTRODEEP catalog (Merlin et al. 2021, yellow shade). The histograms are normalized to allow for easier comparison.

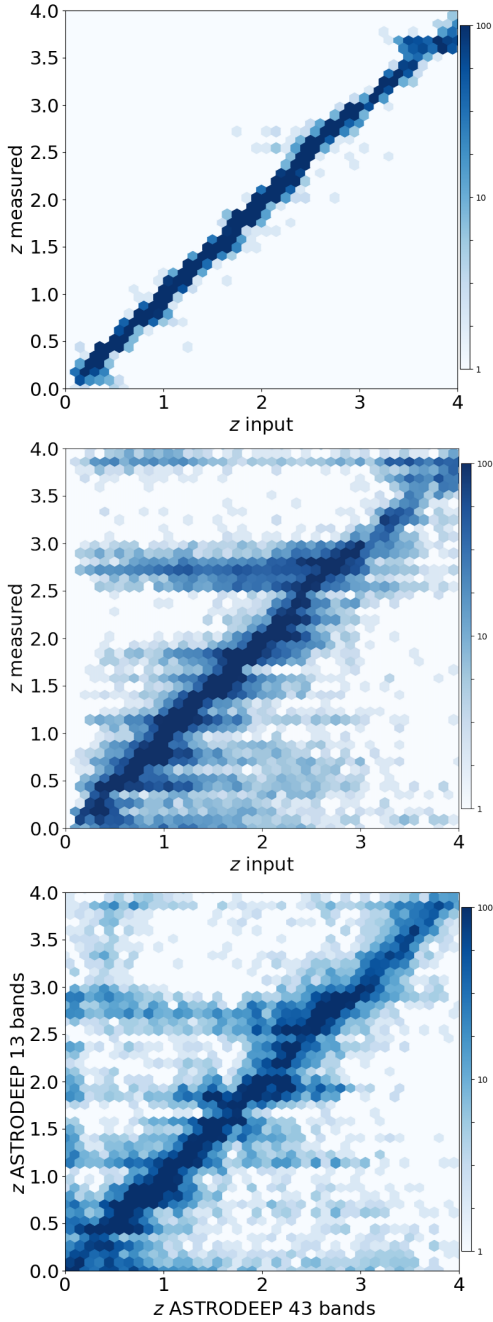
use a library of template galaxy SEDs identical to the one used in Merlin et al. (2021).

#### 3.3.1. Photometric redshift

The distribution of the measured redshifts is shown in Fig. 8, together with the ones from the IU and from ASTRODEEP.

We first checked that the library of models is sufficiently accurate, by estimating the redshifts using the true fluxes of the sources, while keeping the error budget of each source equal to the measured one. The result is in the top panel of Fig. 9. The agreement with the input redshifts is almost perfect, with a mean  $dz = (z_{\text{meas}} - z_{\text{true}})/(1.0 + z_{\text{true}}) = -0.012 \pm 0.022$  for objects with  $|dz| \leq 0.15$ ; the fraction of outliers (that have  $|dz| > 0.15$ ) is very low ( $\eta = 0.21\%$ ).

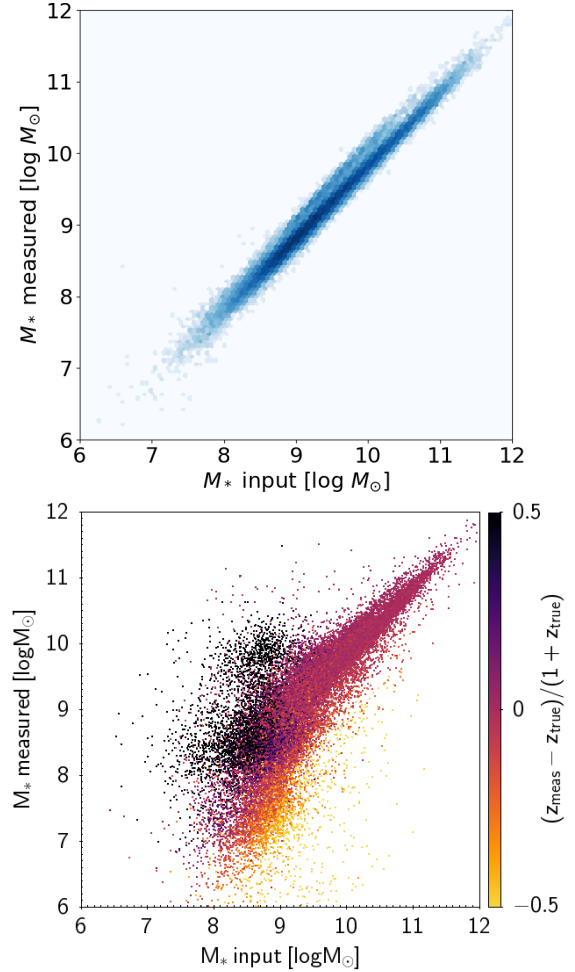
The distribution of the redshifts estimated using the measured fluxes is shown in the central panel of Fig. 9. In general, the estimate seems to be reasonably accurate. We note that the horizontal strips of catastrophic outliers are a typical feature of SED-fitting procedures, caused by a wrong interpretation of galaxy colors by the fitting algorithm, which interprets the red colors of a high redshift source as due to dust-obscured star formation in a low redshift object. The bulk of the objects are well recovered, with a mean  $dz = -0.011 \pm 0.055$ ; however, the fraction of outliers is high ( $\eta = 25.2\%$ ). It suggests that most of the uncertainties and errors are caused by the scatter introduced by the photometric estimates rather than simply by the small number of bands. Of course, a larger number of bands help minimizing the uncertainties in the fit. To further check this issue, we made a final test on the ASTRODEEP catalogs, but using only the same 13 bands simulated for this work to estimate photometric redshift. The results, performed using the same libraries of templates, are in the bottom panel of Fig. 9; however, it must be kept in mind that now we do not have a “true” value of the input redshift, but only the best estimates from spectroscopic or photometric data. While the results are better than the ones for the simulation ( $dz = -0.003 \pm 0.055$ ), there are still a large amount of scatter and outliers ( $\eta = 20.8\%$ ), caused by the limited number of bands.



