A&A, 692, A52 (2024)
https://doi.org/10.1051/0004-6361/202450269
© The Authors 2024

Astronomy & Astrophysics

Confusion of extragalactic sources in the far-infrared: A baseline assessment of the performance of PRIMAger in intensity and polarization

Matthieu Béthermin[1,2,⋆], Alberto D. Bolatto[3], François Boulanger[4], Charles M. Bradford[5,6], Denis Burgarella[2], Laure Ciesla[2], James Donnellan[7], Brandon S. Hensley[5], Jason Glenn[8], Guilaine Lagache[2], Enrique Lopez-Rodriguez[9], Seb Oliver[7], Alexandra Pope[10], and Marc Sauvage[11]

[1] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
[2] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[3] Department of Astronomy, University of Maryland, College Park, MD 20742, USA
[4] Laboratoire de Physique de l'École Normale Supérieure, Université Paris Science et Lettres, Centre National de la Recherche Scientifique, Sorbonne Université, Université de Paris, 75005 Paris, France
[5] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[6] California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
[7] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9QH, UK
[8] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[9] Kavli Institute for Particle Astrophysics and Cosmology (KIPAC), Stanford University, Stanford, CA 94305, USA
[10] Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
[11] AIM, CEA, CNRS, Université Paris-Saclay, Université de Paris, 91191 Gif-sur-Yvette, France

Received 5 April 2024 / Accepted 27 October 2024

ABSTRACT

Aims. Because of their limited angular resolution, far-infrared telescopes are usually affected by the confusion phenomenon. Since several galaxies can be located in the same instrumental beam, only the brightest objects emerge from the fluctuations caused by fainter sources. The PRobe far-Infrared Mission for Astrophysics imager (PRIMAger) will observe the mid- and far-infrared (25–235 μm) sky both in intensity and polarization. We aim to provide predictions of the confusion level and its consequences for future surveys.

Methods. We produced simulated PRIMAger maps affected only by the confusion noise using the simulated infrared extragalactic sky (SIDES) semi-empirical simulation. We then estimated the confusion limit in these maps and extracted the sources using a basic blind extractor. By comparing the input galaxy catalog and the extracted source catalog, we derived various performance metrics as completeness, purity, and the accuracy of various measurements (e.g., the flux density in intensity and polarization or the polarization angle).

Results. In intensity maps, we predict that the confusion limit increases rapidly with increasing wavelength (from 21 μJy at 25 μm to 46 mJy at 235 μm). The confusion limit in polarization maps is more than two orders of magnitude lower (from 0.03 mJy at 96 μm to 0.25 mJy at 235 μm). Both in intensity and polarization maps, the measured (polarized) flux density is dominated by the brightest galaxy in the beam, but other objects also contribute in intensity maps at longer wavelengths (~30% at 235 μm). We also show that galaxy clustering has a mild impact on confusion in intensity maps (up to 25%), while it is negligible in polarization maps. In intensity maps, a basic blind extraction will be sufficient to detect galaxies at the knee of the luminosity function up to $z \sim 3$ and 10^{11} $M_\odot$ main-sequence galaxies up to $z \sim 5$. In polarization for the most conservative sensitivity forecast (payload requirements), ~200 galaxies can be detected up to $z = 1.5$ in two 1500 h surveys covering 1 deg^2 and 10 deg^2. For a conservative sensitivity estimate, we expect ~8000 detections up to $z = 2.5$, opening a totally new window on the high-z dust polarization. Finally, we show that intensity surveys at short wavelengths and polarization surveys at long wavelengths tend to reach confusion at similar depth. There is thus a strong synergy between them.

Key words. techniques: polarimetric – galaxies: high-redshift – galaxies: ISM – galaxies: star formation – infrared: diffuse background – infrared: galaxies

1. Introduction

Far-infrared wavelengths are key to understanding galaxy evolution across cosmic time. Studies of the cosmic infrared background (CIB, e.g., Hauser & Dwek 2001; Lagache et al. 2005; Dole et al. 2006; Berta et al. 2011; Béthermin et al. 2012b) revealed that more than half of the relic emission from galaxies and their host nuclei is located in the 8–1000 μm range, with a peak around 150 μm. The UV photons emitted by young stars are absorbed by dust and their energy is re-emitted in the mid- and far-infrared. These wavelengths thus trace obscured star formation in the Universe.

While the CIB was detected in the 1990s (Puget et al. 1996; Fixsen et al. 1998; Hauser et al. 1998), characterizing all the

⋆ Corresponding author;
matthieu.bethermin@astro.unistra.fr

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

individual galaxy populations producing it remains difficult. Since far-infrared radiation can only be observed from space and the stratosphere, the diameter of the telescope primary mirror is limited. The angular resolution ($\theta \approx \lambda/D_{\rm tel}$, where λ is the wavelength and $D_{\rm tel}$ the telescope diameter) is thus severely limited by diffraction, and at the longest wavelengths it can be up to a few tens of arcseconds. At this resolution, several high-z galaxies can be located in the same beam, leading to background fluctuations, and only the brightest objects emerge above the fluctuations from unresolved faint galaxies. This phenomenon is called confusion (e.g., Condon 1974; Lagache et al. 2003; Dole et al. 2004), and must be taken into account when designing mid- and far-infrared telescopes and surveys.

With its actively cooled 85 cm mirror, the *Spitzer* space telescope (Werner et al. 2004) resolved more than 80% of the CIB at 24 µm into individual sources (e.g., Papovich et al. 2004; Béthermin et al. 2011), but only a small fraction ($\lesssim$10%) emerged above the confusion around the peak of the CIB at 160 µm (e.g., Dole et al. 2004; Frayer et al. 2006). Thanks to its 3.5 m mirror, the *Herschel* space observatory (Pilbratt et al. 2010) allowed us to resolve the majority of the CIB into individual sources at its peak (Berta et al. 2011; Magnelli et al. 2013), but not at longer wavelengths (Oliver et al. 2010; Béthermin et al. 2012b). The passive cooling of the mirror (~85 K) caused a high background and limited the sensitivity. *Herschel* could thus not reach the confusion limit below 100 µm, and the sizes of the deep field were limited between 100 and 200 µm (Lutz et al. 2011). For both *Spitzer* and *Herschel*, sub-confusion flux density regimes were probed using statistical methods as stacking of galaxy population known from shorter wavelengths (e.g., Dole et al. 2006; Béthermin et al. 2012b) or P(D) analysis, the analysis of the one-point distribution of intensity in the maps (e.g., Glenn et al. 2010; Berta et al. 2011), or using source extractors relying on priors from shorter wavelengths (e.g., Magnelli et al. 2009; Roseboom et al. 2010; Hurley et al. 2017).

The PRobe far-Infrared Mission for Astrophysics (PRIMA) project uses a 1.8 m space-based telescope cryogenically cooled to 4.5 K with a new generation of detectors that take full advantage of the low thermal background. One of the two payload instruments is the PRIMA imaging camera (PRIMAger, Burgarella et al. 2023; Meixner et al. 2024). PRIMAger has two main bands. The first one is an hyperspectral band (PRIMAger Hyperspectral Imaging: PHI) that provides imaging with a linear variable filter at a spectral resolution of $R \sim 10$ over 25 to 80 µm. The second band (PRIMAger Polarization Imaging: PPI) provides four broadband filters between 91 and 235 µm, which are sensitive to polarization. PRIMAger will operate with 100 mK cooled kinetic inductance detectors, which allows for an incomparable improvement of sensitivity in the far-infrared.

An observatory like PRIMA will cover a wide range of science topics such as, but not limited to, the origins of planetary atmospheres, the evolution of galaxies, and build-up of dust and metals through cosmic time (Moullet et al. 2024). In addition, it will offer for the first time spaceborne high-sensitivity far-infrared polarimetric capabilities in the far-infrared. So far, no high-z source has been polarimetrically detected at these wavelengths, and only one strongly lensed starburst galaxy at $z = 2.6$ has been published in the submillimeter (Geach et al. 2023). In addition, Chen et al. (2024) reported in a preprint a second submillimeter object at $z = 5.6$ exhibiting kiloparsec-scale ordered magnetic fields while we were revising this paper.

Since high-z galaxies will not be spatially resolved, we shall only detect the integrated polarization, if the magnetic fields driving the dust polarization in the various regions of a galaxy are ordered and their polarized flux densities add up at least partially. Otherwise, in the disordered case, the signals from the various regions cancel out, leading to a very small integrated polarization. The integrated polarization fraction can vary with various physical parameters such as the intrinsic dust polarization, the geometry of the galaxy, or the depolarization caused by turbulence (see, e.g., Sect. 6.3 of André et al. 2019). Pioneering studies in the local Universe have shown that two main mechanisms lead to organized polarization patterns, and thus significant integrated polarized fractions in star-forming galaxies: organized magnetic fields in disk galaxies (Lopez-Rodriguez et al. 2022) and starburst-driven outflows (Lopez-Rodriguez 2023). In addition, active galactic nuclei (AGNs) can also exhibit high polarization fractions in the far-infrared (e.g., Lopez-Rodriguez et al. 2018; Marin et al. 2020), and could also lead to polarized outflows.

The goal of this paper is to assess the impact of the confusion on PRIMAger's performance both in intensity and polarization based on the simulated infrared dusty extragalactic sky (SIDES, Béthermin et al. 2022) tools. We also demonstrate the feasibility of high-z polarization surveys.

It is important to realize that confusion arising from simple source extractions, like that assessed here, is not an ultimate limit for well-designed surveys and instrumentation. In this paper, we focus on the "classical" confusion limit for basic blind source extractors. Super-resolution techniques can certainly break through the classical confusion limit determined in these calculations to extract accurate spectral energy distribution (SED) information from fainter sources. More importantly, prior information derived from catalogs at un-confused wavelengths is very effective at improving the flux extraction to much fainter limits. PRIMAger's hyperspectral architecture is especially designed to take advantage of this type of technique to break through confusion. In this paper, we focus on performances expected from basic source extractors, while prior-based extractors, which show improvements by a factor of 2–3 beyond the confusion noise established here, will be discussed in a companion paper (Donnellan et al. 2024).

In Sect. 2, we introduce the SIDES simulation and describe how it was adapted to perform PRIMA forecasts. In particular, we describe the extension of SIDES to polarization in Sect. 2.3. We then describe in Sect. 3 the methods used to extract sources from the confusion-driven simulations and to assess the expected performances. We then present our results in intensity in Sect. 4 and in polarization in Sect. 5. Finally, we discuss the impact of confusion on future PRIMAger surveys and the expected number of detections in polarization in Sect. 6. We conclude in Sect. 7.

In this paper, we use the terminology "in intensity" to describe all the quantities associated with standard photometric surveys and usually derived from specific intensity maps (e.g., flux density of point sources). The term "in polarization" corresponds to quantities extracted from the polarization maps (Sect. 2.3) such as polarized flux density or the polarized fraction of the flux density (shortened hereafter to "polarization fraction").

2. Description of our simulation

Confusion depends on both the intrinsic nature of the sources being observed and the telescope and instrument providing the data. We discuss these two aspects of our simulation in turn.

2.1. The SIDES simulation

The confusion phenomenon is highly connected to the flux density distribution of galaxies (e.g., Condon 1974; Dole et al. 2004) and mildly by their spatial distribution, as we shall show in this study. To produce accurate forecasts of the confusion limit in the far-infrared, we thus need a realistic model of the statistical source properties at these wavelengths.

SIDES (Béthermin et al. 2017[1]) is a semi-empirical model populating dark-matter halos from numerical simulations using recent observed physical relations. In this paper, we use the 2 deg^2 version of SIDES. It connects the halo mass to the stellar mass using a sub-halo abundance matching technique (e.g., Behroozi et al. 2013). A fraction of the galaxies are drawn to be star-forming based on their stellar mass and redshift, and emits in the far-infrared. Their star formation rate (SFR) is then drawn based on the evolution of the main-sequence of star-forming galaxies, which is the relation between SFR and the stellar mass evolving with redshift (e.g, Schreiber et al. 2015). The observed scatter around this relation is taken into account by SIDES, and a population of high-SFR outliers is labeled as starbursts. Different SEDs are then attributed to galaxies depending on whether they are starbursts or not. These SEDs evolve with redshift following the observations of warmer dust at higher redshift (e.g., Béthermin et al. 2015). A temperature scatter on the SED templates is also included in the simulation. The AGN contribution is not included in the model, but Béthermin et al. (2012a) showed that it has a small impact on the galaxy number counts and thus the confusion noise (see Eq. (2) for the link between number counts and confusion noise).

This model reproduces successfully a large set of observables. The source number counts from the mid-infrared to the millimeter are very well reproduced after taking into account the resolution effects leading to the blending of some galaxies (Béthermin et al. 2017; see also Bing et al. 2023 for recent results in the millimeter). The simulation produces the correct redshift distributions and number counts in redshift slices. This capability of reproducing the galaxy flux density and redshift distributions over a large set of wavelengths suggests that both the SED and redshift distribution of galaxies are realistic, and confirms the relevance of our model to derive confusion limits. In addition, statistical measurements suggest that faint sources below the detection limits are also properly modeled. For instance, the histogram of pixel intensities in *Herschel*/SPIRE maps (250–500 µm, Glenn et al. 2010), also called P(D), is well reproduced after taking into account the clustering (Béthermin et al. 2017). However, some flux density and wavelength ranges targeted by PRIMAger were never observed before (e.g., there is a lack of deep observations between 24 and 70 µm), and we have to rely on the capability of our model to extrapolate correctly in these ranges.

Finally, the CIB anisotropies (e.g., Planck Collaboration XXX 2014; Viero et al. 2013), which is the integrated background from dust emission of galaxies at all redshifts, are also correctly recovered by SIDES, including cross-power spectra between different wavelengths (Béthermin et al. 2017; Gkogkou et al. 2023). The CIB anisotropies at small scale are dominated by the shot noise from galaxies below the detection threshold, while the signal at large scale is caused by the galaxy clustering. This agreement enhances our confidence that our model characterizes both the clustering and flux distribution of faint sources. The faithfully reproduced cross-power spectra indicate that the galaxy colors are also reasonable.

[1] The material associated with SIDES can be found at https://data.lam.fr/sides/home

Table 1. Properties of the various PRIMAger filters used in our analysis.

Band name	Central wavelength [µm]	Filter width [µm]	Beam FWHM [″]	With polarization?
PHI1_1	25.0	2.5	4.1	No
PHI1_2	27.8	2.8	4.3	No
PHI1_3	30.9	3.1	4.6	No
PHI1_4	34.3	3.4	4.9	No
PHI1_5	38.1	3.8	5.2	No
PHI1_6	42.6	4.3	5.7	No
PHI2_1	47.4	4.7	6.2	No
PHI2_2	52.3	5.2	6.7	No
PHI2_3	58.1	5.8	7.3	No
PHI2_4	64.5	6.5	8.0	No
PHI2_5	71.7	7.2	8.8	No
PHI2_6	79.7	8.0	9.7	No
PPI1	96.3	23.0	11.6	Yes
PPI2	126	31.0	15.0	Yes
PPI3	172	43.0	20.3	Yes
PPI4	235	61.0	27.6	Yes

Notes. As is explained in Sect. 2.2, the PHI1 and PHI2 bands are linear-variable filters and are represented in our analysis by 6 representative sub-filters (e.g., PHI1_X for the Xth representative filter of band PHI1).

2.2. Simulated PRIMAger maps

The confusion limit is the faintest flux density at which we can extract sources reliably in the limit of zero instrumental noise. In this paper, the instrumental noise refers both to the noise coming from the instrument itself and the photon noise from the various diffuse astrophysical backgrounds. To estimate the confusion limit, we produced simulated PRIMAger maps without instrumental noise using SIDES. Our simulations do not contain Galactic cirrus, which can potentially impact the photometry but are usually very faint in fields chosen for the deep surveys. However, in polarization, Galactic cirrus are expected to have a typically 5 times higher polarization fraction than unresolved galaxies (5% versus 1%; see Planck Collaboration XII 2020 and Sect. 2.3). The survey footprints will thus have to be chosen very carefully to minimize the cirrus contamination in polarization.

We assume a Gaussian beam with a full width at half maximum (FWHM) as listed in Table 1. These model beams are only determined by the optics and do not take into account the impact of the pointing accuracy, the scanning strategy, nor the future map-making pipeline process. The impact of these effects are discussed in Appendix B.

