

The Farmer:  
**A reproducible profile-fitting photometry package for deep galaxy surveys**

2 J. R. WEAVER,<sup>1,2,3</sup> L. ZALESKY,<sup>4,2</sup> V. KOKOREV,<sup>5,1,2</sup> C. J. R. MCPARTLAND,<sup>1,4,2</sup> N. CHARTAB,<sup>6,7</sup> K. M. L. GOULD,<sup>1,2</sup>  
3 M. SHUNTOV,<sup>8</sup> I. DAVIDZON,<sup>1,2</sup> A. FAISST,<sup>9</sup> N. STICKLEY,<sup>9</sup> P. L. CAPAK,<sup>1</sup> S. TOFT,<sup>1,2</sup> D. MASTERS,<sup>10,9,\*</sup> B. MOBASHER,<sup>11</sup>  
4 D. B. SANDERS,<sup>4</sup> O. B. KAUFFMANN,<sup>12</sup> H. J. McCracken,<sup>8</sup> O. ILBERT,<sup>12</sup> G. BRAMMER,<sup>1,2</sup> AND A. MONETI<sup>8</sup>

5 <sup>1</sup>*Cosmic Dawn Center (DAWN)*

6 <sup>2</sup>*Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark*

7 <sup>3</sup>*Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA*

8 <sup>4</sup>*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*

9 <sup>5</sup>*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands*

10 <sup>6</sup>*The Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101, USA*

11 <sup>7</sup>*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

12 <sup>8</sup>*Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France*

13 <sup>9</sup>*Infrared Processing and Analysis Center, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA, 91125, USA*

14 <sup>10</sup>*Jet Propulsion Laboratory, California Institute of Technology 4800 Oak Grove Drive Pasadena, CA 91109, USA*

15 <sup>11</sup>*Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA*

16 <sup>12</sup>*Aix Marseille Univ, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France*

17 (Received July 25, 2023; Revised July 25, 2023; Accepted July 25, 2023)

18 Submitted to ApJS

19 ABSTRACT

20 While space-borne optical and near-infrared facilities have succeeded in delivering a precise and  
21 spatially resolved picture of our Universe, their small survey area is known to under-represent the  
22 true diversity of galaxy populations. Ground-based surveys have reached comparable depths but at  
23 lower spatial resolution, resulting in source confusion that hampers **accurate** photometry extractions.  
24 What once was limited to the infrared regime has now begun to challenge ground-based ultra-deep  
25 surveys, affecting detection and photometry alike. Failing to address these challenges will mean for-  
26 feiting a representative view into the distant universe. We introduce **The Farmer**: an automated,  
27 reproducible profile-fitting photometry package that pairs a **library of smooth parametric models**  
28 from **The Tractor** (Lang et al. 2016) with a decision tree that chooses a best-fit model in concert  
29 with neighboring sources. Photometry is measured by fitting the models on bands of interest leaving  
30 the brightness free to vary. The resulting photometric measurements are naturally total, and no aper-  
31 ture corrections are required. Supporting diagnostics (e.g.  $\chi^2$ ) enable measurement validation. **As**  
32 **fitting models is relatively time intensive**, The Farmer is built with high-performance computing  
33 routines. We benchmark **The Farmer** on a set of realistic COSMOS-like images and find **accurate**  
34 photometry, number counts, and galaxy shapes. **The Farmer** is already being utilized to produce cat-  
35 alogs for several large-area deep extragalactic surveys where it has been shown to tackle some of the  
36 most challenging optical and near-infrared data sets available, with the promise of extending to other  
37 ultra-deep surveys expected in the near future.

38 **Keywords:** photometry – astrostatistics – catalogs

Corresponding author: John R. Weaver

[john.weaver.astro@gmail.com](mailto:john.weaver.astro@gmail.com)

\* NASA Postdoctoral Program (NPP) Fellow

39 1. INTRODUCTION

40 For most of its history, astronomy has been defined  
41 by the use of electro-magnetic waves to measure sources

detected in the night sky. What began as a purely visual study was transformed in the late 19<sup>th</sup> century with the advent of photographic plates that enabled precise observations from which the brightness of sources could be measured (Bigourdan 1888). It was with such comparatively primitive technology that the first variable stars in Andromeda were identified, leading to the discovery of the ‘island universes’ and later the expansion of the universe (Hubble 1926, 1929). Now almost a century later, all scientific astronomical observations are captured on Charge-Coupled Devices, or CCDs (Lesser 2015), further enhancing the accuracy and precision of photometry.