**Fig. 9.** Results of the SED-fitting photometric redshifts estimate using (from top to bottom): (i) ZPHOT and the true fluxes of the input catalog; (ii) ZPHOT and fluxes measured on the  $H$  band simulated image; (iii) ZPHOT on GOODS-South ASTRODEEP data. The plot shows the estimated redshifts versus the true redshifts from the input catalog (top and central panel); the redshifts estimated with the 13 bands used in the simulations versus the ones in the ASTRODEEP 43 bands catalog (Merlin et al. 2021, bottom panel).

### 3.3.2. Galaxy stellar mass

The top panel in Fig. 10 shows the comparison between the stellar masses estimated with ZPHOT, fixing the redshift to the IU values and using the IU fluxes, and the true masses from the input catalog (defined as the sum of the masses of all the SSPs belonging to a given subhalo). Again, in this case the estimated values are in tight agreement with the true values in the mass estimates. We then replicate the same plot, but using measured



**Fig. 10.** Results of the SED-fitting stellar mass estimate using (from top to bottom): (i) ZPHOT and measured fluxes and redshifts, color coded by error in redshift estimate,  $(z_{\text{meas}} - z_{\text{true}})/(1 + z_{\text{true}})$ . The plot shows the estimated stellar mass versus the true one from the input catalog.

redshifts and fluxes to estimate the masses of the sources with ZPHOT (bottom panel of Fig. 10). Here the points are color-coded by the error the estimated redshift. The agreement is still good in general, although the scatter is quite large toward fainter masses. We note two interesting populations of sources: a group having masses underestimated by one order of magnitude, in the range  $10^8 < M_{*,\text{true}} < 10^9$ , and another group with masses overestimated by one order of magnitude ( $10^8 < M_{*,\text{true}} < 10^9$ ). We see that the vast majority of the sources with underestimated masses also have underestimated redshifts (often by a factor  $(z_{\text{meas}} - z_{\text{true}})/z_{\text{true}} \leq -0.5$ ): being considered as closer to the observer than they really are, they must be fitted with a low stellar mass to match the measured fluxes. A specular line of reasoning can be applied to objects with overestimated masses. We also note a group of sources with overestimated redshift which has masses correctly estimated (the dark points lying on the bisector of the distribution). We found that this feature is due to the underestimation of the  $H$  flux for faint sources (see Sect. 3.2.2), which causes their distance to be overestimated.

We also checked that estimating the masses using the measured fluxes but the IU true redshifts a less evident but non-negligible scatter is still present. We conclude that it is to be attributed to the uncertainties introduced by the photometric

measurements, which will deserve further analysis in future work.

#### 4. Code and data release

The FORECAST code (see Appendix B for further details) is available to the community on our website<sup>3</sup>, together with the CANDELS-like dataset analyzed in this work and a simulated dataset emulating JWST observations, which is described in this section.

##### *Data release of a JWST-like survey*

Together with the CANDELS mock observations, we have also produced and made public a second dataset, this time emulating JWST observations of the same FoV we presented in the previous section. The set of data consists of a galaxy catalog and ten astronomical images. To construct this dataset, we extended the light cone realized for the CANDELS emulation up to  $z_s = 20$ , again exploiting the ILLUSTRISTNG simulation. The pixel scale is 0.031 arcsec (the typical value of JWST short wavelength detectors), yielding images on a grid of  $\sim 750$  million pixels. We emulated ten JWST bands: eight from NIRCam (F090W, F115W, F150W, F200W, F277W, F356W, F410M, F444W), and two from MIRI (F560W, F770W). For the post-processing we used the PSF models provided by STScI in the WebbPSF webpage<sup>4</sup>. For this emulation, we also decided to change the synthetic stellar population model by adopting Gutkin et al. (2016), which includes rest-frame ultraviolet and optical nebular emission from HII regions in star-forming galaxies in a wide range of chemical compositions.

The image in the JWST/NIRCam F090W filter is simulated with resolution and limit magnitude adopted by Merlin et al. (2022) for the JWST GLASS survey (Treu et al. 2022). Additionally, the remaining seven broad-band images in NIRCam filters emulate the JWST CEERS survey (Finkelstein et al. 2023); we created two images in the CEERS JWST/MIRI bands with resolution and depth adopted by Papovich et al. (2023). We did not perform any analysis on this dataset, leaving it to future work. All the released data products are available on our website<sup>5</sup>.

#### 5. Summary and conclusions

We have presented FORECAST, a new software package that performs forward modeling of the output of cosmological hydrodynamical simulations to create realistic synthetic astronomical images. Starting from the physical properties of the simulated stellar resolution elements provided in the output snapshots of a hydrodynamical simulation, the software computes their expected fluxes, accounting for  $k$  correction, attenuation by dust and by the intergalactic medium, and arranges them to produce images to which background noise, PSF smoothing and potentially other observational features can be added. The simulated galaxies are built particle by particle, and therefore they do not have smooth, analytical light profiles; instead, they have realistic morphologies and fluxes, computed from their complex star formation histories. The simulated images can be processed and

**Table 5.** Summary of the instrumental PSF and depths adopted for the image simulations in the JWST filters from Merlin et al. (2022); Finkelstein et al. (2023); Papovich et al. (2023).