The flux densities in the PHI1 and PHI2 bands are derived assuming rectangular filters with spectral resolution $R = \lambda/\Delta\lambda = 10$. These bands employ linearly variable filters that, for simplicity, we decided to represent with six effective filters spanning the wavelength range of each band (PHI1_1 to PHI1_6 for the PHI1 band and PHI2_1 to PHI2_6 for the PHI2 band). This simplification has little consequence for confusion since it is caused by galaxies at various redshifts, and the polycyclic aromatic hydrocarbon (PAH) spectral features are thus smoothed. It would be consequential in other circumstances, such as redshift

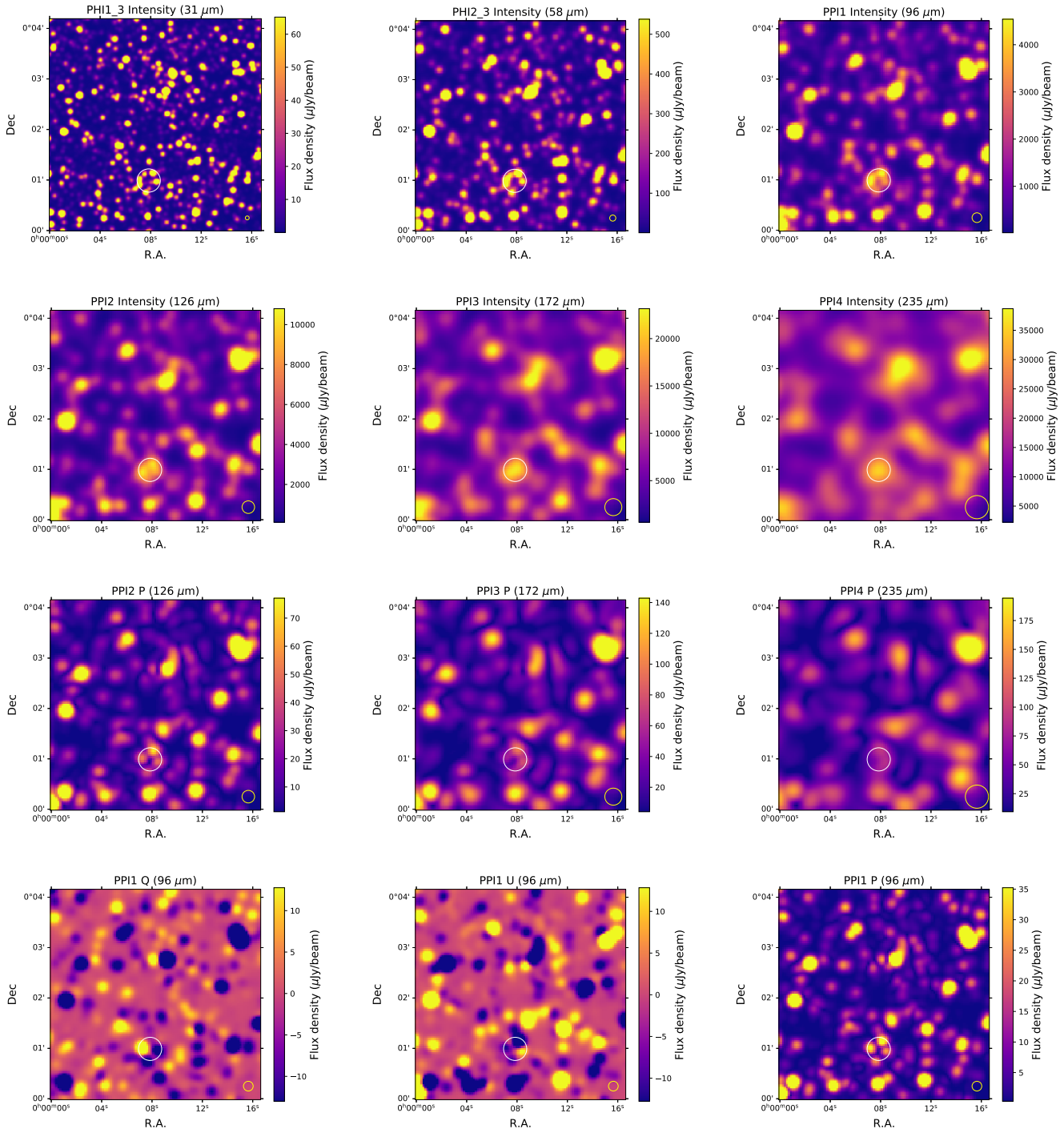

Fig. 1. Cutouts of our simulated PRIMAger noiseless maps produced by SIDES. The first two rows present the intensity maps of the various bands. The third row show the polarized flux ($P = \sqrt{Q^2 + U^2}$) maps of the PPI2, PPI3, and PPI4 bands, which can be compared with intensity maps in the same bands in the second row (see discussion in Sect. 2.3). The fourth row contains the Q, U, and P maps in PPI1 band, which illustrates how the Q and U maps combined into the P map. The PPI4 source indicated with a white circle is discussed in Sect. 2.3. The instrumental beam size is indicated using a yellow circle in the bottom right corner.

estimation from PAH emission bands. For PPI bands, we also assume rectangular filters, but with the currently planned widths that are listed in Table 1.

We produce simulated maps without instrumental noise using the integrated map maker of SIDES (`make_maps.py`, Béthermin et al. 2022). To ensure a sufficient sampling of the beam while keeping the data volume of the maps reasonable, we chose pixel sizes of 0.8″ in the PHI1 band, 1.3″ in PHI2 band, and 2.3″ in the PPI bands. This allows us to have at least 5 pixels per FWHM.

The cutouts (1/20 of the field width, 4.24′) of the resulting maps are presented in Fig. 1 (first two rows for the intensity maps). For simplicity, only one representative filter per band is shown for PHI. As was expected, because of the higher resolution at shorter wavelengths, the number of clearly defined (unconfused) sources is much higher in PPI1 than in PPI4.

2.3. Adding integrated polarization to SIDES

The PPI bands will measure the linear polarization. To study the impact of confusion on polarization measurements, we added a simple polarization model to SIDES motivated by observations of the local Universe.

We consider only the integrated polarization coming from the entire galaxy, because PRIMAger will not spatially resolve distant galaxies ($z \gtrsim 0.1$). As the plane-of-the-sky magnetic field orientations vary across the galaxy, the integrated dust polarization fraction from the entire galaxy is measured to be lower than those from individual lines of sight. Observations of a sample of local galaxies with the SOFIA telescope (the SALSA survey, Lopez-Rodriguez et al. 2022) demonstrated that the average integrated polarization fraction is ∼1%. A similar value ($p = 0.6 \pm 0.1\%$) has been measured in a lensed high-z galaxy (Geach et al. 2023). Since the physics of the dust polarization is very complex and necessitates physical parameters not included in SIDES, we chose a semi-empirical probabilistic approach anchored to the local Universe observations. In addition, more physical galaxy evolution models tend to struggle to reproduce far-infrared observables in intensity and their capability to produce dust polarization has never been tested.

In SIDES, we draw the galaxy integrated polarization fractions (p) from a Gaussian model of the distribution observed in the local Universe by Lopez-Rodriguez et al. (2022) centered on 1% and with a standard deviation of 0.3%. The method used to derive these values is presented in Appendix A. Since the 3-σ confidence interval of the central value of the polarized fraction goes from 0.7% to 1.3%, we shall also discuss scenarios with these extreme values to illustrate the uncertainties on our forecasts. For simplicity, we assume that the relative scatter (σ/μ) on the polarized fraction is constant (0.3) in all the scenarios to be able to apply a simple rescaling to the polarized maps. In the current version of the model, we do not consider a dependence of the polarization with wavelength or the presence of a starburst in the galaxy. Lopez-Rodriguez (2023) showed that starburst outflows can produce a specific wavelength-dependent signature. However, SIDES does not contain a model for outflow, and we could not calibrate an empirical dependence on these parameters using the SALSA sample since we did not find any statistically significant effect on the integrated polarization (see Appendix A).

As in Lagache et al. (2020), we neglect the intrinsic alignments between the integrated polarization angles of galaxies. This is primarily motivated by the small (<5%) probability of a spiral galaxy to be aligned with its neighbors found both in observations (Singh et al. 2015) and in simulations (Codis et al. 2018). We thus drew randomly the polarization angle α from a uniform distribution between 0 and π. The various integrated Stokes parameter (Q, U) and the polarized flux density (P) are then derived for each simulated galaxy using:

$$\begin{aligned} Q &= pI \cos 2\alpha, \\ U &= pI \sin 2\alpha, \\ P &= \sqrt{Q^2 + U^2} = pI, \end{aligned} \tag{1}$$

where I is the intensity.

The Q and U maps are built using the method described in Sect. 2.2 using the flux densities in Q and U instead of I. Contrary to I, Q and U can be negative, and the flux of the sources in the same beam do not add up systematically. It is thus important to compute the Q and U maps before generating the observed P map by combining them quadratically. The polarized maps are presented in Fig. 1 (last two rows). The comparison between the intensity and polarized maps (second and third rows) demonstrates immediately that the polarized maps are not rescaled versions of the intensity maps. While the flux of blended sources add up in intensity map to produce extended bright blobs, this is not the case in polarization. If the polarization angles are not aligned, the flux of two sources can potentially lead to depolarization. A good example can be found around the coordinates (00h00m08s, +00°01′00″). It is highlighted with a white circle in Fig. 1. The source is relatively bright in intensity in the PPI4 band, but barely visible in polarization in the same band. The PPI1 Q, U, and P maps (last row), which have a better resolution than the PPI4 map, help to understand the origin of this effect. The PPI1 P map exhibits four components in the beam. In the Q map, the eastern component has a positive signal, while the southern and western ones are negative. The northern component has no significant Q signal. In the U map, the northern component has a strong negative signal, the southern and eastern ones have a weaker negative signal, and the western component has a positive signal. When the beam is larger, the four components merge partially canceling each other Q and U signal. This leads to a small polarized flux density P.

3. Source extraction and determination of our confusion metrics

3.1. Our basic source extractor

In this paper, our goal is to determine the baseline performance that can be expected from PRIMAger at the confusion limit; that is, when the instrumental noise is negligible. For this baseline we purposefully chose a basic source detection and photometry method, making use of the standard `photutils` package (Bradley et al. 2023). This method is not intended to be optimal, and instead provide a robust estimate of the performance expected with a basic blind extraction algorithm.

We expect significantly better performance for a more advanced photometry method (XID+, Hurley et al. 2017) relying on priors and this is presented in Donnellan et al. (2024).

Point source detection and photometry methods typically employ an image filter. It is common for this to be optimized for the detection of isolated sources in the presence of uncorrelated instrumental noise, in which case the image is convolved with the PSF. Since we work at the instrumental-noise-free limit and confusion is our main concern this filter is not optimal. Indeed, filtering broadens the effective beam and increases the confusion noise. Optimal filters are discussed in Donnellan et al. (2024). Here, for simplicity, we choose to apply no filtering.

The sources are detected by searching for the brightest pixel above a given threshold (choice discussed in Sect. 3.3) within a local region using the `find_peaks` algorithm. The local region is defined to be 5×5 pixel square corresponding roughly to the beam half-light radius. Our maps are in units of Jy/beam, and the flux of the sources is estimated by recording the value of the central pixel in the background subtracted map. Finally, we determined the sub-pixel centroid of the sources using the `centroid_sources` algorithm using the same 5×5 region size used in detection.

3.2. Background

To detect sources and perform photometry, it is crucial to evaluate the background. This is not trivial for data from far-infrared observatories, since they are not usually absolute photometers.

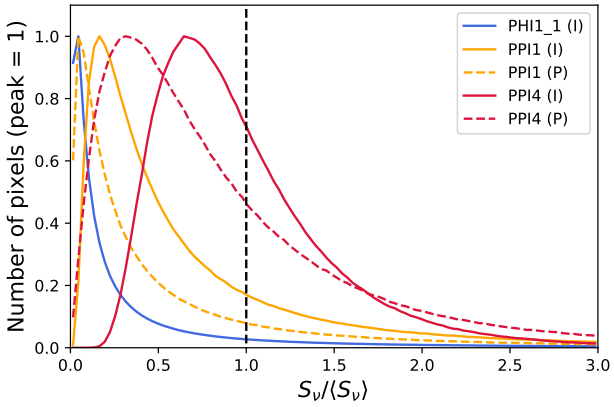

Fig. 2. Histogram of pixel flux densities of our simulated PRIMAger maps in various bands (PHI1_1 in blue, PPI1 in orange, and PPI4 in red) in intensity I (solid lines) and polarization P (dashed lines). The x axis is normalized by mean of the map, while the y axis is normalized to unity at the peak. The vertical dashed line corresponds to the mean of the map.

Real maps often have a zero mean enforced in order to filter out instrumental and celestial backgrounds and foregrounds. In addition, because of confusion, no region is free from sources to be suitable for estimating the background. To illustrate this, we show the flux density distribution from our PRIMAger simulated maps in Fig. 2. These simulated maps have a true zero, which corresponds to the absence of any galaxy emission. However, diffuse background (e.g., cosmic microwave background) and foregrounds (e.g., zodiacal light) are not included in the simulation and nor do they include instrumental backgrounds. In practice, we cannot use this zero for the photometry, since it cannot be measured in real data. Furthermore the faint sources provide an unresolved background that we should remove for unbiased photometry.

We thus chose to use the mode of the distribution to define the zero point. In Fig. 2, the x axis is normalized by the overall mean of the map. The mode is well below the mean (dashed vertical line). At short wavelength (PHI1 in blue), where the beam is small, the mode is very close to the true zero, since most beams contains only very faint sources. At long wavelengths (PPI4 in red), few pixels are close to zero and the mode is closer to the mean, since each beam contains a significant number of sources creating a background. Similar behavior is observed in polarization.

3.3. Classical confusion limit estimates and source extraction threshold

In the absence of clustering, the variance of the background fluctuations coming from sources below the detection limit ($\sigma_{\rm conf}$) can be computed using (Condon 1974; Lagache et al. 2000):

$$\sigma^2_{\rm conf} = \int \int b^2 d\Omega \times \int_0^{S_{\rm lim}} S^2 \frac{dN}{dS} dS, \quad (2)$$

where b is the beam function, $S_{\rm lim}$ is the flux limit above which sources can be detected, $\frac{dN}{dS}$ the number of galaxies per flux interval and per solid angle (usually called differential number counts). This equation is implicit, since $S_{\rm lim}$ is usually defined to be $5\sigma_{\rm conf}$ in the confusion limited case and needs to be solved; for example, by iteration. The limit is essential as without the integral would be divergent. However, the choice of where to place the limit is inherently subjective and the choice of 5σ is a convention without a strong rationale[2].

Equation (2) is no longer valid if we take into account galaxy clustering (Lagache et al. 2020). Béthermin et al. (2017, Sect. 5.1 and Fig. 12) show that the clustering tends to broaden the histogram of pixel flux densities; in other words, increase the fluctuations. This is expected since bright sources tend to bunch up together leading to stronger positive fluctuations, while low-density area tends to be even emptier. Since the analytical computation would be very complex, we decided to use an iterative map-based method inspired by Eq. (2). We compute the initial standard deviation of the map ($\sigma_{\rm map,0}$), and then recomputed it iteratively after masking the pixels $5\,\sigma_{\rm map,k}$ above background (see Sect. 3.2), where $\sigma_{\rm map,k}$ is the standard deviation at the k-th iteration. After a few tens of iterations, $\sigma_{\rm map}$ converges on the confusion noise, $\sigma_{\rm conf}$. The results are summarized in Table 2.

To extract sources from the simulated maps using the method described in Sect. 3.1, we set the detection threshold to $5\sigma_{\rm conf}$ after subtracting the mode. The photometry is also performed on the mode-subtracted map. The surface density of the detected sources goes from 43 962 sources per deg^2 in the PHI1_1 band to 306 sources per deg^2 in the PPI4 band. All the values are provided in Table 2. The surface density of sources above the confusion limit is slightly higher in polarization than in intensity in the same band (see discussion in Sect. 5.3). The confusion metrics expected for a different mean polarization fraction μ_p can

Table 2. $5\sigma_{\rm conf}$ classical confusion limits (Sect. 3.3) obtained with our minimal source extractor in intensity (I) and polarization (P), together with the 50% and 80% completeness flux densities (Sect. 3.5) and the source surface density.

Band name	Type	Central wavelength µm	5-σ limit mJy	50% comp. mJy	80% comp. mJy	Surface density deg^{-2}
PHI1_1	I	25.0	0.021	0.02	0.046	43 962
PHI1_2	I	27.8	0.027	0.027	0.055	39 602
PHI1_3	I	30.9	0.037	0.037	0.07	35 070
PHI1_4	I	34.3	0.051	0.051	0.091	30 680
PHI1_5	I	38.1	0.072	0.071	0.11	26 357
PHI1_6	I	42.6	0.11	0.1	0.16	22 080
PHI2_1	I	47.4	0.16	0.16	0.25	18 089
PHI2_2	I	52.3	0.25	0.25	0.35	15 281
PHI2_3	I	58.1	0.4	0.39	0.54	12 722
PHI2_4	I	64.5	0.66	0.64	0.82	10 498
PHI2_5	I	71.7	1.1	1.1	1.4	8545
PHI2_6	I	79.7	1.9	1.8	2.1	6890
PPI1	I	96.3	4.6	4.4	5.1	4346
PPI2	I	126.0	12	11	13	2227
PPI3	I	172.0	28	27	30	908
PPI4	I	235.0	46	42	49	306
PPI1	P	96.3	0.03	0.032	0.041	5898
PPI2	P	126.0	0.078	0.082	0.096	3326
PPI3	P	172.0	0.17	0.18	0.21	1582
PPI4	P	235.0	0.25	0.27	0.32	697

Notes. The polarized flux density is assumed to be in average 1% of the photometric flux density of a galaxy.