Photometry itself has for decades been performed using apertures. That is, the integrated flux or total brightness of a source is computed within apertures of a fixed size. This is especially useful for isolated, unresolved, point-like sources like stars, quasars, and distant galaxies whose spatial appearance is well-described by the point-spread function (PSF) determined by the optical train of the telescope. While larger apertures ensure all of the light is captured and are less susceptible to noise, they may unintentionally capture light from other nearby sources which is usually mitigated by smaller apertures, although with typically greater uncertainties. Images with high source density, arising either from physically compact structures (e.g., star clusters) or from background and foreground sources appearing in close proximity on the sky, may require apertures smaller than the PSF (or alternative mitigation strategies, see Stetson 1987; Bertin & Arnouts 1996). Recovering the total flux in such cases requires scaling the aperture-integrated flux proportional to the total extent of the PSF, which often involves complicated strategies to characterize the PSF stability across the detector or co-added mosaic. Transitioning from monochrome photometry of a single band to photometering multi-wavelength images presents its own challenge as PSFs and pixel sizes typically vary with the filter as well as telescope, instrument, and observing conditions. The solution has been a procedure known as PSF homogenization (or matching) whereby each image is convolved with a kernel that maps the PSF of that particular image to that of a target PSF, typically requiring re-sampling images to a common pixel scale. Not only is the choice of the target PSF not always well-defined, especially in cases where the PSF characteristics vary significantly between bands, but re-sampling images often induces or increases pixel-to-pixel covariance.

For applications in extragalactic studies, the deepest wide-field ground-based near-infrared survey at the time of writing is UltraVISTA (McCracken et al. 2012)

which at a uniform  $K_s \approx 26$  AB depth captures  $\sim 150$  sources per arcmin<sup>2</sup> over 2 deg<sup>2</sup> with resolution set by its  $\sim 0.51''$  PSF at FWHM. Consequently, modest apertures of 3'' diameter can be contaminated by neighboring sources. In the corresponding source catalog of Weaver et al. (2022), 2'' diameter apertures are adopted when measuring photometry to be used in spectral fitting, which in the case of some high-redshift ( $z \gtrsim 7$ ) galaxies remain contaminated such that interloping blue light does not permit a high-redshift solution (Kauffmann et al. 2022). While manually removing such interlopers in small samples is possible, doing so for several thousand becomes impractical, and risks imposing human biases. Until the operation of space observatories such as *Euclid* and *Roman*, surveys with the large area and near-infrared bands necessary to detect large numbers of rare, high-redshift galaxies will continue to be conducted by ground-based facilities at significantly lower spatial resolution, and so these challenges to aperture-based methods will only become more difficult, e.g. 40% of Rubin/LSST sources will be blended at above a 5% level (see Fig. 19 of Faisst et al. 2021). As we will also demonstrate, **the success of aperture photometry becomes more limited with deeper surveys of crowded galaxy fields.**

These challenges must be met with appropriate solutions now if we are to continue exploring not only the high-redshift universe, but **pursuing** any study whose success relies upon contending with crowded fields and faint sources (e.g., cosmology, transients). Successfully approaches will necessarily be robust to contamination from neighboring sources, provide reliable limits on non-detections, and be consistently applicable over a wide range of spatial resolution, wavelength, seeing, and sensitivity.

An attractive class of alternative photometric techniques called “profile-fitting” photometry has enjoyed great success overcoming these very challenges. They work by fitting a model (parametric or non-parametric) that describes and can be reliably fit to the surface brightness profile of a source. Usually the total brightness is a parameter of the model, or can somehow be derived from it. Commonly used parametric implementations of profile-fitting involve a source model parameterized by flux, position, and for resolved sources also size, axis ratio, position angle, and light profile (i.e. Sérsic index; Sérsic 1963) which is then convolved with a known PSF and fit to the surface brightness profile of a given source. This approach has significant advantages over traditional apertures. Firstly, the flux reported is the total brightness of the source in that particular band, avoiding aperture corrections and related

systematics. Secondly, the PSF is a property of the model, which is a more tractable solution compared to PSF homogenisation which manipulates the measurement image. This means that the fitted properties of resolved sources are the intrinsic, PSF-deconvolved values. Thirdly, positions are not simply determined as the peak or centroid of an image but are rather fitted parameters, subsequently achieving greater precision over commonly-used peak-finding routines in photometry software (e.g., **Source Extractor**; Bertin & Arnouts 1996). Lastly, sources that have some fraction of their flux overlapping can be accurately photometered by fitting an appropriate number of simultaneous models. This forward-modeling ability to de-blend sources is unique to profile-fitting photometry and means that sources easily differentiated in high-resolution images can be accurately photometered in low-resolution bands such as *Spitzer*/IRAC.