Instrument	Filter	PSF(arcsec)	$5\sigma$ depth AB
NIRCam	F090W	0.035	28.78 <sup>(a)</sup>
	F115W	0.066	29.20 <sup>(b)</sup>
	F150W	0.070	29.00 <sup>(b)</sup>
	F200W	0.077	29.20 <sup>(b)</sup>
	F277W	0.123	29.20 <sup>(b)</sup>
	F356W	0.142	29.20 <sup>(b)</sup>
	F410M	0.155	28.40 <sup>(b)</sup>
	F444W	0.161	28.60 <sup>(b)</sup>
MIRI	F560W	0.240	26.50 <sup>(c)</sup>
	F770W	0.280	27.10 <sup>(c)</sup>

**Notes.** <sup>(a)</sup>The limiting magnitude is measured in a 0.1" radius aperture, from Merlin et al. (2022). <sup>(b)</sup>The limiting magnitudes are measured in a 0.1" radius aperture, from Finkelstein et al. (2023). <sup>(c)</sup>These depths are derived from tests on the MIRI images. They are measured in 0.45" diameter apertures (see Papovich et al. 2023 for further details).

analyzed with the same methods and tools used in real data analysis, and directly compared in a fully consistent way to the results from real observational data. FORECAST is a flexible tool that can produce realistic images, enabling the analysis of possible systematics and biases arising in observations due to image processing, the choice and limitations of the algorithms used to detect, deblend, and measure galaxy fluxes, as well as the physical assumptions in a SED-fitting procedure.

To test FORECAST, we built a light cone between  $z = 0$  and  $z = 7$ , emulating the GOODS-South CANDELS field, creating scientific images in eight HST bands (ACS: B435, V606, I814, and Z850; and WFC3: Y105, J125, JH140, and H160), one VLT band (HAWK-I  $K_s$ ), and four *Spitzer* bands (IRAC CH1, CH2, CH3, CH4), considering the deep region of the field as a reference for the synthetic image depths.

The simulated field of view has an area of 200 sq. arcmin, and the fluxes are mapped on a grid of 200 million of pixels, resulting in a pixel scale of 0.06 arcsec, a typical resolution of real HST bands (we created all bands with the same pixel scale, avoiding rebinning procedures). This light cone includes a great diversity of galaxies over a large range of mass and star formation rates, metallicities, ages, star formation histories generated by the complex interplay of the diverse astrophysical processes (cooling, star formation, feedback, and dynamical evolution and interactions). The final products, the images in the 13 bands, are generated adding noise and PSF to the outputs of FORECAST with our post-processing procedure. We found that the simulated images offer a realistic representation of many observational features; we verified this using standard techniques used for the photometric analysis of real images.

We performed the detection on the simulated  $H$  band using SEXTRACTOR, and we then measured the fluxes of the detected sources using aperture photometry with A-PHOT on the 13 simulated images. The flux of the sources is generally well recovered in all bands, with a slight underestimation at faint magnitudes due to the measurements on the  $H$  band, which is used to derive the fluxes in the remaining bands.

We then checked the number counts of the simulated sources, comparing the counts as a function of the  $H$  magnitude between a sample of simulated objects (with fluxes taken both from the

<sup>3</sup> <http://www.astrodeep.eu/FORECAST>

<sup>4</sup> <https://jwst-docs.stsci.edu/jwst-near-infrared-camera/nircam-performance/nircam-point-spread-functions>

<sup>5</sup> <http://www.astrodeep.eu/FORECAST>

Input Universe and from the detection process), and a sample of objects in the CANDELS GOODS-South area. We found that the number of objects detected on the simulated image is consistent with the Input Universe (IU) up to  $\text{mag}H \sim 25$ , after which their counts begin to be less than expected, also compared with the trend in CANDELS. We determined that there are two contributing factors to the discrepancy between the counts in the range  $26 < H < 27$ : firstly, the ILLUSTRISTNG IU seems less populated than expected; secondly, there is a large fraction of galaxies that are either blended with larger objects or too faint to be detected due to their fragmented morphologies. However, the identification of significant overdensities in the GOODS-South field, spanning redshifts  $z = 0.6\text{--}3.7$  (Castellano et al. 2007, 2011; Salimbeni et al. 2009; Kurk et al. 2009; Kang & Im 2009), suggests that the galaxy counts in the CANDELS catalog might have been impacted by this clustering, potentially resulting in an increase in the galaxy counts.

We estimated the physical properties of the galaxies detected on the simulated images via SED-fitting. While the redshifts are perfectly recovered in ideal conditions (i.e., using true IU fluxes), a noticeable amount of scatter is introduced using the measured fluxes. The accuracy is almost perfect also in the estimate of the stellar masses if the photometry is ideal (i.e., fitting the true fluxes at the true redshifts). However, a mild scatter emerges if the measured fluxes are fitted at the measured redshifts, mostly caused by the error committed in measuring the  $H$  flux at faint magnitude, which is spread in the other bands; for a subsample of sources, an error of one order of magnitude in the estimate mainly depends on the propagation of the error on the photo- $z$  estimate.

We want to remark that the realization of these synthetic images is the first attempt of forward modeling as much physics as possible from hydrodynamical simulations, and the tests performed in this work must be intended as a first quality check.

Future work will include (i) implementing additional effects in the light cone: adding Milky Way stars and local objects, Active Galactic Nuclei, the absorption due to Milky Way gas and dust, the effect of lensing; (ii) implementing options to allow for more flexibility, e.g., giving the user the possibility to choose a preferential position to extract the light cone.

We make the simulated CANDELS dataset publicly available. We also release a set of images simulated in ten JWST bands, and the corresponding Input Universe catalog containing simulated physical properties and simulated true fluxes of the galaxies.

As new upcoming observational instruments will allow us to probe the Universe to an unexplored extent, numerical tools like FORECAST will help us to capture the significance of their exploration, improving the synergism between observations and theory. The next few years will revolutionize our understanding of the Cosmos and will make us more aware about the Universe we inhabit.