[2] The commonly presented rationale is that sources above the limit can be identified and removed and so would not contribute to the residual fluctuations, but a 5σ threshold will correspond to different detection probabilities depending on the underlying source counts.

be computed applying a (μ_p/1%) scaling factor (except for the surface densities; see Appendix C).

3.4. Matching of the input and output catalogs

Having extracted the sources on the simulated maps, we searched for the source counterparts in the simulated galaxy catalog. Since there can be numerous simulated galaxies in the beam, we chose to define the brightest galaxy in the half-light radius of the beam as the main counterpart. In single-dish far-infrared and submillimeter data, the flux of a source can come from multiple components (e.g., Karim et al. 2013; Hayward et al. 2013; Scudder et al. 2016; Béthermin et al. 2017; Bing et al. 2023). In our approach this multiplicity will become apparent as the observed flux density will be larger than the associated counterpart from the input catalog.

We found no systematic offset between the position of the brightest source and the observed centroid. The peak of the distribution of radial separations is less than the half-light radius by a factor of at least 3. At this cutoff radius, the histogram has a value 2% lower than the peak. This shows that the exact choice of the search radius for the brightest counterpart should have a negligible impact on the final results. The intensity results are presented in Sect. 4.1 and the polarization results are presented in Sect. 5.1.

3.5. Completeness and purity estimates

The completeness as a function of the flux density (polarized or not, in this section we use the term flux density to discuss both the intensity and polarization cases) is a key characterization of the source detection performance. A classical definition of the completeness is the fraction of sources at a given flux density in the input catalog (in our analysis, the SIDES simulated galaxy catalog) that are recovered in the output catalog produced by the source extractor. However, since several simulated galaxies can be found in the beam of a single source, recovery can be ambiguous. In this paper, we use two definitions of a recovered source. In definition A, we consider that a galaxy from the simulated catalog is recovered if it is located in the half-light radius of any source extracted from the associated simulated map. However, in this case, the completeness does not tend to zero at low flux density. At first order (no clustering), it converges on the fraction of the map encircled in the half-light radii of the extracted sources (up to 7%). To avoid this problem, we introduce a definition B, where the galaxy must satisfy the additional condition of being the brightest source in half-light radius. The results are discussed in Sects. 4.2 and 5.3.

The definition of the purity is also nontrivial. The purity is the fraction of "true" sources in a sample of detected sources, with the remained being artifacts caused by noise. In our case, there is no instrumental noise, and the density of galaxy in SIDES is so high that every beam contains several simulated galaxies. To declare a detection as true if there is a simulation counterpart in the beam would thus not be meaningful. Since our goal is to show that the flux of bright individual galaxies can be measured despite the confusion, we chose to consider an extracted source as true if the brightest counterpart is at least half of the measured flux density (definition A). At long wavelengths, the recovered flux density can be systematically overestimated due to the multiplicity described above (see also Sect. 4.1). We thus computed a second estimate of the purity (definition B) after correcting for this systematic bias, scaling by the median flux density ratio between the brightest galaxy in the beam and the extracted source.

4. Confusion in intensity

In this section, we focus on the impact of confusion on intensity data.

4.1. Photometric accuracy

A key question in photometry of sources with confused data is whether the measured flux densities are dominated by a single bright galaxy or come from several objects. Modeling can take into account blending effects before comparing predictions to data (e.g., Bing et al. 2023); however, most conventional astronomy relies on photometry of individual objects. We compared the flux density of the brightest source in the beam to the measured flux in the map using the matching algorithm presented in Sect. 3.4.

In Fig. 3, we show the ratio between the flux density of the brightest galaxy in the beam and the measurement in the map. This is the inverse of the classical output versus input ratio used to illustrate the flux boosting, but our choice has the advantage of immediately providing the relative contribution of the brightest galaxy. In PHI bands, the median ratio is above 0.91 at the classical confusion limit and converges rapidly on unity at higher flux density. We also studied the distributions around the median using the 16–84% and 2.3–97.7% ranges (corresponding to 1 and 2σ for Gaussian distributions). We note that the distribution is highly asymmetrical. Only rare outliers have an underestimated flux density (ratio>1), while the lowest 2.3% can have a flux density underestimated by half in these bands. This is a consequence of the large tail of positive outliers in the histogram of pixel flux densities (see Fig. 2). These overestimated flux densities are mainly caused by the blending of two sources with a similar flux density. Advanced deblending algorithms are usually very efficient at mitigating this effect (Donnellan et al. 2024).

In PPI bands, the contribution from other sources in the beam becomes more significant with larger beams at longer wavelengths. The median flux density ratio at the classical confusion limit decreases with increasing wavelength from 0.90 to 0.72. This effect has already been discussed in the case of the *Herschel*/SPIRE instrument by Scudder et al. (2016) and Béthermin et al. (2017), and is mainly caused by sources at other redshifts, while a small contribution ($\lesssim$5%) comes from physically related sources. The dispersion of the ratio around the median value also becomes more symmetrical with increasing wavelength. This is expected, since the histogram of pixel flux densities becomes more symmetrical at longer wavelengths (see Fig. 2). This is a consequence of the central limit theorem as the approximately Poisson distribution of source fluxes becomes more Gaussian with a larger number of sources per beam.

We investigated if a different choice of background could impact the measured flux density excess using the extreme example of the PPI4 band. As is discussed in Sect. 3.2, the mode is the most natural choice in a noiseless case, but a higher background could reduce the excess. The median flux density excess is 15.5 mJy. Choosing a higher background as the median or the mean would have led to an excess of 12.9 and 10.4 mJy, respectively. It does not change the results qualitatively.

We also estimated the flux uncertainties using half of the 84–16% interval (corresponding to 1σ for a Gaussian distribution). At the classical confusion limit, in PHI bands, the relative flux

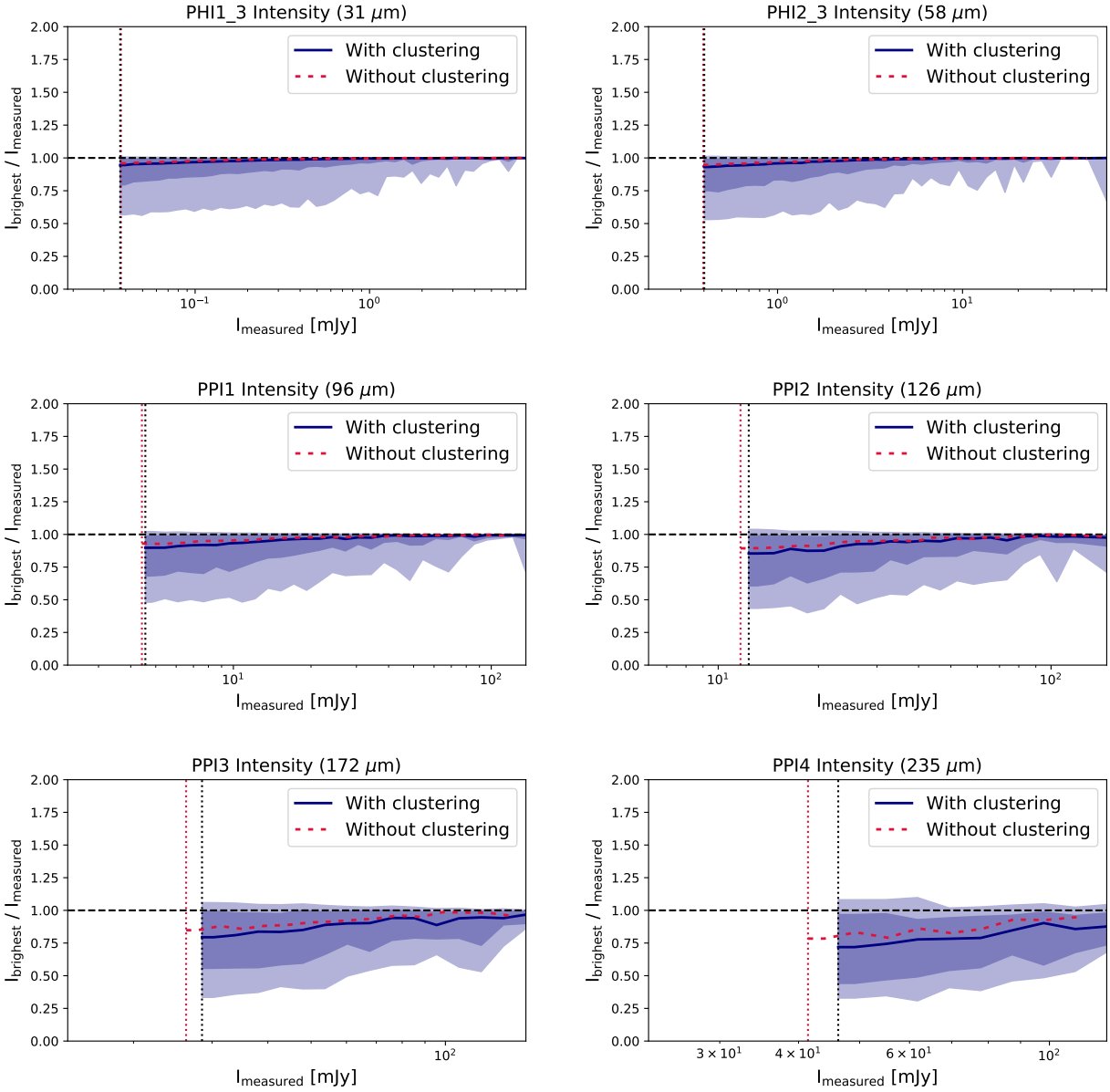

Fig. 3. Ratio between the flux density of the brightest simulated galaxy in the beam (see Sect. 3.4) and the measured flux density in the simulated map as a function of the measured flux density. The various panels corresponds to the various PRIMAger bands in intensity (see title above the panel). For PHI bands, we show only the third representative filters. The solid dark blue line is the median value. The dark and light blue areas represent the 16–84% and 2.3–97.7% ranges, respectively, which are equivalent to 1σ and 2σ in the Gaussian case. The horizontal dashed line is the one-to-one ratio. The vertical dotted black and red lines are the classical confusion limit estimated in Sect. 3.3 used as the detection threshold to produce the output catalog with and without clustering, respectively. The dashed red line represents the median flux density ratio in absence of clustering.

uncertainties range from 8% in band PHI1_1 to 15% in band PHI2_6, and are better than the 20% expected from our 5 σ construction of the extraction threshold. This is because the distribution is highly non-Gaussian and the clipped variance is more sensitive to outliers. These relative uncertainties below 20% confirm that our estimate of classical 5σ confusion level is conservative in these bands, and deeper catalogs can be obtained with more advanced extraction methods (see Donnellan et al. 2024). In PPI bands, the performance degrades with increasing wavelength (15% in PPI1, 19% in PPI2, 22% in PPI3, and 26% in PPI4). The accuracy in PPI3 and PPI4 bands is slightly worse than the 20% expected. In these bands, blind-extracted catalogs will be challenging to use in intensity, and their interpretation will either require complex statistical corrections of the fluxes or incorporating the effects of angular resolution through statistical models (e.g., Hayward et al. 2013; Cowley et al. 2015; Béthermin et al. 2017; Bing et al. 2023). In contrast, techniques such as prior-based source extraction will enable accurate measures of the fluxes (e.g., Donnellan et al. 2024).

Finally, we investigated the impact of clustering on the flux density bias by comparison with results following randomization of the source positions (without clustering). This illustrated in Fig. 3 by the dashed red line. The effect is almost negligible in PHI bands. In PPI bands and in absence of clustering, the brightest galaxy in the beam of sources at the classical confusion limit contributes to 3%, 4%, 6%, and 8% more than in the clustering case. The clustering does not explain fully the effect discussed previously, but it reinforces it.

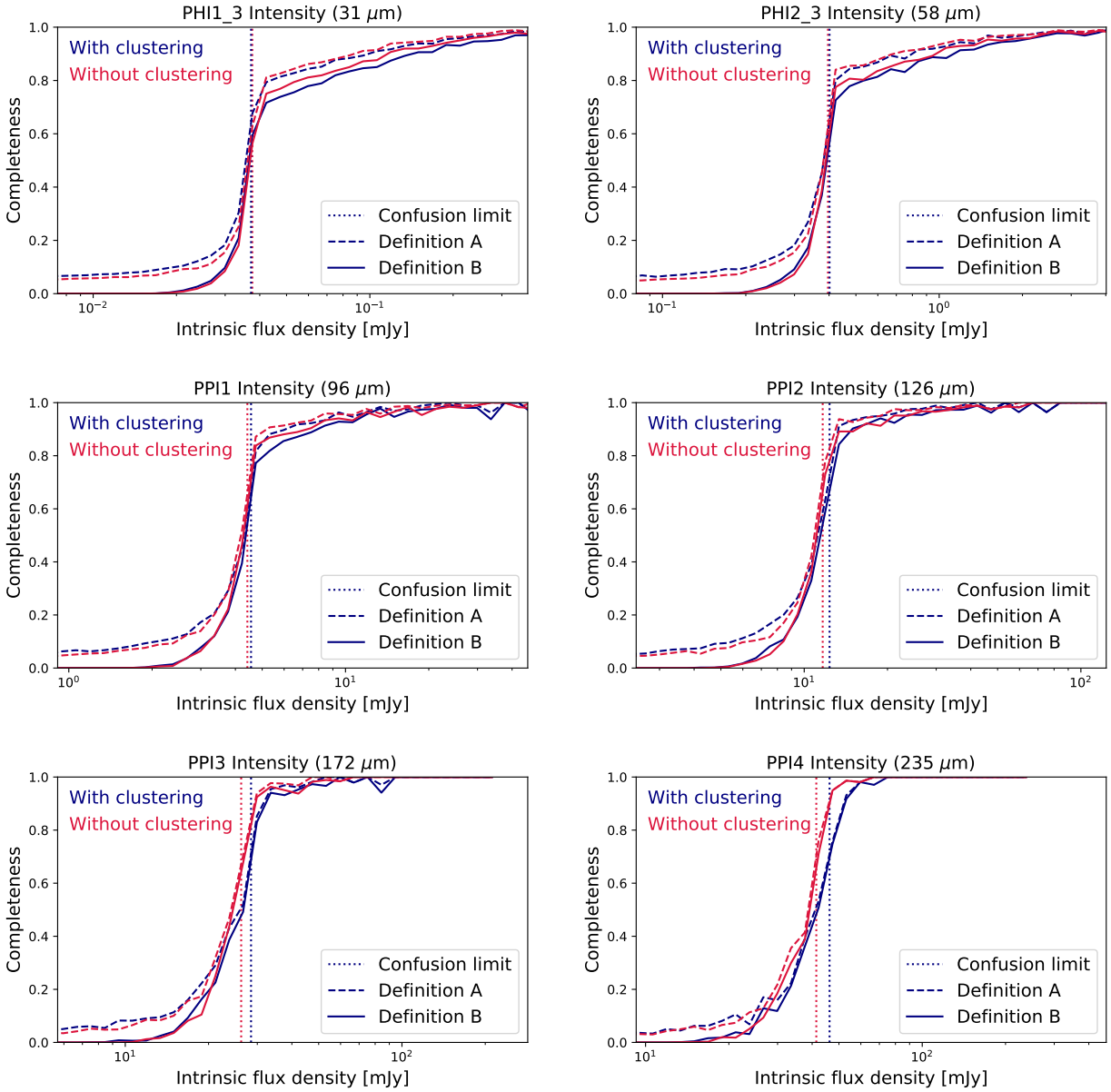

Fig. 4. Completeness as a function of the intrinsic flux density of a galaxy in intensity in the input simulated catalog. The six panels correspond to the same bands as in Fig. 3. The red and blue lines show the results with and without clustering, respectively. The dotted and solid lines correspond to definitions A (a detected galaxy is in the beam of an extracted source) and B (a galaxy is detected only if it is the brightest in the beam of the extracted source) of the completeness described in Sect. 3.5. The vertical dotted line is classical confusion limit computed in Sect. 3.3.

4.2. Completeness

The probability that a survey will be able to detect sources as a function of flux density, called completeness, is also an important performance criterion. In Fig. 4, we present the completeness obtained using our minimal extractor in simulated noiseless PRIMAger maps. We show the two definitions of completeness introduced in Sect. 3.5 for both the clustered and randomized cases.

At low flux densities, the completeness converges on a nonzero value with definition A. As is discussed in Sects. 3.4 and 3.5, this is caused by the faint galaxies in the beam of a brighter object being considered as detected. We do not observe this behavior for definition B, where only the brightest galaxy in the beam is considered to be detected.

Both definitions reach 50% close to the classical 5σ confusion limit used as threshold by our source extractor (see Sect. 3.3). In the case of Gaussian noise (or any symmetrical noise), we would expect to have exactly 50% at the extraction threshold, since only half of the sources will be on a positive noise realization. However, in the PPI3 and PPI4 bands, the completeness is slightly larger (~60%). This could be that in these bands the flux is not coming from a single object and the flux is boosted by the neighbors (Sect. 4.1 and Fig. 3).

At short wavelengths, the completeness curve increases only slowly above the 5-σ classical confusion limit (one full flux density decade from 80% to 95% in PHI1), especially for definition B of the completeness. This suggests that the sources close to bright sources tend to be missed, since this second definition does not consider a faint galaxy to be detected when in the beam of a brighter one. At long wavelengths, the transition is much

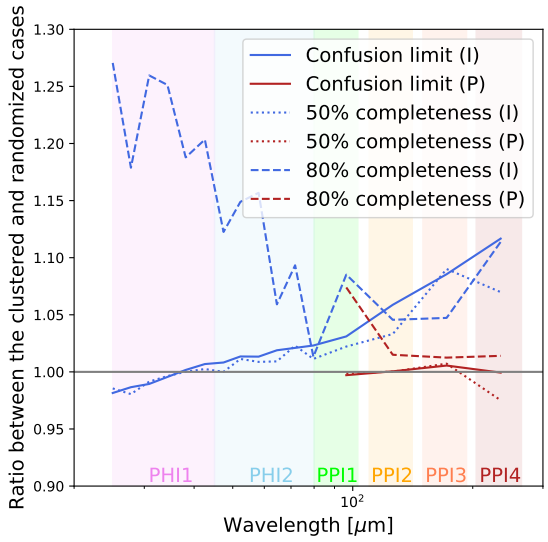

Fig. 5. Ratio between values found using the simulation with (original SIDES) and without (randomized positions) clustering for the following performance metrics: 5σ classical confusion limit (solid line), 50% completeness flux density (dotted line), and 80% completeness flux density (dashed line). The intensity is in blue and the polarization in dark red. The intensity is discussed in Sect. 4.2 and the polarization in Sect [illegible].

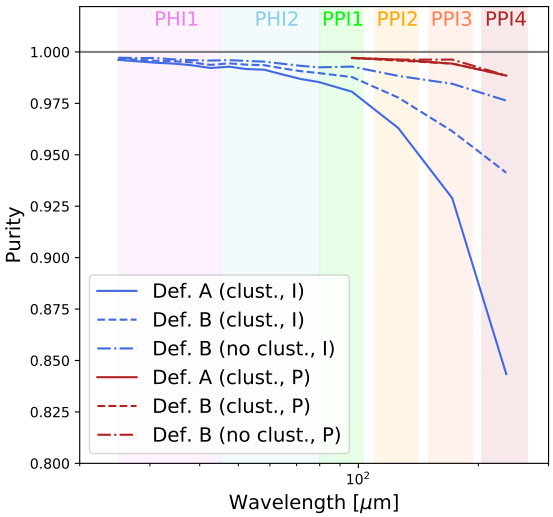

Fig. 6. Purity of the sample extracted from the simulated map above the $5\sigma_{\rm conf}$ threshold in intensity (blue) and polarization (dark red). The solid lines correspond to definition A of purity (see Sect. 3.5), where a source is considered "true" if the brightest galaxy in the beam is at least half of the measured flux. The dashed and dash-dotted lines are corresponding to definition B with and without clustering, respectively. In this second definition, we lower the minimal flux density to consider a source to be true by the flux density excess factor measured in Sect. 4.1. We discuss the intensity in Sect. 4.3 and the polarization in Sect. 5.3.

sharper and converges rapidly on unity. There is also not much dependence on the definition.

Because definition A at faint flux density does not converge on zero and the difference between the two definitions remains small around the confusion limit, we use only definition B of the completeness in the rest of this paper. In Table 2, we summarize the classical confusion limits and the 50% and 80% completeness levels found in the various bands.

Finally, we investigated the wavelength dependence of the clustering impact on completeness. In Fig. 5, we compare the value of the completeness determined in our clustered simulation and after randomizing the positions (no clustering). The ratio between the clustered and the random case increases with increasing wavelength for both the 5σ classical confusion limit and the 50% completeness flux density, and reaches ∼10% in the PPI4 band. This is expected, since the clustering tends to broaden the pixel flux histogram (e.g., Béthermin et al. 2017). The behavior on the short-wavelength side of the PHI1 band, where the clustered case has a lower classical confusion limit, is less intuitive. This is a small effect (∼2%), and could be due to the strong blending of the bright sources in the clustered case decreasing the density in the rest of the field. Finally, the 80% completeness flux density has a more complex U-shaped trend with wavelength, with a minimum between band PHI2 and PPI1. The rise above band PPI1 has a similar explanation as for the other quantities. A possible explanation for the strong impact (∼25%) of clustering at short wavelengths is that the bright sources tend to cluster with each other and a small fraction of objects well above the global classical confusion limit are missed, since they are in the vicinity of brighter objects. At longer wavelengths, the flux density ratio between the brightest and faintest detectable sources is smaller, reducing the impact of this effect.

4.3. Purity

The third criterion to evaluate the quality of the catalogs is the purity. Surveys usually aim for 80–95% depending on whether the scientific goal is a pure statistical measurement or building a clean sample for detailed follow-up studies. As is discussed in Sect. 3.5, the definition of purity in confusion-limited data is not trivial. In Fig. 6, we show the purity as a function of wavelength for our two definitions.

In PHI bands, the purity is always excellent (>98%) whatever the definition, even though it slightly decreases with increasing wavelength. This suggests that we were conservative in our choice of extraction threshold. We could thus expect to go deeper for statistical studies in confusion-limited data using more aggressive source extraction algorithms.

In PPI bands, the purity degrades rapidly with increasing wavelength. In the case of definition A (source considered true if the brightest counterpart produces more than half of the measured flux density), it drops to 84% in PPI4. This is mainly because the median ratio between the flux density of the brightest galaxy in the beam and the measured flux is below unity (see Sect. 4.1). If we use definition B of the purity for which we correct the measured flux densities by the median ratio, the purity rises to 94%. This demonstrates that it is the main reason of the lower purity at longer wavelengths. In the case of real data, this average correction could be calibrated using artificial sources injections in the data or using end-to-end simulations. Finally, if we use definition B in absence of clustering, the result increases to 97.6% and is close to the performance reached in PHI bands. The clustering thus has also a mild impact on the degradation of PPI-band purity.

4.4. Detection probability in the SFR-z plane in intensity

In the previous sections, we characterized the classical confusion limit only in term of flux density. However, to understand its impact on the observatory science, it is essential to consider the impact on intrinsic physical properties. We thus computed the probability of detecting a galaxy above the classical confusion

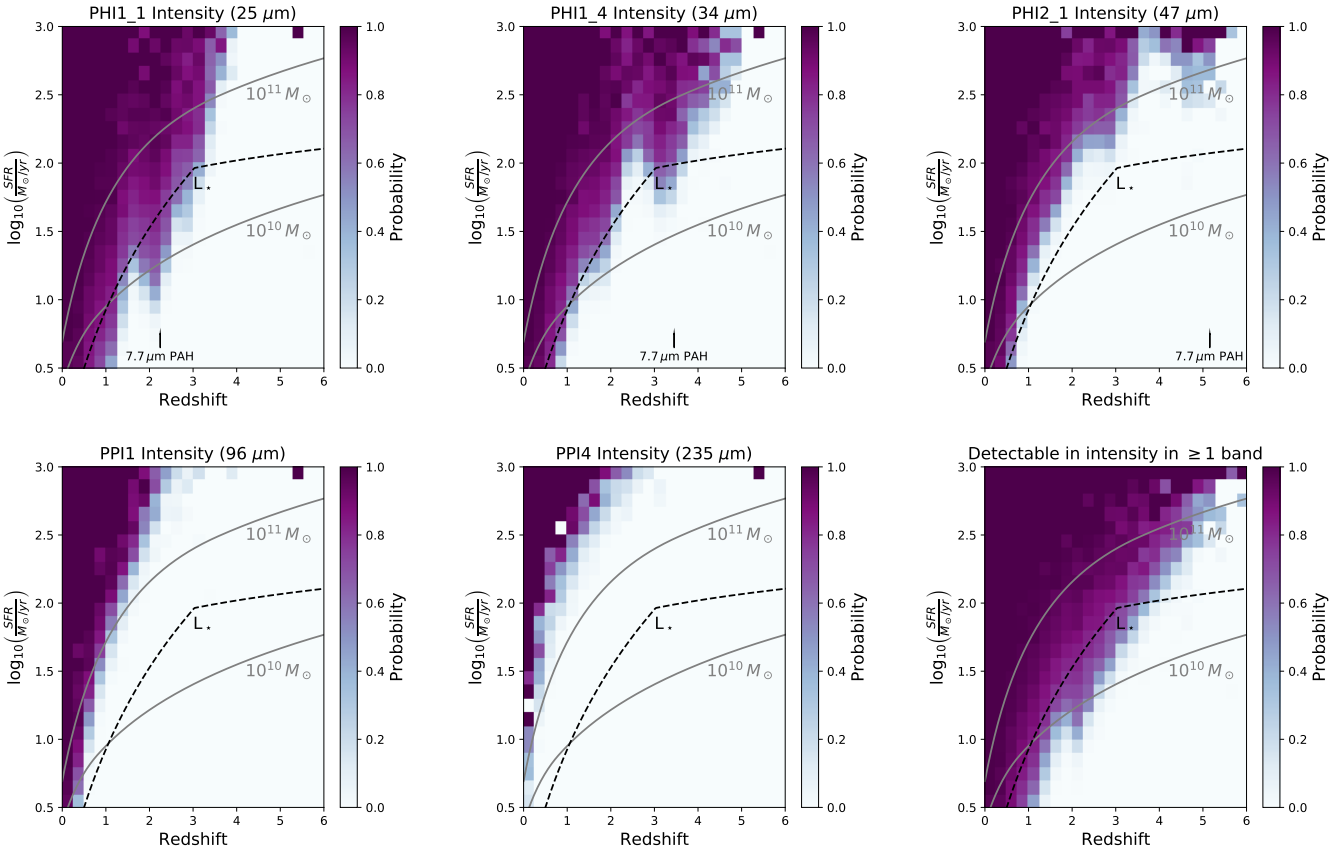

Fig. 7. Probability (color-coded) of detecting a galaxy with our basic blind source extractor in an intensity map affected only by confusion as a function of its position in the SFR-z plane. The lower right panel is the probability of detecting the source in at least one band, while the other panels are for a selection of single bands. The two gray tracks show the position of a galaxy exactly on the main-sequence relation (Schreiber et al. 2015) for various stellar masses. The dashed black line shows the evolution of the knee of the infrared luminosity function $L_\star$ measured by Traina et al. (2024).

limit as a function of SFR and redshift. The border between the regions of low and high probability is blurred, since our model has a diversity of SEDs and our simulation produces a completeness curve that has a continuous transition from 0 to 1 (see Sect. 4.2). The results are presented in Fig. 7 together with the tracks corresponding to galaxies of various stellar masses following the main sequence relation of Schreiber et al. (2015) and the evolution of the knee of the infrared luminosity function ($L_\star$, Traina et al. 2024).

For the shortest wavelength (PHI1_1 band centered on 25 μm, upper left corner of Fig. 7), the $10^{10}\,M_\odot$ and $10^{11}\,M_\odot$ main-sequence galaxies are recovered up to $z \sim 2.5$ and $z \sim 3.5$, respectively. The $L_\star$ galaxies are slightly above the detection border up to $z \sim 3$ and undetected above. Overall, the border between detections and non-detections moves toward higher SFR with increasing redshift. However, we can identify some specific features. At $z \lesssim 2$, we are probing the 10 μm rest-frame dip in SED between the various PAH bands, and galaxies are harder to detect (higher SFR limit). At $z \gtrsim 2$, the 7.7 μm PAH band enters the representative filters making the galaxies easier to detect (lower SFR limit). Around $z = 3$, the typical SFR at which galaxies are detected increases sharply, and only some rare outliers can be detected. This is the consequence of the absence of strong dusty features below 6 μm rest-frame in the SED of star-forming galaxies.

At longer wavelengths in PHI bands, at z <1.5, the SFR sensitivity decreases with increasing wavelength. At $z > 1.5$, the PAH features boost the SFR sensitivity in some specific redshift range (e.g., $z \sim 3.4$ in PHI1_4 at 34 μm and $z \sim 5.1$ in PHI2_1 at 47 μm). In PPI bands, the PAH corresponds to very high redshifts and the border between the detection and non-detection areas is less complex.

Finally, we combined all the bands to derive the probability of detecting galaxies in at least one band (lower right corner of Fig. 7). This illustrates the parameter space, which could be probed by PRIMAger above the classical confusion limit. A galaxy exactly on the main-sequence relation and with a stellar mass of $10^{10}\,M_\odot$ and $10^{11}\,M_\odot$ can be detected up to $z \sim 2.5$, and $z \sim 5$, respectively. The $L_\star$ galaxies are detected up to $z \sim 3.5$. The border between detection and non-detection is almost featureless, since the PAH slides through the various representative sub-filters. This illustrates how hyperspectral imaging can help to deal with confusion. However, the dip at $z \sim 1.5$ seen in PHI1_1 is still present, since there is no shorter wavelength to observe this redshift range around 7.7 μm rest-frame.

5. Confusion in polarization

In this section, we discuss the impact of confusion on polarization data. We do not take into account the instrumental noise in this section. If we assume a different mean polarized fraction μ_p, both the confusion noise and the source polarized flux density scales as (μ_p/1%). Consequently, the x axis of Figs. 8 and 9 and the values in Table 3 must be shifted by this factor, while Fig. 10 is unchanged.

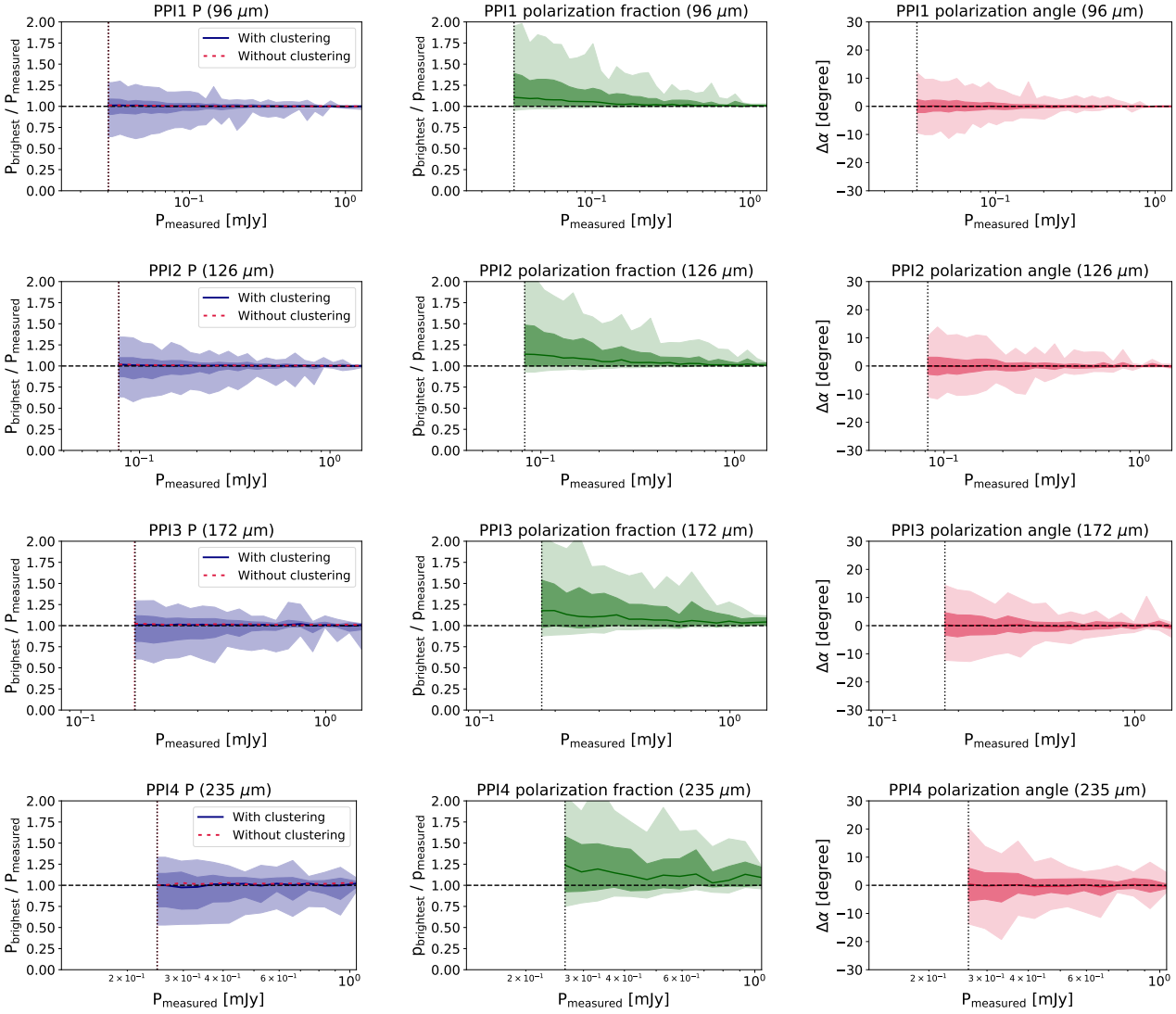

Fig. 8. Performance of the integrated polarization measurements. Left panels: ratio between the polarized flux density P of the brightest simulated galaxy in the beam and the measured one in the P map. The symbols are the same as in Fig. 3. Central panels: same thing for the polarized fraction p. Right panels: difference between the polarization angle α of the brightest source in the beam (see Sect. 2.3) and the measured angle. The rows correspond to the various polarized band from PPI1 to PPI4.