*Acknowledgements.* The IllustrisTNG simulations were undertaken with computational time awarded by the Gauss Centre for Supercomputing (GCS) under GCS Large-Scale Projects GCS-ILLU and GCS-DWAR on the GCS share of the supercomputer Hazel Hen at the High-Performance Computing Center Stuttgart (HLRS), as well as on the machines of the Max Planck Computing and Data Facility (MPCDF) in Garching, Germany. We thank the CINECA award under the ISCR initiative, for the availability of high-performance computing resources and support. We thank the INAF computing system PLEIADI, for the availability of high-performance computing resources and support. Carlo Giocoli acknowledges support from the PRIN-MIUR 2017 WSCC32 ZOOMING, the ASI no. 2018-23-HH.0, the INAF grant under the “Bando PrIN 2019”, PI: Viola Allevato, the INAF theory grant 2022: Illuminating Dark Matter using Weak Lensing by Cluster Satellites, PI: Carlo Giocoli.

## References

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2003, *AJ*, **126**, 2081
- Anders, E., & Grevesse, N. 1989, *Geochim. Cosmochim. Acta*, **53**, 197
- Aoyama, S., Hou, K.-C., Hirashita, H., Nagamine, K., & Shimizu, I. 2018, *MNRAS*, **478**, 4905
- Baes, M., Davies, J. I., Dejonghe, H., et al. 2003, *MNRAS*, **343**, 1081
- Baes, M., Verstackpen, J., De Looze, I., et al. 2011, *ApJS*, **196**, 22
- Bagla, J. S. 2002, *J. Astrophys. Astron.*, **23**, 185
- Barrientos Acevedo, D., van der Wel, A., Baes, M., et al. 2023, *MNRAS*, **524**, 907
- Behroozi, P., Conroy, C., Wechsler, R. H., et al. 2020, *MNRAS*, **499**, 5702
- Bekki, K. 2015, *MNRAS*, **449**, 1625
- Bernyk, M., Croton, D. J., Tonini, C., et al. 2016, *ApJS*, **223**, 9
- Bertin, E. 2009, *Mem. Soc. Astron. Italiana*, **80**, 422
- Bertschinger, E., & Gelb, J. M. 1991, *Comput. Phys.*, **5**, 164
- Blaizot, J., Wadadekar, Y., Guiderdoni, B., et al. 2005, *MNRAS*, **360**, 159
- Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, **200**, 13
- Bravo, M., Lagos, C. d. P., Robotham, A. S. G., Bellstedt, S., & Obreschkow, D. 2020, *MNRAS*, **497**, 3026
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Bryan, G. L., Norman, M. L., O’Shea, B. W., et al. 2014, *ApJS*, **211**, 19
- Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, *ApJ*, **429**, 582
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
- Castellano, M., Salimbeni, S., Trevese, D., et al. 2007, *ApJ*, **671**, 1497
- Castellano, M., Pentericci, L., Mencì, N., et al. 2011, *A&A*, **530**, A27
- Castellano, M., Yue, B., Ferrara, A., et al. 2016, *ApJ*, **823**, L40
- Castellano, M., Fontana, A., Treu, T., et al. 2022, *ApJ*, **938**, L15
- Chabrier, G. 2003, *PASP*, **115**, 763
- Charlot, S., & Fall, S. M. 2000, *ApJ*, **539**, 718
- Choi, Y.-Y., Park, C., Kim, J., et al. 2010, *ApJS*, **190**, 181
- Cimatti, A., Pozzetti, L., Mignoli, M., et al. 2002, *A&A*, **391**, A1
- Cochrane, R. K., Hayward, C. C., Anglés-Alcázar, D., & Somerville, R. S. 2023, *MNRAS*, **518**, 5522
- Colless, M. 1999, *Philos. Trans. Roy. Soc. Lond. A*, **357**, 105
- Croft, R. A. C., Di Matteo, T., Davé, R., et al. 2001, *ApJ*, **557**, 67
- Cui, W., Davé, R., Peacock, J. A., Anglés-Alcázar, D., & Yang, X. 2021, *Nat. Astron.*, **5**, 1069
- Curtis-Lake, E., Carniani, S., Cameron, A., et al. 2023, *Nat. Astron.*, **7**, 622
- Daddi, E., Cimatti, A., Renzini, A., et al. 2004, *ApJ*, **617**, 746
- da Silva, A. C., Barbosa, D., Liddle, A. R., & Thomas, P. A. 2000, *MNRAS*, **317**, 37
- da Silva, A. C., Barbosa, D., Liddle, A. R., & Thomas, P. A. 2001a, *MNRAS*, **326**, 155
- da Silva, A. C., Kay, S. T., Liddle, A. R., et al. 2001b, *ApJ*, **561**, L15
- Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, *MNRAS*, **486**, 2827
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, **292**, 371
- Devriendt, J. E. G., & Guiderdoni, B. 2000, *A&A*, **363**, 851
- Devriendt, J., Rimes, C., Pichon, C., et al. 2010, *MNRAS*, **403**, L84
- Di Cesare, C., Graziani, L., Schneider, R., et al. 2023, *MNRAS*, **519**, 4632
- Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, *MNRAS*, **399**, 497
- Donnan, C. T., McLeod, D. J., Dunlop, J. S., et al. 2023, *MNRAS*, **518**, 6011
- Draine, B. T., & Lee, H. M. 1984, *ApJ*, **285**, 89
- Draine, B. T., Li, A., Hensley, B. S., et al. 2021, *ApJ*, **917**, 3
- Drakos, N. E., Villaseñor, B., Robertson, B. E., et al. 2022, *ApJ*, **926**, 194
- Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, *Astron. Geophys.*, **50**, 5.12
- Dunne, L., Gomez, H. L., da Cunha, E., et al. 2011, *MNRAS*, **417**, 1510
- Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2023, *ApJ*, **946**, L13
- Fontana, A., D’Odorico, S., Poli, F., et al. 2000, *AJ*, **120**, 2206
- Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, *AJ*, **111**, 1748
- Genel, S., Vogelsberger, M., Springel, V., et al. 2014, *MNRAS*, **445**, 175
- Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, **600**, L93
- Giocoli, C., Metcalf, R. B., Baldi, M., et al. 2015, *MNRAS*, **452**, 2757
- Gordon, K. D., Misselt, K. A., Witt, A. N., & Clayton, G. C. 2001, *ApJ*, **551**, 269
- Graziani, L., Schneider, R., Ginolfi, M., et al. 2020, *MNRAS*, **494**, 1071
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35
- Guiderdoni, B., & Rocca-Volmerange, B. 1987, *A&A*, **186**, 1
- Guidi, G., Scannapieco, C., Walcher, J., & Gallazzi, A. 2016, *MNRAS*, **462**, 2046
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, **207**, 24
- Gutkin, J., Charlot, S., & Bruzual, G. 2016, *MNRAS*, **462**, 1757
- Harikane, Y., Ouchi, M., Oguri, M., et al. 2023, *ApJS*, **265**, 5
- Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, *ApJ*, **457**, L51
- Hopkins, P. F. 2015, *MNRAS*, **450**, 53
- Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, *MNRAS*, **442**, 1805
- Ji, Z., & Giavalisco, M. 2022, *ApJ*, **935**, 120