5.1. Photometric accuracy

Since the polarized flux density, P, is the quadratic combination of Q and U (Eq. (1)), the pixel values of the P map are always positive, while Q and U pixels can be either positive or negative (with a zero mean in the absence of alignment between galaxies). However, contrary to the intensity maps where flux densities always add up, two bright sources at the same position and with the same polarized flux density, P, can, in principle, lead to a null P flux intensity map if their polarization angles, α, differ by $\pi/2$, since the sum of their Q and U values will be zero. As is illustrated by Fig. 2, the mode of the P-map histogram is strictly positive, and even in polarization it is important to define carefully the background.

In Fig. 8 (left panels), we show the ratio between the brightest galaxy polarized flux density, P, in the beam and the measured value in the simulated map. In contrast with the intensity maps (Sect. 4.1 and Fig. 3), we do not observe any bias in the median flux ratio. This is likely to be because the contribution of several sources in the beam is not fully additive if their polarization angles are not aligned. However, similarly to the intensity, we still observe an increase in the half width of the $1\,\sigma$ confidence region from 9% to 20% from PPI1 to PPI4, but overall the dispersion is slightly lower than it is for intensity.

We can thus recover the polarized flux density of sources just above the classical confusion limit with a good accuracy, while this is not the case for intensity (see Sect. 4.1). However, the polarized flux density is much weaker than the intensity, and detecting it will require much deeper data. In addition, as is discussed in Sect. 2.3, we did not include galaxy alignments, which could produce a small polarized flux density excess similar to what happens in intensity.

5.2. Recovering polarized angles and polarized fraction

The polarized fraction, p (P/I), is another useful quantity to characterize distant unresolved galaxies. We derived p for each galaxy detected in the P map, extracting the value of I at the same position in the intensity map. In Fig. 8 (central panels), we show the polarized fraction ratio between the brightest galaxy

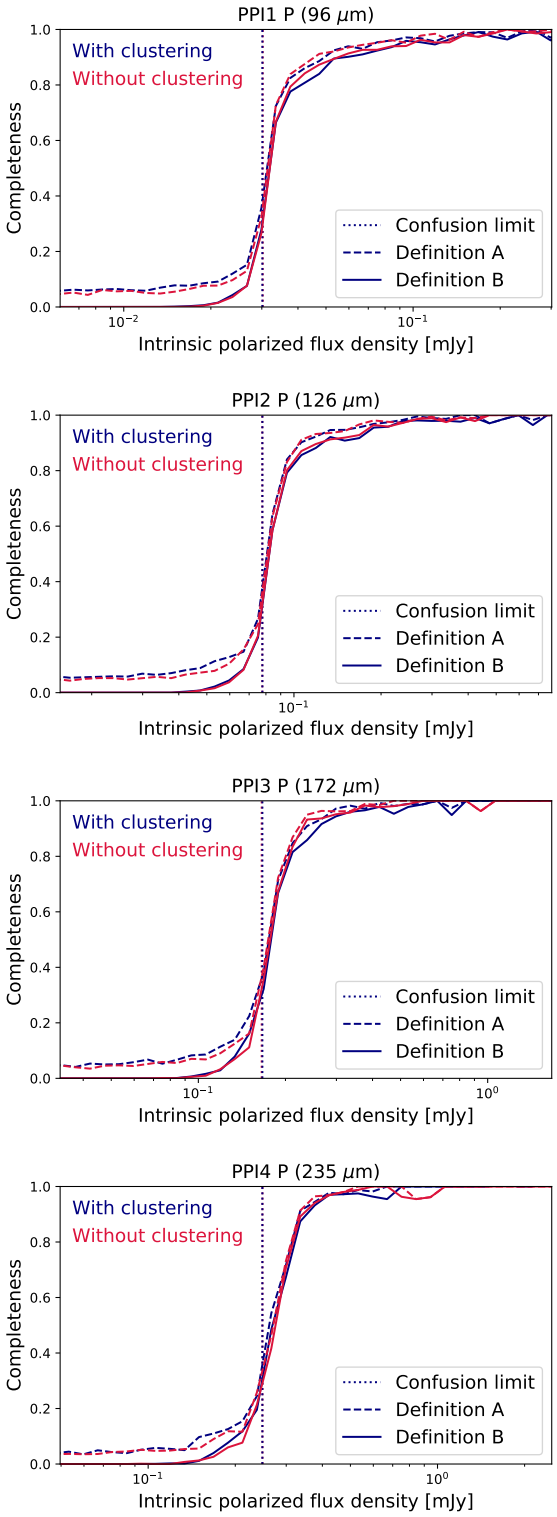

Fig. 9. Same as Fig. 4, but for the four polarized bands.

in the beam (highest P) and the measurement in the simulated map. At low polarized flux, the intrinsic polarized fraction of the brightest galaxy is significantly larger than the measured one. This is a natural consequence of the negligible bias found for the polarized flux density measurements (P, Sect. 5.1) and the significant bias found in intensity (I, Sect. 4.1 and Fig. 3), since $p_{\rm brightest}/p_{\rm measured} = (P_{\rm brightest}/P_{\rm measured}) \times (I_{\rm measured}/I_{\rm brightest})$ with the first factor being close to one and the second being significantly above (i.e., the inverse of the quantity shown in Fig. 3).

We also tested our ability to recover the polarization angle, α. We measured it from the Q and U values found at the position of sources detected in the P maps:

$$\alpha = \begin{cases} \frac{1}{2}\arctan(\frac{U}{Q}) & \text{if } Q \geq 0, \\ \frac{1}{2}\arctan(\frac{U}{Q}) + \frac{\pi}{2} & \text{if } Q < 0. \end{cases} \quad (3)$$

We then computed the difference between the intrinsic polarization angle of the brightest galaxy in the beam and the measured angle ($\Delta\alpha$). Since $\Delta\alpha$ is defined modulo π, we shifted all the values between $-\pi/2$ and $+\pi/2$. The results are presented in Fig. 8 (right panels). We do not identify any significant bias. Just above the classical confusion limit, the precision remains high and the half width of the 16–84% region is 2, 3, 4, and 6 deg in PPI1, PPI2, PPI3, and PPI4, respectively. However, the region equivalent to $2\,\sigma$ is more than two times broader (11, 11, 13, and 17 deg, respectively), highlighting that the impact of the confusion noise on angle measurements is non-Gaussian.

5.3. Purity and completeness

The purity of the samples extracted from polarization maps is excellent (>98%; see Fig. 6), and is barely affected by the choice of definition for the clustering. This is not surprising as the low level of flux boosting by the neighbors on P means definitions A and B consider very similar matches. Finally, the clustering is also not expected to have a strong impact, since the polarized flux density does not add up as it does for the intensity data.

The completeness curves as a function of the intrinsic galaxy P have a rather similar shape to those found for I, but the transition between low and high completeness appears at lower flux densities (see Fig. 9). As is shown in Table 2 summarizing the completeness in both intensity and polarization, the polarized flux density limits are up to a factor of 1.8 lower than the product of the limits in intensity by the mean polarization fraction, μ_p. This is again probably caused by the non-additivity of the polarized flux density inside a beam in polarization mitigating slightly the blending problems. Consequently, the surface density of sources above the classical confusion limit in a given band is higher in polarization than in intensity (see Table 2).

Finally, we find that the impact of clustering on the 50% completeness polarized flux density and the classical confusion limit is negligible (see Fig. 5, red lines). This is the consequence of the confusion in polarization being driven by chance polarization alignments rather than the local source density. There is a small impact (<10%) on the 80% completeness polarized flux density, which could have the same cause as the effect seen at short wavelengths in intensity (see discussion in Sect. 4.2).

5.4. Polarized detection probability in the SFR-z plane above the classical confusion limit

We also studied the probability of recovering a source above the classical confusion limit using the same method as is described in Sect. 4.4 for the intensity. The results are shown in Fig. 10.

In the PPI bands, only $z < 2.5$ galaxies emerge from the confusion, similarly to intensity. The PPI1 band probes $10^{10}\,M_\odot$, $10^{11}\,M_\odot$, and $L_\star$ galaxies up to $z \sim 0.5$, $z \sim 1.5$, and $z \sim 0.5$, respectively. Above $z = 2.25$, the probability of detection remains small, even for the most strongly star-forming galaxies. Although they are less sensitive at low redshift, the PPI2 and PPI3 bands are slightly better at catching these extreme

Table 3. Estimated fluctuations caused by the CIB in polarized surface brightness density maps in absence and in presence of a strong foreground (see Sect. 5.5), minimal foreground surface brightness to obtain a 10% precision on the foreground polarized color (confusion noise only), and correlation coefficient of the polarized CIB signal between two polarized bands at this minimal surface brightness.

	PPI1	PPI2	PPI3	PPI4
Polarized CIB 1σ fluctuations at native resolution in kJy/sr	1.7	2.6	3.0	2.4
1σ confusion noise at native resolution in kJy/sr (strong foreground case)	1.8	2.9	3.4	2.9
Polarized CIB 1σ fluctuations at PPI4 angular resolution in kJy/sr	1.2	1.8	2.3	2.4
1σ confusion noise at PPI4 angular resolution in kJy/sr (strong foreground case)	1.4	2.2	2.8	2.9
Polarized surface brightness limit to obtain a 10% uncertainty on the color with PPI4 in kJy/sr	11	19	20	–
Correlation coefficient between a band and PPI4 at the polarized surface brightness limit	0.79	0.87	0.96	–

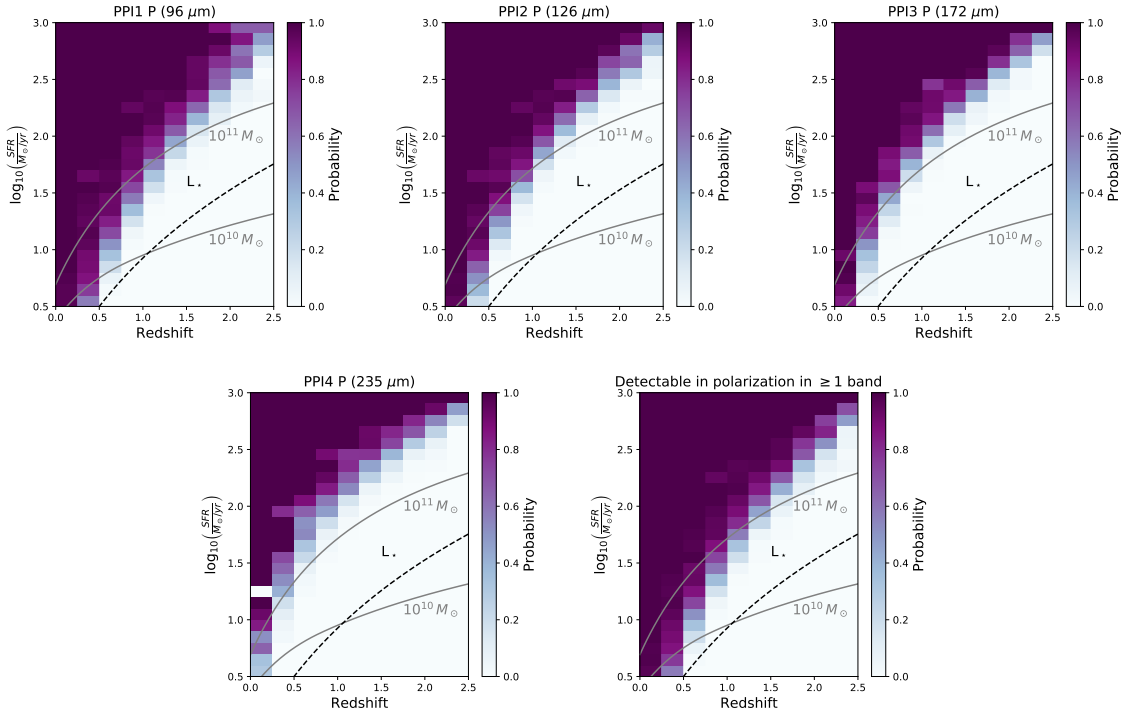

Fig. 10. Same as Fig. 7 but for polarization.

sources, since they observe them closer to their peak of emission. Consequently, the probability of detecting a source in at least one band is very similar to the probability of detecting it in the PPI1 band with the exception of the tail at $z > 2.25$ and SFR $\sim 1000\,M_\odot$/yr.

5.5. Impact of high-z galaxy confusion on measurements of Galactic and low-z diffuse emission in polarization

The confusion noise is not only a problem for studying high-redshift galaxies. The fluctuations of the polarized CIB can impact both diffuse foreground and background measurements. The case of the cosmic microwave background has already been extensively discussed by Lagache et al. (2020). While the contribution of astrophysical components in intensity is additive, the polarization is a vectorial quantity and leads to a more complex combination of the various components. In intensity, the confusion noise from the CIB is only sufficient to estimate its impact on the foreground science (e.g., diffuse Galactic emission and nearby or spatially resolved galaxies). In contrast, as we show in this section, the impact of the CIB in polarization depends on the foreground polarized surface brightness.

To estimate the fluctuations caused by background sources, we converted our simulated $Q_{\rm CIB}$ and $U_{\rm CIB}$ maps to MJy/sr and co-added them with a constant polarized foreground. For simplicity, we assume that this foreground is oriented on the Q direction and denote this constant foreground value as Q_f (by construction $U_f = 0$ MJy/sr). The value of the P map combining the two components is thus:

$$P_{\rm tot} = \sqrt{(Q_{\rm CIB} + Q_f)^2 + U_{\rm CIB}^2}. \tag{4}$$

If $Q_f \ll Q_{\rm CIB}$, the results are similar to the case discussed in Sect. 2.3 (except that the units are different). If $Q_f \gg Q_{\rm CIB}$, polarized CIB can be seen as a perturbation of the strong foreground signal:

$$\frac{\partial P_{\rm tot}}{\partial Q_{\rm CIB}} = \frac{Q_{\rm CIB} + Q_f}{P_{\rm tot}} \approx 1 \text{ and } \frac{\partial P_{\rm tot}}{\partial U_{\rm CIB}} = \frac{U_{\rm CIB}}{P_{\rm tot}} \ll 1. \tag{5}$$

We can thus see that the impact on the $P_{\rm tot}$ map depends on whether the CIB vector is aligned with the foreground or not. If they are in the same direction, the CIB component in the Q direction will thus add or remove polarized flux density compared to

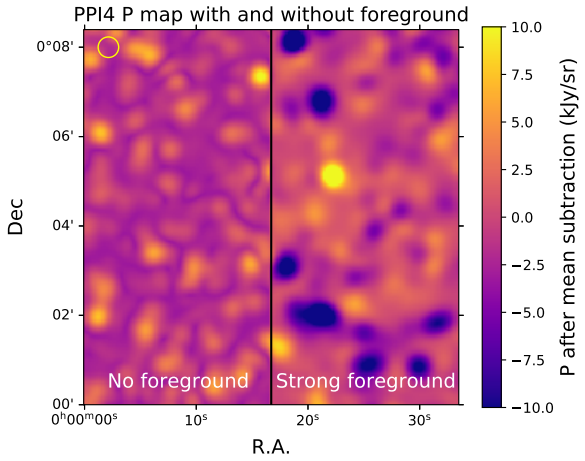

Fig. 11. Comparison between the PPI4 P maps in absence (only CIB, left side) and in presence (right side) of a strong foreground (1000 MJy/sr; see Sect. 5.5) illustrating the different behavior of the CIB confusion noise depending on the foreground strength. Since P is much higher on the right side, we subtracted the mean of each side to obtain a better visualization. The yellow circle in the top left corner shows the instrumental beam size.

the foreground alone. In contrast, the orthogonal component (U in our construction) has no first-order impact on $P_{\rm tot}$.