- Jonsson, P. 2006, *MNRAS*, 372, 2
- Jonsson, P., Groves, B. A., & Cox, T. J. 2010, *MNRAS*, 403, 17
- Kang, E., & Im, M. 2009, *ApJ*, 691, L33
- Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, *MNRAS*, 467, 4739
- Kitzbichler, M. G., & White, S. D. M. 2007, *MNRAS*, 376, 2
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36
- Kurk, J., Cimatti, A., Zamorani, G., et al. 2009, *A&A*, 504, 331
- Laigle, C., Davidzon, I., Ilbert, O., et al. 2019, *MNRAS*, 486, 5104
- Lilly, S., Le Fèvre, O., Renzini, A., et al. 2007, *ApJS*, 172, 70
- Liu, Z., Morishita, T., & Kodama, T. 2023, *ApJ*, submitted, [arXiv:2305.10944]
- Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, *ApJ*, 701, 1765
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, *MNRAS*, 480, 5113
- McKinnon, R., Torrey, P., & Vogelsberger, M. 2016, *MNRAS*, 457, 3775
- McKinnon, R., Vogelsberger, M., Torrey, P., Marinacci, F., & Kannan, R. 2018, *MNRAS*, 478, 2851
- Merlin, E., Fontana, A., Ferguson, H. C., et al. 2015, *A&A*, 582, A15
- Merlin, E., Bourne, N., Castellano, M., et al. 2016, *A&A*, 595, A97
- Merlin, E., Pilo, S., Fontana, A., et al. 2019, *A&A*, 622, A169
- Merlin, E., Castellano, M., Santini, P., et al. 2021, *A&A*, 649, A22
- Merlin, E., Bonchi, A., Paris, D., et al. 2022, *ApJ*, 938, L14
- Merlin, E., Castellano, M., Huertas-Company, M., & Bretonniere, H. 2023, *A&A*, 671, A101
- Merson, A. I., Baugh, C. M., Helly, J. C., et al. 2013, *MNRAS*, 429, 556
- Misselt, K. A., Gordon, K. D., Clayton, G. C., & Wolff, M. J. 2001, *ApJ*, 551, 277
- Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, *ApJ*, 808, 101
- Morishita, T., & Stiavelli, M. 2023, *ApJ*, 946, L35
- Murray, N. 2011, *ApJ*, 729, 133
- Murray, N., Quataert, E., & Thompson, T. A. 2010, *ApJ*, 709, 191
- Naidu, R. P., Oesch, P. A., van Dokkum, P., et al. 2022, *ApJ*, 940, L14
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, 477, 1206
- Nanni, L., Thomas, D., Trayford, J., et al. 2023, *MNRAS*, 522, 5479
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, 475, 624
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *Comput. Astrophys. Cosmol.*, 6, 2
- Oke, J. B. 1974, *ApJS*, 27, 21
- Overzier, R., Lemson, G., Angulo, R. E., et al. 2013, *MNRAS*, 428, 778
- Papovich, C., Cole, J. W., Yang, G., et al. 2023, *ApJ*, 949, L18
- Park, C., Choi, Y.-Y., Kim, J., et al. 2012, *ApJ*, 759, L7
- Parsotan, T., Cochrane, R. K., Hayward, C. C., et al. 2021, *MNRAS*, 501, 1591
- Pentericci, L., McLure, R. J., Garilli, B., et al. 2018, *A&A*, 616, A174
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, *MNRAS*, 475, 648
- Pillepich, A., Springel, V., Nelson, D., et al. 2018b, *MNRAS*, 473, 4077
- Planck Collaboration XIII. 2016, *A&A*, 594, A13
- Plazas, A. A., Meneghetti, M., Maturi, M., & Rhodes, J. 2019, *MNRAS*, 482, 2823
- Price, S. H., Kriek, M., Feldmann, R., et al. 2017, *ApJ*, 844, L6
- Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2014, *A&A*, 563, A31
- Roberts-Borsani, G., Morishita, T., Treu, T., et al. 2022, *ApJ*, 938, L13
- Robertson, B. E., Tacchella, S., Johnson, B. D., et al. 2023, *Nat. Astron.*, 7, 611
- Robitaille, T. P. 2011, *A&A*, 536, A79
- Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, *MNRAS*, 449, 49
- Roncarelli, M., Moscardini, L., Tozzi, P., et al. 2006, *MNRAS*, 368, 74
- Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015, *Astron. Comput.*, 10, 121
- Salimbeni, S., Castellano, M., Pentericci, L., et al. 2009, *A&A*, 501, 865
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, *ApJ*, 801, 97
- Scaramella, R., Cen, R., & Ostriker, J. P. 1993, *ApJ*, 416, 399
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, 521
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2017, *A&A*, 602, A96
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, 172, 1
- Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, *MNRAS*, 452, 575
- Snyder, G. F., Lotz, J. M., Rodriguez-Gomez, V., et al. 2017, *MNRAS*, 468, 207
- Snyder, G. F., Peña, T., Yung, L. Y. A., et al. 2023, *MNRAS*, 518, 6318
- Somerville, R. S., Olsen, C., Yung, L. Y. A., et al. 2021, *MNRAS*, 502, 4858
- Springel, V. 2005, *MNRAS*, 364, 1105
- Springel, V. 2010, *MNRAS*, 401, 791
- Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, *MNRAS*, 328, 726
- Springel, V., Di Matteo, T., & Hernquist, L. 2005, *MNRAS*, 361, 776
- Springel, V., Pakmor, R., Pillepich, A., et al. 2017, *MNRAS*, 475, 676
- Taghizadeh-Popp, M., Fall, S. M., White, R. L., & Szalay, A. S. 2015, *ApJ*, 801, 14
- Taylor, M. B. 2005, in *ASP Conf. Ser.*, 347, *Astronomical Data Analysis Software and Systems XIV*, eds. P. Shopbell, M. Britton, & R. Ebert, 29
- Teyssier, R. 2002, *A&A*, 385, 337
- Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, *ApJ*, 783, 85
- Treu, T., Roberts-Borsani, G., Bradac, M., et al. 2022, *ApJ*, 935, 110
- Vogelsberger, M., Genel, S., Sijacki, D., et al. 2013, *MNRAS*, 436, 3031
- Vogelsberger, M., Marinacci, F., Torrey, P., et al. 2018, *MNRAS*, 474, 2073
- Vogelsberger, M., Nelson, D., Pillepich, A., et al. 2020, *MNRAS*, 492, 5167
- Wadsley, J. W., Stadel, J., & Quinn, T. 2004, *New Astron.*, 9, 137
- Weinberger, R., Springel, V., Pakmor, R., et al. 2018, *MNRAS*, 479, 4056
- Xu, G. 1995, *ApJS*, 98, 355
- Yan, H., Cohen, S. H., Windhorst, R. A., et al. 2023, *ApJ*, 942, L8
- Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., et al. 2023, *MNRAS*, 519, 1578