The impact of this asymmetry generated by the strong foreground is illustrated in Fig. 11. While the pure CIB map has mainly positive fluctuations, the sum of the CIB and the foreground exhibits both positive fluctuations (CIB and foreground polarization in the same direction) and negative fluctuations (orthogonal direction) at the position of the bright sources. We can also see that the fluctuations around the mean are larger in presence of a strong foreground.

We derive the classical confusion limit in the presence of the CIB and a foreground using a similar method as in Sect. 3.3. However, since negative sources can appear when the foreground is included, we iteratively mask all the 5σ outliers instead of only the positive ones. In Table 3, we tabulate the 1σ fluctuations generated by the CIB in absence and in the presence of a strong foreground. For the case of the strong foreground, we adopt $Q_f = 1000$ kJy/sr at each wavelength, which is more than five orders of magnitude above CIB fluctuations. The values obtained for a fainter foreground would be between these two extreme cases. Finally, since some science cases will need color maps with a matched resolution, we also derived the same quantity after degrading the beam size to the PPI4 resolution.

The fluctuations measured in the presence of a strong foreground are up to 20% higher than in the CIB-only case. Hence, this is a small but non-negligible effect. If we had masked only the positive 5σ outliers, the CIB fluctuations would have been up to 50% higher in the strong foreground case, but unchanged in the pure CIB case, since there are no strong negative fluctuations. Our table also shows that CIB fluctuations at PPI4 resolution are lower than at native resolution. In polarized surface brightness units, the signal from the constant foreground does not vary with the beam size, while a larger beam contains more sources and reduces the stochastic fluctuations.

Finally, we explored the impact of CIB on foreground color measurements. We use the "astrodust" model of dust emission and polarization (Hensley & Draine 2023) and have assumed that the dust is heated by a radiation field appropriate for diffuse atomic gas, to derive nominal input values of the PPI1/PPI4, PPI2/PPI4, and PPI3/PPI4 foreground colors of 0.52, 1.02, and 1.25, respectively. We varied the foreground polarized surface brightness fixing the input color, and derived the relative uncertainty on the measured foreground color produced by CIB fluctuations. We then interpolated between these values to determine the polarized surface brightness sensitivity limit corresponding to a 10% uncertainty (see Table 3). The foreground polarized surface brightness limits to reach a 10% precision on the foreground color are lower than ten times the 1σ CIB fluctuations, which is the limit expected based only on the numerator part of the color computation. However, as is shown in the last row of Table 3, the confusion noise is highly correlated. Positive and negative fluctuations of the CIB are thus expected to impact both bands in a similar way, reducing their impact on the ratio. The correlation between bands thus mitigates the confusion noise in such analyses.

We thus showed that the confusion noise from the polarized CIB depends on the properties of the foreground and is also strongly correlated between bands. Our work provides first estimates of the impact of CIB to study Galactic emission and nearby galaxies in polarization. More complex simulations including full foreground models will be key to preparing these science cases.

6. Consequences for surveys

6.1. Expected impact of confusion in intensity

In the previous sections, we discussed only the noiseless case corresponding to the best possible performance we could obtain for a given telescope diameter. However, it is crucial to compare the classical confusion limit with the expected instrumental performance. If the instrumental noise is much higher, the confusion can be ignored. If the instrumental noise is well below the confusion noise, advanced deblending methods will be required to make the most of the intrinsic sensitivity, but the performance may never fully match the instrumental noise.

In our analysis, we have considered two cases for the instrumental noise. The required payload sensitivity is the guaranteed performance of the instrument. It is a very conservative estimate to ensure high confidence in meeting the PI science goals. The real performance is expected to be much better, with a margin of at least 60% and in some cases by a much larger factor. We considered an intermediate estimate termed the conservative estimated sensitivity, which lies between the payload requirement and the actual estimated performance. To be consistent with the confusion, we also considered a 5σ limit. We treated the case of two fields observed 1500 h each. The deep field has a 1 deg^2 area, while the wide field covers 10 deg^2.

In Fig. 12, we compare the classical confusion limit with the instrumental sensitivity. The confusion has a steeper rise with increasing wavelength than the sensitivity for both sensitivity estimates. Consequently, the short wavelengths are noise-limited and the long wavelengths are confusion-limited. For the wide survey and a required payload sensitivity, the curves cross at 65 µm around the center of band PHI2. As is shown in Sect. 4.4 and Fig. 7, the red end of the PHI2 and the PPI bands at the confusion limit are not probing a part of the SFR-z space missed by the other sub-bands. However, PPI data are important to characterize the physics of the objects, and PHI priors will be crucial to deblend them. In the deep field with a required payload sensitivity, the classical confusion limit is reached at around 45 µm, and the PHI2 band will thus be affected by it. At this wavelength, the 7.7 µm PAH feature can be seen up to $z \sim 5$. (see Sect. 4.4).

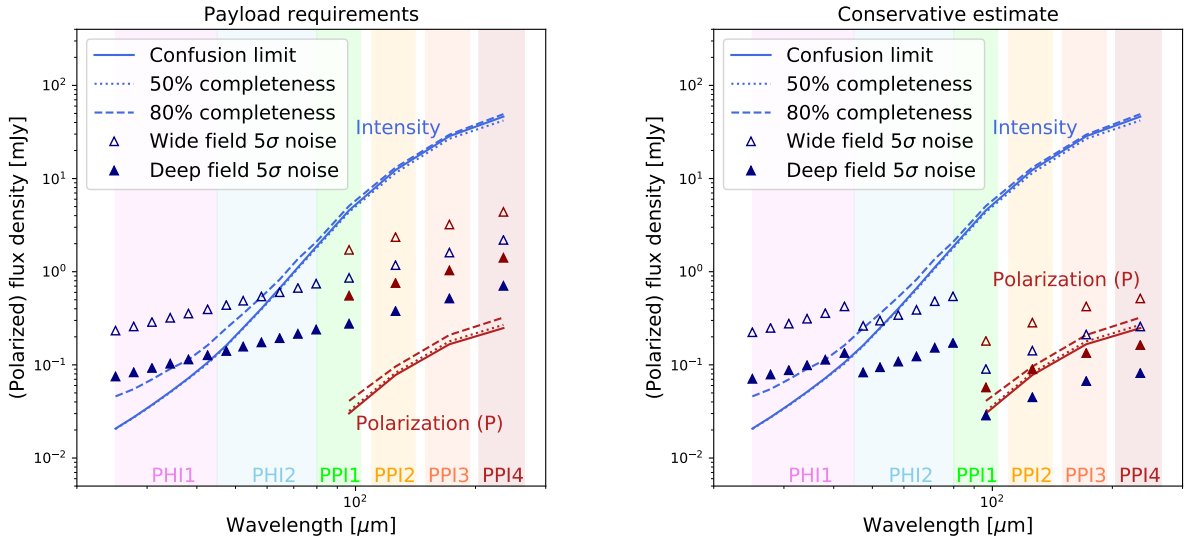

Fig. 12. Summary of the maximal depth reachable at the classical confusion limit as a function of wavelength and comparison with the expected PRIMAger instrumental depth. The left panel shows the survey depth for the required payload sensitivity, and the right panel corresponds to the conservative estimated sensitivities predicted by the instrumental teams. The open and filled upward triangles correspond to the 5σ instrumental sensitivity in the wide and deep surveys, respectively. The solid, dotted, and dashed lines are the classical confusion limit, 50%, and 80% completeness flux densities, respectively. The blue symbols correspond to quantities derived from intensity maps (discussed in Sect. 6.1) and the red from polarization maps (see Sect. 6.2). Note that the flux density of a given galaxy is a factor of ∼100 lower in polarization than in intensity, since the mean polarization fraction is 1%.

This ensures that we shall benefit from very deep priors up to this redshift before reaching the confusion, which will be essential to deblend the PHI2 and PPI bands and obtain more accurate physical constraints.

For the conservative estimated sensitivity, the confusion is reached at ∼55 µm and ∼45 µm in the wide and deep fields, respectively. We are thus confusion-limited in the PHI2 band, except in its bluest part for the wide fields. The SFR-z space probed will thus be similar to the pure confusion case discussed in Sect. 4.4 and Fig. 7. We also note that in PPI bands, the instrumental noise will be ∼2.5 orders of magnitude below the confusion. These data will thus be extremely close to the noiseless case discussed in this paper and will be ideal for applying the modern deep learning deblending algorithm (e.g., Lauritsen et al. 2021).

6.2. Feasibility of dust polarization surveys of distant galaxies

Far-infrared blank-field polarization surveys are in a totally uncharted territory. With our analysis, we can now set constraints on the expected classical confusion limit and we have demonstrated that we can recover constraints on the polarized flux density and angle of a galaxy (Sect. 5). However, since the signal will be fainter than in intensity, it is important to check if the instrumental sensitivity will be good enough to detect a large sample of sources. In this section, we discuss only our standard simulation assuming a mean polarization fraction μ_p of 1%. In Appendix C, we consider the alternative cases with 3σ-lower ($\mu_p = 0.7\%$) and 3σ-higher ($\mu_p = 1.3\%$) values (see Sect. 2.3).

In Fig. 12, we also compare the confusion and the instrumental limits in polarization. In PPI bands, the required payload and the conservative estimated sensitivities are very different. In the first case, the sensitivity limit in the deep fields are about an order of magnitude above the classical confusion limit (1.5 dex above for the wide). In the second more optimistic case, the wide field is noise-limited. In the deep field, the PPI1 and PPI2 band are close to the classical confusion limit, while the other bands have a sensitivity limit slightly below the classical confusion limit. This means that we should be very close to the confusion limited case in the SFR-z plane discussed in Sect. 7.

To evaluate the impact of these two hypotheses on the sensitivity, we used the SIDES simulation to predict the number of detections expected in the various cases. We combined quadratically the 5σ confusion and 5σ instrumental noise to obtain a secure polarized flux density limit, $P_{\rm lim}$, and used it to select the detectable sources in SIDES. Since the wide field is a factor of 5 larger than our simulation, we applied a scaling factor to the number of SIDES detections. A factor of 0.5 was applied for the deep field. The number of detections and their mean redshift and SFR are listed in Table 4. We also show the redshift and SFR distributions in Fig. 13.

For the required payload sensitivity, we expect ∼100 galaxies per field in PPI1, but fewer than 10 in PPI4. Since it is a totally unexplored parameter space, it will open a new window with small but statistical significant samples. In the deep field, half of the sources detected in polarization are below $z \leq 0.2$, and only a select number are above $z \geq 0.7$ with a tail up to $z \sim 1.5$. We thus trace mainly intermediate redshifts, though the dust polarization properties of galaxies at these epochs are currently totally unexplored. In terms of SFRs, we span a large range of SFRs from nearby 0.2 $M_\odot$/yr to high-z 1000 $M_\odot$/yr galaxies. In the wide field, objects are detected only up to $z = 0.4$. As was expected, the lowest SFR will not be probed. Paradoxically, we also observe fewer >100 $M_\odot$/yr systems than in the deep field. This is driven by the very low number density of these high-SFR objects at $z < 0.4$ and the polarized flux density limit being too high to be able to detect any high-z system.

With the conservative estimated sensitivity, the confusion and instrumental noise will be similar in the PPI1 and PPI2 bands. Several thousands of galaxies will be detected both in the deep and wide fields (Table 4). The number of detections in the deep field is a factor of ∼2 smaller than in the wide field. The PPI1 band will provide the highest number of detections, but

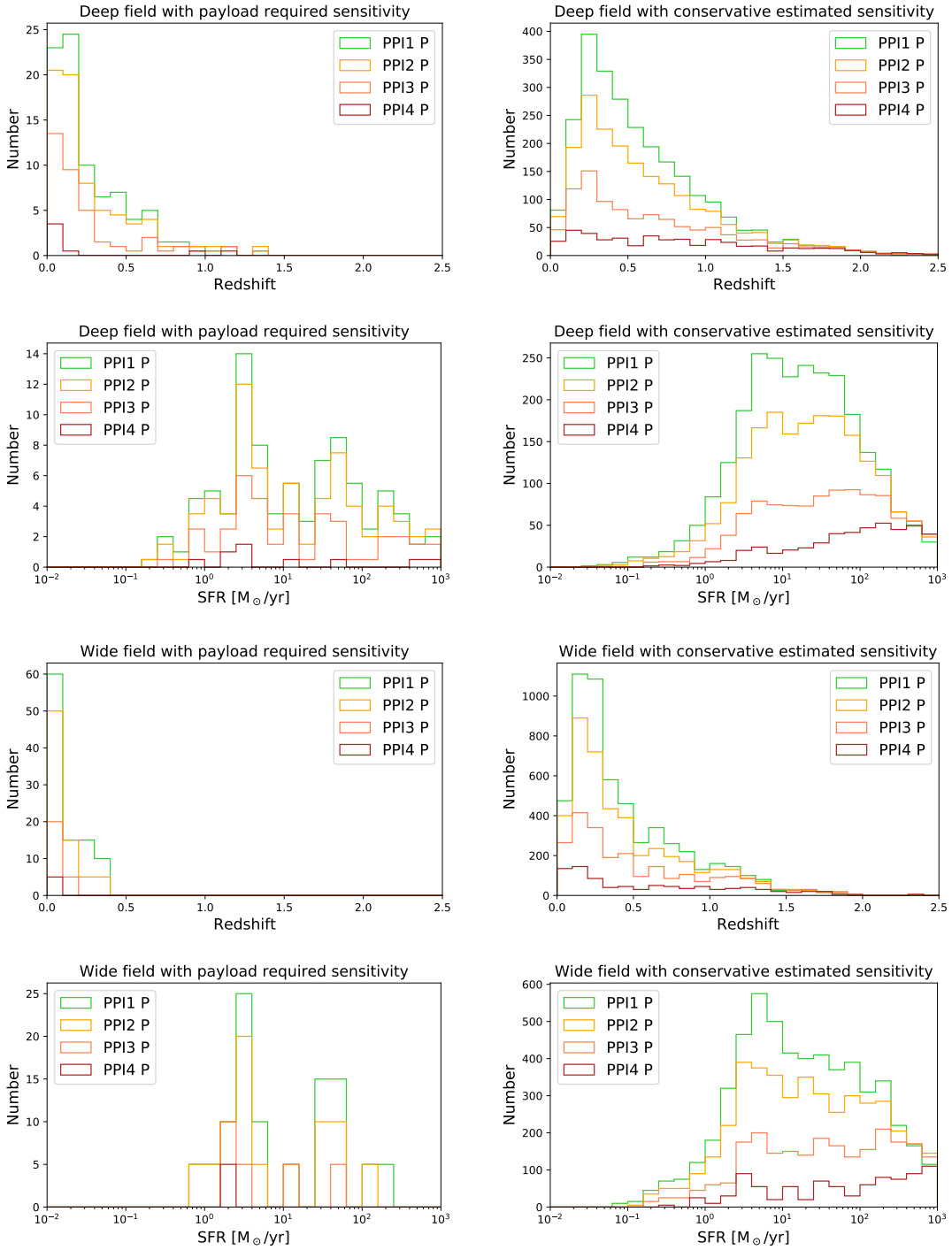

Fig. 13. Redshift and SFR distributions of the sources above the detection limit in polarized flux density $P_{\rm lim}$ (see Table 4 and Sect. 6.2). The left columns correspond to the required payload sensitivity and the right ones to the conservative estimated sensitivity. The rows are from top to bottom: redshift distribution in the deep field, SFR distribution in the deep field, redshift distribution in the wide field, SFR distribution in the wide field. The bands are color-coded, as is indicated in the figure.

even the PPI4 band will detect several hundred sources, enabling statistical studies of the polarized SEDs. Both deep and wide fields have a peak redshift distribution around $z \sim 0.4$ with large tail up to $z = 2.5$. In terms of SFRs, we cover five orders of magnitude from 0.01 to 1000 $M_\odot$/yr.

The required payload sensitivity would open a new window on the dust polarization of high-redshift galaxies with more than 100 detections up to $z \sim 1.5$. With the conservative estimated sensitivity, the results would be totally transformational by opening this new window directly with several thousands of sources up to $z = 2.5$.

6.3. Synergies between intensity and polarization surveys

Independently of the sensitivity scenario, the PHI1 band will be dominated by the instrumental noise, while the PPI bands will always be confusion-limited in intensity. However, the classical confusion limit in polarization is more than 100 times smaller

Table 4. Number of expected detections (N_{det}) above the polarized flux density limit P_{lim} (quadratic combination of the 5σ confusion and 5σ instrumental noise; see Sect. 6.2), their mean redshift, and their mean SFR for a deep and a wide 1500 h PRIMAger survey assuming the required payload sensitivity and the conservative estimated sensitivity.