## Appendix A: Mock light cone with IllustrisTNG100

In order to convert the output of the ILLUSTRISTNG100-1 simulation into mock astronomical images, we extracted the physical properties of stellar particles and gas cells required by FORECAST, as described in Sects. 2.1.3 and 2.1.4. In particular, for stellar particles ("PARTTYPE4") we extracted "Coordinates"; "Masses"; "GFM\_InitialMass"; "GFM\_Metallicity"; "GFM\_StellarFormationTime", the latter to compute the ages of the SSPs. FORECAST also requires information about the subhalo membership of the particles (as it is defined by the simulation procedure), but for TNG it is not stored in the snapshots for stellar particles and gas cells, while subhalos in subhalo catalogs have information about their particle and cell members. Thus, we reconstructed the membership of stellar particles and gas cell in reverse, with a designed algorithm that uses SUBFIND subhalo catalogs and Offset files, exploiting the specific organization of halos and subhalos within the catalog files.

Concerning gas cells ("PARTTYPE0"), we extracted their "Coordinates"; "Masses"; "Density"; "GFM\_Metallicity"; "ElectronAbundance", which is gas cell fractional electron number density  $x_e$ , with respect to the total hydrogen number density, that is  $n_e = x_e n_H$ , and "InternalEnergy", which is gas cell internal (thermal) energy per unit mass  $u$ , both needed to derive the neutral hydrogen column density  $N_{HI}$  (see Eq. 6). The IDs of gas cells are exploited to derive their subhalo membership, as done for star particles.<sup>6</sup>

The neutral hydrogen column density within each gas cell is not available in the simulation output, so we estimated it as follows, taking advantage of already accessible properties. The fraction of neutral hydrogen with respect to the total hydrogen number density within each gas cell, "NEUTRALHYDROGENABUNDANCE" alias  $x_{HI}$ , necessary to compute the neutral hydrogen column density, is available in the ILLUSTRISTNG output only at certain snapshots, the so-called "full snapshots"; the remaining "mini snapshots" only have a subset of particle fields available, and they do not include  $x_{HI}$ . To overcome this shortage, we perform a fourth-degree least squares polynomial fit on the relation between the neutral hydrogen-to-gas mass ratio  $M_{HI}/M_g$  and the temperature of the gas  $T_g$  in the "full snapshots"

When performing the polynomial fit to the distribution of gas cells' values in the "full snapshots", we find that the red curve in Fig. A.1, representing the best fit at  $z = 0$ , fits the relation up to  $z = 7$  with sufficient accuracy, so we decide to use the polynomial coefficients of the best fit at  $z = 0$  to derive  $M_{HI}/M_g$  throughout the full light cone (in each snapshot). To maintain the fitting curve within a valid physical range, we set to 0 the values of  $\log(M_{HI}/M_g)$  that exceed 0. The best fit coefficients of the relation between  $M_{HI}/M_g$  and  $T_g$  are given in Table A.1.

Gas number density of each gas cell is derived as  $n_g = \rho_g/m_p$ , where  $m_p$  is the proton mass. Therefore, it is possible to compute the neutral hydrogen number density  $n_{HI} = x_{HI} x_H n_g$ , from which we derive the neutral hydrogen column density of each gas cell, assuming their volume as cubic with side  $L_g$

$$N_{HI} = n_{HI} \cdot L_g. \quad (\text{A.1})$$

<sup>6</sup> The available fields for stellar particles and gas cells, their units, and descriptions are available at <https://www.tng-project.org/data/docs/specifications/>.