Band	P_{lim} μJy	N_{det}	Mean z	Mean SFR $M_\odot$/yr
Deep field (1500 h, 1 deg^2) with required payload sensitivity				
PPI1	553	86	0.27	71
PPI2	760	72	0.28	84
PPI3	1045	37	0.27	94
PPI4	1432	5	0.26	153
Wide field (1500 h, 10 deg^2) with required payload sensitivity				
PPI1	1718	100	0.12	28
PPI2	2353	75	0.10	22
PPI3	3215	25	0.07	12
PPI4	4394	5	0.03	2
Deep field (1500 h, 1 deg^2) with conservative estimated sensitivity				
PPI1	65	2546	0.59	61
PPI2	119	1940	0.62	78
PPI3	214	1050	0.68	119
PPI4	297	474	0.86	202
Wide field (1500 h, 10 deg^2) with conservative estimated sensitivity				
PPI1	183	5510	0.45	79
PPI2	294	4295	0.47	98
PPI3	455	2375	0.54	145
PPI4	573	880	0.63	215

than in intensity, and the confusion will only be reached in the deep field in the optimistic sensitivity scenario. This opens the opportunity for synergistic strategies between intensity and polarization.

If we undertake deep integrations that approach the classical intensity confusion limit in the PHI1 band, the PHI2 and PPI band will be limited by confusion in intensity. However, depending on the exact sensitivity ratio between bands, PPI bands may still not be affected by confusion in polarization. In addition, it will also be possible to deblend the PHI2 band using a prior-based extraction algorithm (Donnellan et al. 2024). All the bands will thus be used efficiently in such a strategy, and we shall fully exploit the high PPI sensitivity through the polarization. The risk of attempting a first high-z deep polarization survey will also be mitigated, since we shall get extremely deep intensity data at shorter wavelengths at the same time. PRIMAger is thus a very promising instrument that is able to open two new windows of survey parameter space with a single deep field observation.

7. Conclusion

We produced simulated PRIMAger data (Fig. 1) using the SIDES simulation to study how confusion impacts the performance of basic blind source extractors, both in intensity and polarization. With this, we determined the classical confusion limit for all PRIMA bands, which increases steeply with increasing wavelength (Fig. 12).

For the conservative estimated sensitivities of the PRIMAger wide and deep surveys, the classical confusion limit curve crosses the sensitivity limit at approximately 60 μm and 45 μm, respectively. Taking advantage of the available instrument sensitivity at longer wavelengths requires the use of deblending methods. The PRIMAger hyperspectral architecture, which produces finely sampled $R = 10$ SEDs, is particularly good at enabling these methods by providing priors for sources detected at shorter, unconfused wavelengths. A companion paper (Donnellan et al. 2024) analyzes the performance of a particular deblending approach (XID+, Hurley et al. 2017), showing that its application will recover fluxes out to $\lambda = 100$ μm and beyond for astrophysical SEDs. Moreover, we show that in polarization the confusion limit is more than two orders of magnitude lower than in intensity. Surveys will thus be limited by instrument sensitivity, except at $\lambda > 150$ μm in the deep field.

We have studied the effect of galaxy clustering, showing that it has a mild impact on confusion in intensity (<25%), while its effect on polarization is very small (Fig. 5). This difference in behavior is explained by the respective scalar and vectorial natures of intensity and polarization.

The measured flux density in intensity for $\lambda > 100$ μm is on average larger than the flux density of the brightest galaxy in the beam, because of contamination from confused sources (Fig. 3). In contrast, the polarized flux density and polarization angle measurements are essentially unaffected (Fig. 8). The polarization fraction measurements are affected, however, because they are derived from both intensity and polarization measurements.

We computed the probability of detecting a galaxy above the classical confusion limit as a function of its position in the SFR-z plane (Figs. 7 and 10). In intensity, galaxies at the knee of the infrared luminosity function ($L_\star$) will be above the classical confusion limit in at least one band up to $z \sim 3$, while massive (10^{11} $M_\odot$) main-sequence galaxies can be recovered up to $z \sim 5$. In polarization, PRIMager opens up a brand new parameter space by enabling studies for $L_\star$ and massive main-sequence galaxies that are brighter than the classical confusion limit up to $z \sim 0.5$ and $z \sim 1.5$, respectively. We can also observe a tail of extreme objects up to $z \sim 2.5$.

We estimate the effect of the background of polarized emission due to high-z galaxies on measurements of polarized emission of extended low-z foreground sources (Table 3). The 1-σ noise on polarized surface brightness can vary from 1.2 to 3.4 kJy/sr depending on the band and how bright the foreground is. This noise is correlated between bands, which has an impact on the uncertainties of foreground color measurements. We estimated that a minimum foreground polarized surface brightness of 11 kJy/sr at 96 μm is necessary to obtain a 10% precision on the foreground P_{96}/P_{235} color.

Finally, assuming the PRIMAger conservative estimated sensitivity, we expect several thousands of detections of the integrated dust polarization of high-z galaxies (up to $z \sim 2.5$) in both the deep and wide fields (Fig. 13 and Table 4). Considering that polarization has been reported on only two lensed high-z galaxys so far (Geach et al. 2023; Chen et al. 2024), PRIMAger has the potential to revolutionize this field of study by producing large statistical samples.

Detecting sources in polarization requires a ~100 times better sensitivity than necessary to detect the same sources in intensity at the same wavelength. Opening this new window on high-z dust polarization thus obviously has a cost in observing time. However, it can be offset by the detector intrinsic sensitivity. Indeed, for a deep PRIMAger survey (1 deg^2 in 1500 h) and our conservative estimated sensitivities, both the short-wavelength

data in intensity (PHI1) and the polarization maps at longer wavelengths (PPI) will be close to the confusion limit of the instrument. There is thus a strong synergy between pushing the instrument to this limit in intensity at short wavelengths, and performing deep polarization surveys at long wavelengths. At this depth, the long-wavelength data in intensity will be fully dominated by confusion rather than by the extremely low instrumental noise, but analysis methods using short-wavelength priors are expected to allow us to deblend sources efficiently using short-wavelength data (Donnellan et al. 2024). A single survey will thus obtain well-sampled mid- to far-infrared SEDs in intensity, opening up new opportunities for star formation and AGN science (Bisigello et al. 2024), together with long-wavelength far-infrared data in polarization, providing new constraints on galactic magnetic fields.

Acknowledgements. This research made use of Photutils, an Astropy package fordetection and photometry of astronomical sources (Bradley et al. 2023). James Donnellan was supported by the Science and Technology Facilities Council, grant number ST/W006839/1, through the DISCnet Centre for Doctoral Training. Seb Oliver acknowledges support from the UK Space Agency through ST/Y006038/1. Guilaine Lagache This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788212). Portions of this work were carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). The contributions of the authors using the Contributor Roles Taxonomy (CRediT) were as follows: M. Bethermin: Conceptualization, Methodology, Software, Validation, Investigation, Writing – original Draft; A. Bolatto, L. Ciesla, J. Glenn, B. Hensley, G. Lagache, E. Lopez-Rodriguez, S. Oliver, and A. Pope: Conceptualization, Writing – Review & Editing; D. Burgarella and J. Donnellan: Writing – Review & Editing; C. Bradford, F. Boulanger, and M. Sauvage: Conceptualization.

References

André, P., Hughes, A., Guillet, V., et al. 2019, PASA, 36, e029
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Berta, S., Magnelli, B., Nordon, R., et al. 2011, A&A, 532, A49
Béthermin, M., Dole, H., Lagache, G., Le Borgne, D., & Penin, A. 2011, A&A, 529, A4
Béthermin, M., Daddi, E., Magdis, G., et al. 2012a, ApJ, 757, L23
Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012b, A&A, 542, A58
Béthermin, M., Daddi, E., Magdis, G., et al. 2015, A&A, 573, A113
Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, A&A, 607, A89
Béthermin, M., Gkogkou, A., Van Cuyck, M., et al. 2022, A&A, 667, A156
Bing, L., Béthermin, M., Lagache, G., et al. 2023, A&A, 677, A66
Bisigello, L., Gruppioni, C., Bolatto, A., et al. 2024, A&A, 689, A125
Bradley, L., Sipőcz, B., Robitaille, T., et al. 2023, astropy/photutils: 1.8.0, https://zenodo.org/records/7946442
Burgarella, D., Ciesla, L., Sauvage, M., et al. 2023, Am. Astron. Soc. Meeting Abstr., 55, 160.22
Chen, J., Lopez-Rodriguez, E., Ivison, R. J., et al. 2024, ArXiv e-prints [arXiv:2407.14596]
Codis, S., Jindal, A., Chisari, N. E., et al. 2018, MNRAS, 481, 4753
Condon, J. J. 1974, ApJ, 188, 279
Cowley, W. I., Lacey, C. G., Baugh, C. M., & Cole, S. 2015, MNRAS, 446, 1784
Dole, H., Le Floc'h, E., Pérez-González, P. G., et al. 2004, ApJS, 154, 87
Dole, H., Lagache, G., Puget, J., et al. 2006, A&A, 451, 417
Donnellan, J. M. S., Oliver, S. J., Béthermin, M., et al. 2024, MNRAS, 532, 1966
Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, ApJ, 508, 123
Frayer, D. T., Fadda, D., Yan, L., et al. 2006, AJ, 131, 250
Geach, J. E., Lopez-Rodriguez, E., Doherty, M. J., et al. 2023, Nature, 621, 483
Gkogkou, A., Béthermin, M., Lagache, G., et al. 2023, A&A, 670, A16
Glenn, J., Conley, A., Béthermin, M., et al. 2010, MNRAS, 409, 109
Hauser, M. G., & Dwek, E. 2001, ARA&A, 39, 249
Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, ApJ, 508, 25
Hayward, C. C., Behroozi, P. S., Somerville, R. S., et al. 2013, MNRAS, 434, 2572
Hensley, B. S., & Draine, B. T. 2023, ApJ, 948, 55
Hurley, P. D., Oliver, S., Betancourt, M., et al. 2017, MNRAS, 464, 885
Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, MNRAS, 432, 2
Lagache, G., Haffner, L. M., Reynolds, R. J., & Tufte, S. L. 2000, A&A, 354, 247
Lagache, G., Dole, H., & Puget, J.-L. 2003, MNRAS, 338, 555
Lagache, G., Puget, J., & Dole, H. 2005, ARA&A, 43, 727
Lagache, G., Béthermin, M., Montier, L., Serra, P., & Tucci, M. 2020, A&A, 642, A232
Lauritsen, L., Dickinson, H., Bromley, J., et al. 2021, MNRAS, 507, 1546
Lopez-Rodriguez, E. 2023, ApJ, 953, 113
Lopez-Rodriguez, E., Alonso-Herrero, A., Diaz-Santos, T., et al. 2018, MNRAS, 478, 2350
Lopez-Rodriguez, E., Mao, S. A., Beck, R., et al. 2022, ApJ, 936, 92
Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, A90
Magnelli, B., Elbaz, D., Chary, R. R., et al. 2009, A&A, 496, 57
Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132
Marin, F., Le Cam, J., Lopez-Rodriguez, E., et al. 2020, MNRAS, 496, 215
Meixner, M., Burgarella, D., Ciesla, L., et al. 2024, Am. Astron. Soc. Meeting Abstr., 56, 457.13
Moullet, A., Kataria, T., Mills, E., et al. 2024, Am. Astron. Soc. Meeting Abstr., 56, 360.30
Oliver, S. J., Wang, L., Smith, A. J., et al. 2010, A&A, 518, L21
Papovich, C., Dole, H., Egami, E., et al. 2004, ApJS, 154, 70
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1
Planck Collaboration XII. 2020, A&A, 641, A12
Planck Collaboration XXX. 2014, A&A, 571, A30
Puget, J., Abergel, A., Bernard, J., et al. 1996, A&A, 308, L5
Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, MNRAS, 409, 48
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Scudder, J. M., Oliver, S., Hurley, P. D., et al. 2016, MNRAS, 460, 1119
Singh, S., Mandelbaum, R., & More, S. 2015, MNRAS, 450, 2195
Traina, A., Gruppioni, C., Delvecchio, I., et al. 2024, A&A, 681, A118
Viero, M. P., Wang, L., Zemcov, M., et al. 2013, ApJ, 772, 77
Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1

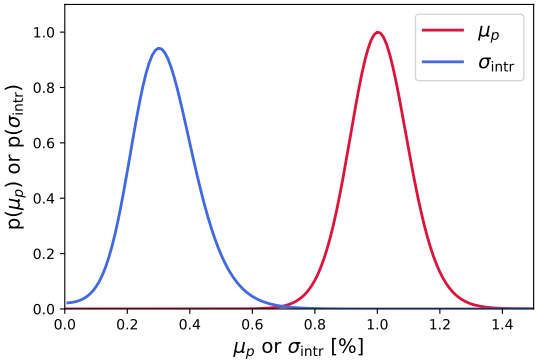

Fig. A.1. Probability density function of the central value of the polarization fraction (μ_p) and the scatter ($\sigma_{\rm intr}$) around it. These distributions were determined based on the local Universe sample of Lopez-Rodriguez et al.(2022, see Sect. 2.3).

Appendix A: Calibration of the polarization fraction using data from the local Universe

To calibrate the intrinsic distribution of the polarization fraction, we used the local-Universe SALSA sample from Lopez-Rodriguez et al. (2022). We discarded the Circinus value at 214 µm, which is a very strong outlier (8.4 %). It is a Seyfert object, which may not be representative of typical galaxy populations. It showed large polarization fractions located in the interarm regions a few kiloparsecs from the central AGN. This specific observation only had a few dozen polarization measurements due to a shallow integration time, and the significance of final polarization measurements is 3–5 σ.

We fitted the mean polarization fraction (μ_p) and the intrinsic scatter around it ($\sigma_{\rm intr}$) using their measured integrated polarization fractions (their Table 5). The likelihood $\mathcal{L}$ is computed using:

$$\ln(\mathcal{L}) = \sum_{\rm i=1}^{N_{\rm sample}} -\frac{\ln(2\pi) + \ln(\sigma_{\rm intr}^2 + \sigma_{\rm mes,i}^2)}{2} - \frac{(p_{\rm mes,i} - \mu_p)^2}{2(\sigma_{\rm intr}^2 + \sigma_{\rm mes,i}^2)}, \quad \text{(A.1)}$$

where $p_{\rm mes,i}$ is the integrated polarized fraction of the i-th object, and $\sigma_{\rm mes,i}$ is measurement uncertainty on it. By combining all the objects and the wavelengths, we found $\mu_p = 1.0^{+0.1}_{-0.1}$ % and $\sigma_{\rm intr} = 0.3^{+0.1}_{-0.1}$ %. The marginalized probability density functions of both parameters are shown in Fig. A.1. We thus drew p from a Gaussian with $\mu_p = 1.0$ and $\sigma_{\rm intr} = 0.3$, and replaced negative values with zero.

In the current version of SIDES including the polarization, we use the same probability function to draw the polarization fraction p for all galaxies (see Sect. 2.3). In Fig. A.2, we show the probability density of the means and standard deviations obtained for various subsamples from the SALSA survey (Lopez-Rodriguez et al. 2022). The starbursts have marginally higher mean polarization fraction and standard deviation than non-starbursts, but the offset between the two probability density is much smaller than the uncertainties. The probability of the values found for the full sample corresponds to 0.95 times the peak probability density of the subsamples. There is thus no significant dependence detected by our data-driven approach.

The wavelength-dependence is more complex to interpret. Both 53 and 89 µm have peak probability densities around 1.1 % of polarization, while 154 µm peaks around 0.8 % and 214 µm peaks around 1.6 % with large uncertainties. It could be a hint of the signature from inner outflows (Lopez-Rodriguez 2023,, their Fig. 10 and 11). However, the best-fit value of 1 % found combining all wavelengths has a probability larger than 0.2 at all wavelengths. It is thus hard to conclude with such a low probability. Lopez-Rodriguez (2023) used the resolved polarization data to consolidate their results, but in this paper we consider only integrated properties since high-z galaxies will not be spatially resolved.

Appendix B: Impact of the instrument pixels

In our main analysis, we used a beam FWHM driven only the optics. However, the current design of PRIMAger has pixel sizes that undersample the PSF (particularly at short wavelengths). Pixel sizes range from $1.25\lambda/D$ to $0.7\lambda/D$ for the PHI pixels in the wavelength ranges 23 to 45 µm (PHI1) and 45 to 82 µm (PHI2). We repeated our analysis using a beam that is broadened by the response of the pixel. To compute this broaden beam, we convolve the diffraction-limited response of the system with the response of the pixel, modeled as a top-hat function with the diameter corresponding to the wavelength range. This estimate is conservative since the effects of pixel size can be mitigated in part by the use of sub-pixel dithers and drizzle-type map making algorithms. These calculations are thus an upper limit on the possible impact. Moreover, the instrument pixel size will likely be further optimized to reduce undersampling effects.