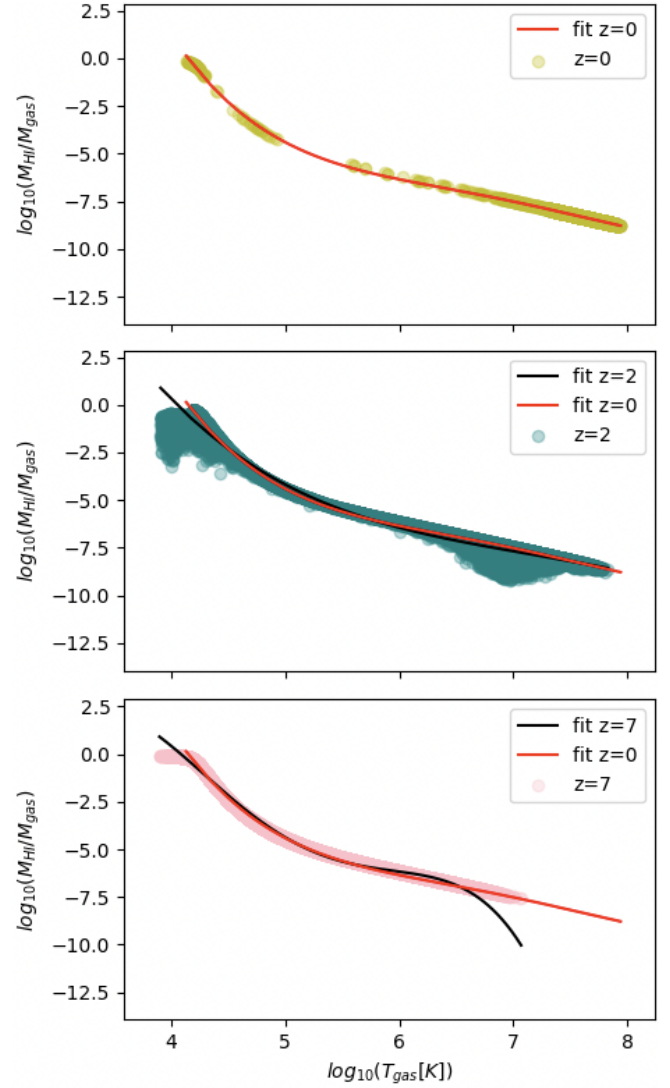


Fig. A.1: Scaling relation between neutral hydrogen-to-gas mass ratio  $M_{HI}/M_g$  and the temperature of the gas  $T_g$  in the "full snapshots" of the ILLUSTRISTNG100 simulation, where the neutral hydrogen fraction is available ( $z = 0, 2, 7$ ). The red curve is the best fourth-grade polynomial fit at  $z = 0$ ; black curves are best fits for the data points (dots) at the redshift of each panel.

## Appendix B: FORECAST code

FORECAST is a robust and flexible code. Its main body, written in C and C++, is supported by independent libraries to make the code more readable and user-friendly. It requires the following C/C++ standard libraries: gsl<sup>7</sup>, openBLAS<sup>8</sup>, LAPACK<sup>9</sup>, CCfits<sup>10</sup>, CFITSIO<sup>11</sup>, FFTW<sup>12</sup>, Eigen<sup>13</sup>, Armadillo<sup>14</sup>, H5C++<sup>15</sup>, HDF5 C++<sup>16</sup>, and a gcc compiler.

<sup>7</sup> <https://www.gnu.org/software/gsl/>

<sup>8</sup> <https://www.openblas.net/>

<sup>9</sup> <http://www.netlib.org/lapack/>

<sup>10</sup> <https://heasarc.gsfc.nasa.gov/fitsio/CCfits/>

<sup>11</sup> <https://heasarc.gsfc.nasa.gov/fitsio/>

<sup>12</sup> <https://www.fftw.org/>

<sup>13</sup> <https://eigen.tuxfamily.org/>

<sup>14</sup> <http://arma.sourceforge.net/>

<sup>15</sup> <https://ess-dmcs.github.io/h5c++/>

<sup>16</sup> <https://www.hdfgroup.org/>

Table A.1: Fourth degree polynomial fit coefficients at  $z = 0$ .

$\frac{M_{HI}}{M_g} = a_1 T_g^4 + a_2 T_g^3 + a_3 T_g^2 + a_4 T_g + a_5$					
$z$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
0	2.31105e-02	-8.91893e-01	1.14885e+01	-6.27073e+01	1.18956e+02

In the input configuration file of the code, the user chooses the image simulation parameters (e.g., the dimension of the FoV, the filters, the hydrodynamical simulation; see Sect. 2.1.1).

The code requires the input files of the chosen hydrodynamical simulation to build the light cone and produce the final images. The data products of the numerous currently available simulations are organized differently, changing from one to another simulation, and are stored with different formats; as example ILLUSTRISTNG stores a single snapshot in multiple .hdf5 files, while in the EAGLE simulation (Schaye et al. 2015) the snapshots are available for public download via an SQL web interface. Thus, the code requires these data products to be uniformed in a specific format in order to be easily read and processed.

The code pipeline consists of four modules. The architecture of each module is not intrinsically parallel (e.g., it does not exploit MPI protocols), but it has been designed to allow the user to independently run it on multiple snapshots to realize multiple light-cone partitions simultaneously.

We release a beta version of the code, which can be read and improved by the scientific community with a request for access to its source through our website<sup>17</sup>.

### B.0.1. Pipeline

The FORECAST code, currently available in beta version, is structured into four interconnected modules, where each module relies on the output of the previous one. The first and third modules require as input file the snapshots of a hydrodynamical simulation to extract relevant properties of the simulated stellar and gas resolution elements. Currently, the code reads input files in the ILLUSTRISTNG format. Users are required to convert their data into the TNG format (.hdf5 files with the same column names as TNG columns) to ensure the effective utilization of the code. Future updates to the code will include additional scripts to read input files from multiple hydrodynamical simulations in different formats, such as the EAGLE Project (Schaye et al. 2015) and the Simba Simulation (Davé et al. 2019) data products.

The first module handles the construction of the light cone, exploiting the data products of a chosen hydrodynamical simulation (see Sect. 2.1.2). The resulting output is an ASCII file containing the properties of the SSP elements within the field of view, including their IDs, coordinates, redshift, and physical characteristics. This file serves as the input for the subsequent module. This step might be skipped if a user already has their light cone, as long as the input file for the next module is written in the proper format.