In Fig. B.1, we show the ratio between the results obtained with the broader beam including the impact of the pixels and the narrower beam corresponding to the optics only. The impact on the 5σ classical confusion limit and the 50 % completeness flux density is similar. There is also no significant difference between intensity and polarization in PPI bands. We find a difference by a factor of 1.6 in the blue side of both PHI bands, decreasing down to 1.2 in the red side of each band. There is thus a strong jump between band PHI1 and PHI2, which can be explained by the constant pixel size in a given band while the beam size is increasing from the blue to the red. The relative impact of the pixel size is thus stronger in the blue, where the beam is the smallest. In PPI bands, the effect is of the order of 20 %. Finally, the 80 % completeness flux density is slightly less impacted in the PHI1 band, and it could be due to the flatter completeness curves discussed in Sect. 4.2.

Appendix C: Impact of the mean polarization fraction on the expected number of detections

The confusion noise scales linearly with the mean polarization fraction μ_p. Since the polarized flux density is also proportional to μ_p, the expected number of detections in polarization surveys affected only by confusion is independent of μ_p. The situation is more complex in presence of instrumental noise. When μ_p decreases, the polarized flux density of the sources and the polarized confusion noise will both decrease, but the instrumental noise will remain constant. In the regime where the instrumental noise is dominant, the number of detections will thus decrease. In this appendix, we discuss a low (μ_p=0.7 %) and a high (μ_p=1.3 %) mean polarization fraction corresponding to the 3σ bounds of the probability density distribution of μ_p determined in Appendix A.

In Table C.1 and C.2, we present the number of polarized detections expected for a low and high mean polarization fraction, respectively. They can can be compared with the results from our standard simulation (see Table 4). The number of detections remains consistent within a factor of ~2 with the standard

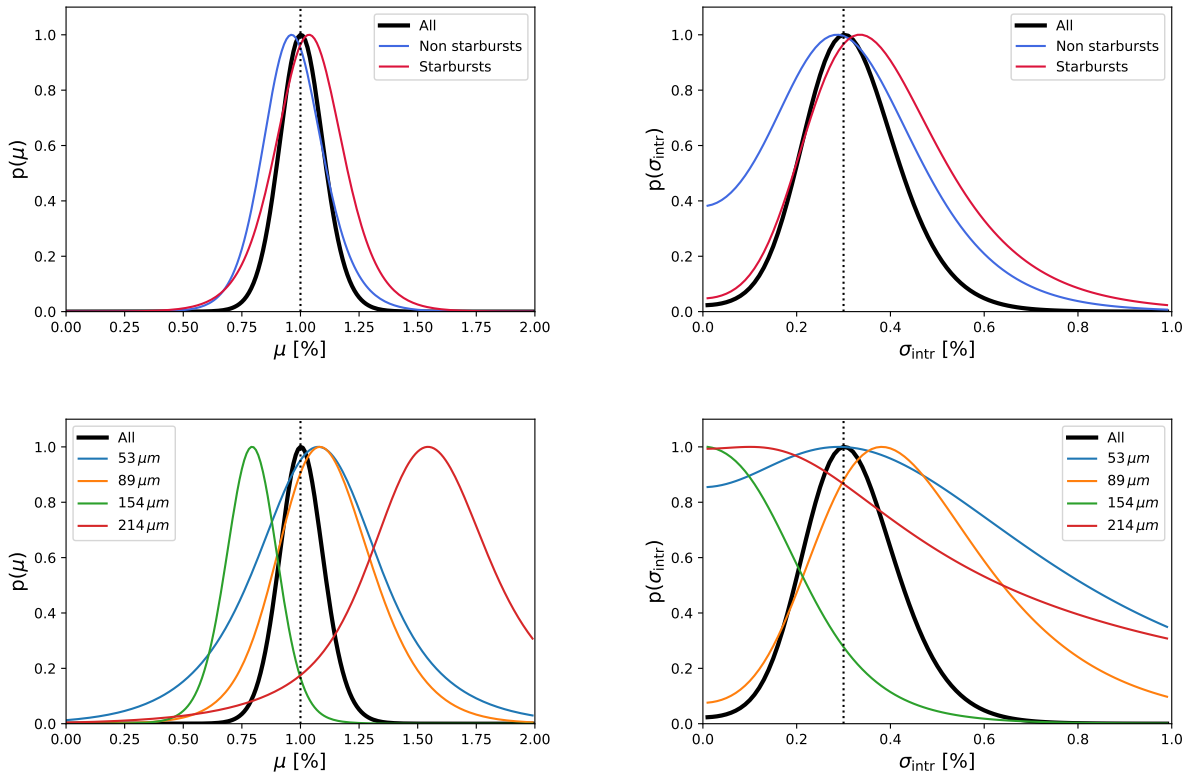

Fig. A.2. Probability density of the mean (left panels) and the standard deviation (right panels) of the polarization fraction determined using the method described in Sect. 2.3. The top and bottom panels show the dependence with the presence of a starburst and the wavelength, respectively. The thick black curve is the result obtained with the full sample and the dotted vertical line is the most probable value. The colored curve are obtained using subsamples.

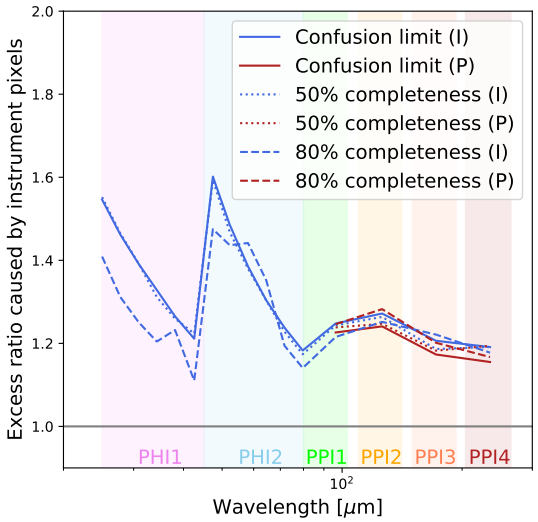

Fig. B.1. Same as Fig. 5 but for the effect of the instrument pixels on the beam FWHM in absence of thin dithers or other mitigations. The ratio corresponds to the same clustered input catalog, but with different beam FWHM.

simulation, except for the wide field in the PPI3 and PPI4 bands assuming required payload sensitivity for which the number of detection is very low in the low μ_p case. This uncertainty by a factor of ∼2 is much smaller than the variations by more than an order of magnitude found for our two different sensitivity scenarios.

Finally, in Fig. C.1 and C.2, we present the redshift and SFR distributions for our various polarization and sensitivity scenarios. The impact of μ_p is mild. As is shown in Table C.1 and C.2, the mean redshift and SFR varies by less than a factor of ∼2 (after excluding the cases with $\lesssim$10 objects, which are not statistically significant). For the conservative estimated sensitivity with larger numbers of detections, the variations are lower than ∼20 %.

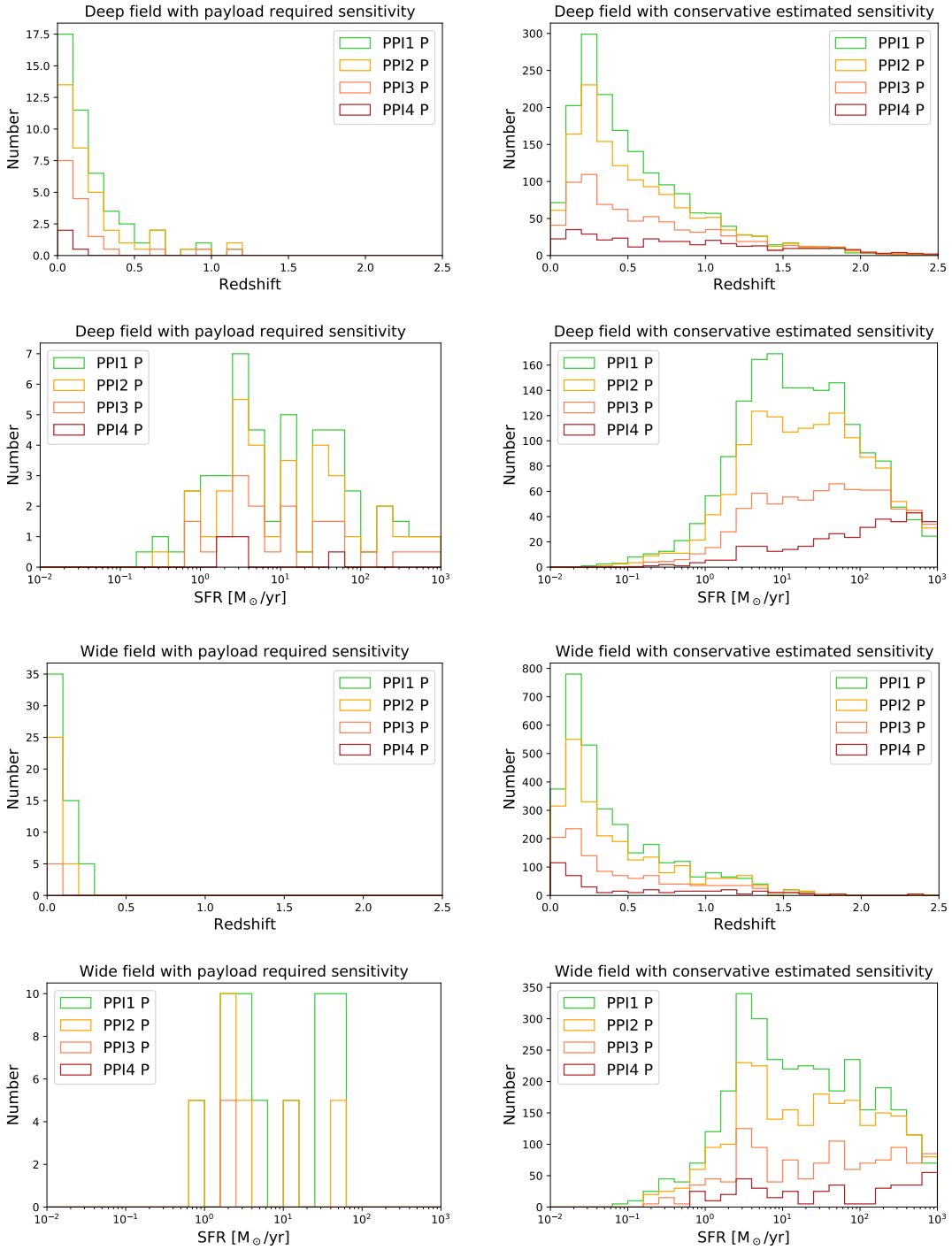

Fig. C.1. Same as Fig. 13 but assuming a mean polarized fraction of 0.7 %.

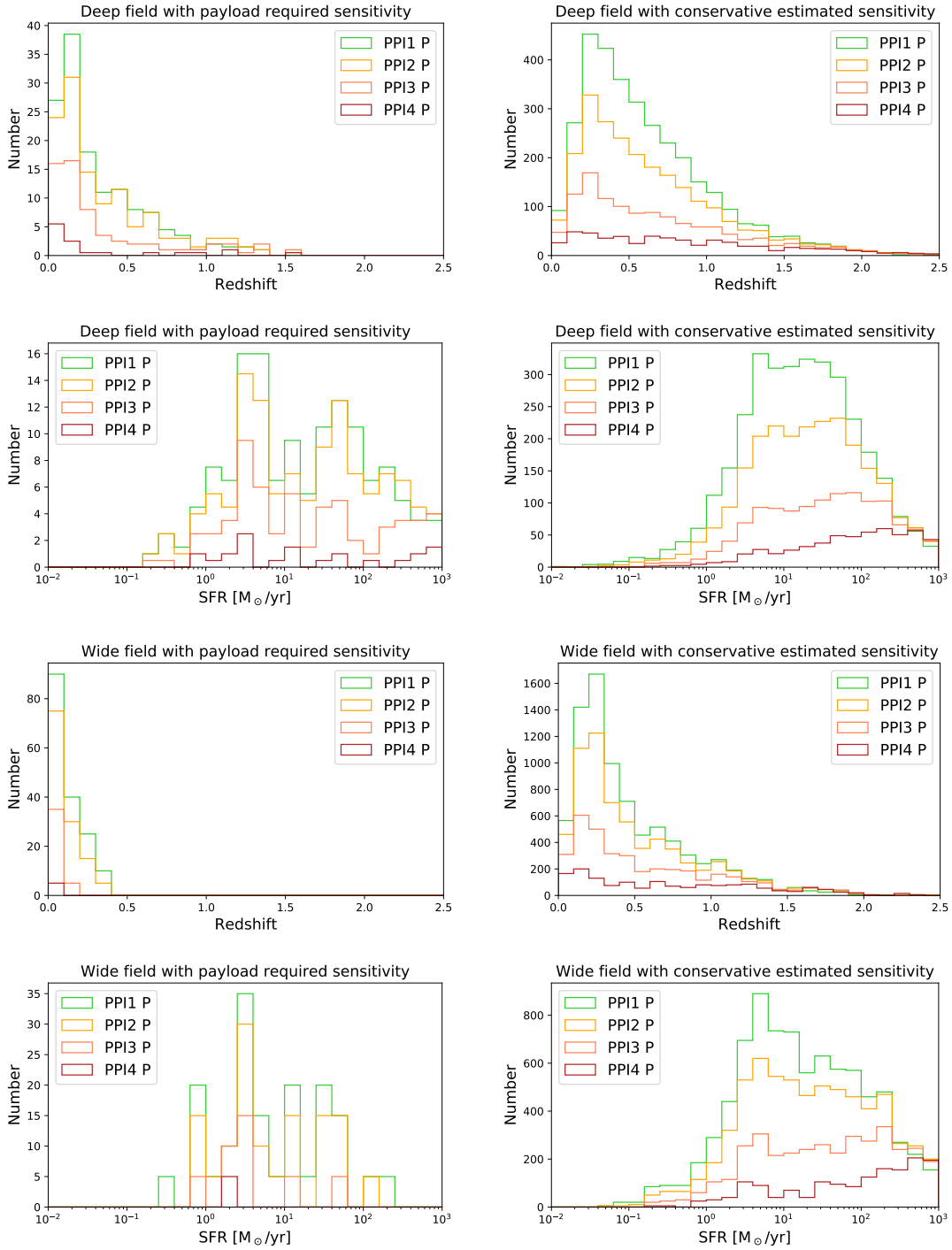

Fig. C.2. Same as Fig. 13 but assuming a mean polarized fraction of 1.3 %.

Table C.1. Same as Table 4, but assuming a mean polarized fraction of 0.7 %.

Band	P_{lim} μJy	N_{det}	Mean z	Mean SFR $M_\odot$/yr
Deep field (1500h, 1deg^2) with required payload sensitivity				
PPI1	552	46	0.22	63
PPI2	758	34	0.23	75
PPI3	1039	16	0.19	72
PPI4	1421	2	0.07	11
Wide field (1500h, 10deg^2) with required payload sensitivity				
PPI1	1718	55	0.09	18
PPI2	2353	30	0.07	11
PPI3	3213	5	0.03	2
PPI4	4390	0	–	–
Deep field (1500h, 1deg^2) with conservative estimated sensitivity				
PPI1	61	1669	0.54	66
PPI2	105	1344	0.58	84
PPI3	178	768	0.66	127
PPI4	238	352	0.86	225
Wide field (1500h, 10deg^2) with conservative estimated sensitivity				
PPI1	182	3145	0.40	85
PPI2	289	2345	0.43	108
PPI3	439	1155	0.47	150
PPI4	545	400	0.53	212

Table C.2. Same as Table 4 but assuming a mean polarized fraction of 1.3 %.

Band	P_{lim} μJy	N_{det}	Mean z	Mean SFR $M_\odot$/yr
Deep field (1500h, 1deg^2) with required payload sensitivity				
PPI1	553	136	0.32	78
PPI2	763	119	0.35	96
PPI3	1054	61	0.36	128
PPI4	1446	12	0.34	177
Wide field (1500h, 10deg^2) with required payload sensitivity				
PPI1	1718	165	0.12	20
PPI2	2354	125	0.11	17
PPI3	3218	40	0.06	9
PPI4	4398	5	0.03	2
Deep field (1500h, 1deg^2) with conservative estimated sensitivity				
PPI1	69	3288	0.61	57
PPI2	135	2366	0.64	75
PPI3	255	1242	0.70	113
PPI4	362	553	0.87	194
Wide field (1500h, 10deg^2) with conservative estimated sensitivity				
PPI1	185	8195	0.47	73
PPI2	301	6565	0.51	94
PPI3	476	3660	0.58	139
PPI4	609	1575	0.75	239