The second module computes the dust-free flux of each SSPs included in the FoV (see Sect. 2.1.3). It assembles an ASCII file with the properties of the star particles and their dust-free fluxes in chosen filters.

The third module addresses the computation of dust-corrected fluxes and it requires the data products of the hydrodynamical simulation to extract the properties of gas resolution elements

belonging to the sources included in the FoV. In output it is given the same file produced by the previous module, also including dust-corrected fluxes for stellar particles, and the gas mass-weighted mean of the gas metallicity and the neutral hydrogen column density. The final module adds the IGM correction to dust-corrected fluxes, producing the final output catalog, which includes the physical properties of the stellar particles and their corrected fluxes, and the mean properties of the gas.

An independent C++ script handles the arrangement of the fluxes on a grid of pixels with  $N_{pix}$ -per-side chosen by the user. Two additional independent scripts, written in python, are made available (i) to build the galaxy catalog, in ASCII format, from the particle catalog given in output by FORECAST (see Sect. B.0.2 for a full description of the output); (ii) to post-process the FORECAST images with our noise and PSF pipeline (see Sect. 2.2).

### B.0.2. Output

The output of the code is the catalog including the physical properties and the true fluxes of the simulated stellar particles. It is used to build the galaxy catalog (see Appendix C for a full description of the fields included in the galaxy catalog). The catalogs, both the particles and the galaxy ones, have different sizes depending on the number of particles (or galaxies) included in the FoV, and typically grow in size as the redshift increases since more structures are included. The total size of output files is  $\sim 72$  GB. The output images are recorded on 16-bit floating-point .fits files. Each plane (projection on a bi-dimensional map of fluxes from a volume of the Universe included in the field of view, in a redshift range) occupies  $\sim 3$  GB, while the size of the final stacked image is  $\sim 5$  GB.

### B.0.3. Memory consumption and running time

We performed some tests and realized the light cone and the images presented in this work on the GALILEO100 supercomputer located at CINECA<sup>18</sup>. The peak of memory consumption is reached during the post-processing of the dust, and in particular, during the intensive manipulation of gas elements: in each run corresponding to a snapshot, hundreds of millions of gas components are tracked in front of millions of stellar particles, and some gas properties might have to be computed and assigned (e.g., neutral hydrogen column density, see Sect. 2.1.4 for more details). The typical memory consumed with these procedures in the third module is currently 180 GB, which is almost five times the amount of memory requested in the second module, during the convolution and integration of the SSPs SED within filters, which are memory-intensive operations (40 GB RAM) performed with well-optimized routines.

The code architecture, which has not been conceived as parallel, can be improved by implementing an MPI protocol to minimize memory consumption.

<sup>18</sup> The description of the GALILEO100 architecture is available at <https://wiki.u-gov.it/confluence/display/SCATUS/HPC+User+Guide>.

<sup>17</sup> [www.astrodeep.eu/FORECAST](http://www.astrodeep.eu/FORECAST)

The building of the light cone structure, the extraction of stellar particles properties from snapshot files, the arrangement of stellar particles in the observing cone following geometrical cuts, and the selection of the stellar particles in the FoV (first module) take  $\sim 2$  to 15 minutes, depending on the number of stellar particles included in the simulation volume. This is true also for the calculations related to the IGM absorption in the fourth module. In the second module, the connection of each SSP included in the FoV with the corresponding synthetic stellar population model is computationally inexpensive, while the conversion of the rest-frame spectrum into the observer-frame flux per unit wavelength, and its convolution and integration with the filter response are time-consuming operations, by order of  $\sim 15$  ms per stellar particles, translating into maximum  $\sim 0.6$  day of running per single snapshot if the particle budget is high ( $\sim 1.8 - 3.5$  million of stellar particles in the FoV). Concerning the implementation of dust effects (third module), the tracking of gas elements in front of star particles and the manipulation of their properties to derive the relevant quantities used to turn dust-free in dust-corrected magnitudes (e.g.,  $N_{HI}$ ) is slower ( $\sim 1$  day per snapshot).

### Appendix C: The Input Universe catalog

The Input Universe catalog is a file that collects all the information on the sources included in the simulated FoV, before they are post-processed and measured, that is their true values. It is built from the particle catalogs, which are the output of the code at each run (see Appendix B). The available fields, with their units and their description, are listed in Table C.1.

Table C.1: Galaxy catalog content - Input Universe

Field	Units	Description
subhaloID	-	identifier of Subhalo at that redshift; not unique; it could repeat throughout the full catalog.
redshift	-	redshift of Subhalo.
stellar mass	$M_{\odot}$	stellar mass of Subhalo.
gas mass	$M_{\odot}$	gas mass of Subhalo.
DM mass	$M_{\odot}$	DM mass of Subhalo.
SFR	$M_{\odot}yr^{-1}$	instantaneous star formation rate of Subhalo.
metallicity	$Z_{\odot}$	stellar mass-weighted metallicity of Subhalo.
age	Gyr	stellar mass-weighted age of Subhalo.
oldest SSP age	Gyr	age of the oldest SSP of Subhalo.
$(x_c, y_c)$	pix	coordinates of the center of Subhalo.
$R_{max}$	pix	maximum radius of Subhalo; it matches the further SSP from the center.
$N_{SSP}$	-	number of SSP in Subhalo.
$F_i$	$\mu J$	observer-frame, integrated flux in filter $i$ .
$F_{i,3pix}$	$\mu J$	observer-frame, integrated flux in filter $i$ , in $R = 3$ pix.
$R_{i,hl}$	pix	half-light radius for $F_{i,R_{i,hl}}$ .
$F_{i,R_{i,hl}}$	$\mu J$	observer-frame, integrated flux in filter $i$ , in $R_{i,hl}$ .