The **Tractor**<sup>1</sup> (Lang et al. 2016) is one such profile-fitting tool. Given a set of initial positions, model profile files (e.g. point-like versus resolved), and image information with per-pixel uncertainties, The **Tractor** optimizes those models for a given set of images whose sources have been already identified from some existing (ideally higher resolution) detection image. The key distinction when utilizing such parametric models is that we can derive a likelihood for the particular model parametrization given the data, as well as estimate uncertainties on those parameters. Key implementations of The **Tractor** include Lang et al. (2016a), Faisst et al. (2021), and Stevens et al. (2021). In addition, Nyland et al. (2017) explored for the first time the capabilities of The **Tractor** to photometer highly blended IRAC sources using models derived from higher resolution VISTA imaging.

We develop a pipeline to perform reproducible profile-fitting photometry built around The **Tractor** called The **Farmer**, which adopts similar principles used in previous work concerning model-based photometry including HSCPipe (Aihara et al. 2019), the DECaLS pipeline (Dey et al. 2019), GaLight (Ding et al. 2021), and SExtractor++, (Bertin et al. 2020; Kümmel et al. 2020). The **Farmer** provides a larger framework within which The **Tractor** can be scaled to large galaxy surveys where source detection must be handled in a statistical manner. Crucially, The **Farmer** includes built-in parallelization methods which enable efficient computational runtimes. The **Farmer** utilizes the optimization routines already provided by The **Tractor** to obtain es-

timates of source flux and positions, as well as galaxy shapes for resolved sources. At no point are fluxes derived through integration over an aperture. Instead, the fluxes are derived directly from the normalization factor required to scale a unit-normalized model to best describe a given source. Parameter uncertainties, including flux, are derived as minimum-variance estimates according to the Cramér-Rao bound (Cramer 1946; Rao 1945). For point-like sources, this equates to the classical variance derived when fitting a pattern using inverse-variance weights.

We begin in Section 2 with a review of the key aspects of The **Tractor**. Section 3 then describes the purpose and design of The **Farmer**. Section 4 presents the results of benchmarking The **Farmer** against a set of simulated COSMOS-like images before concluding in Section 5. Other considerations and discussion is included in Appendix A.

The features, behaviour, and performance of The **Farmer** described in this paper is purposefully consistent with its use in Weaver et al. (2022) so that it can be used as a supporting reference. The software is available on GitHub<sup>2</sup> and is provided ‘as is’. The authors reserve the right to update the software – and its features and performance – at any time, documenting relevant changes. The material presented here is independent of any assumed cosmology. All magnitudes are expressed in the AB system (Oke 1974), for which a flux  $f_\nu$  in  $\mu\text{Jy}$  ( $10^{-29} \text{ erg cm}^{-2}\text{s}^{-1}\text{Hz}^{-1}$ ) corresponds to  $\text{AB}_\nu = 23.9 - 2.5 \log_{10}(f_\nu/\mu\text{Jy})$ .

## 2. REVIEW OF THE TRACTOR METHODOLOGY

The **Tractor** is a recent development aimed at providing a generalized framework for fitting the surface brightness profiles of sources in an image. The approach is generative, that is, The **Tractor** attempts to construct a predictive model based on the science image, a corresponding PSF, and a per-pixel noise estimate (typically a weight map), and optionally a background sky model; as well as initial guesses as to the model parameters such as source positions, shapes, and fluxes. In practice, The **Tractor** optimizes these initial parameters to produce a model image which describes input image within the bounds of the properties provided, separating the source signal from the background noise.

The flux of a given source  $\alpha$  is not measured with apertures, but is rather obtained directly as the normalization of a unit-normalized model profile  $G(\phi)_i$ , where  $\phi$  is the subset of parameters describing the position and

<sup>1</sup> <https://github.com/dstndstn/tractor>

<sup>2</sup> [https://github.com/astroweaver/the\\_farmer](https://github.com/astroweaver/the_farmer)

<sup>243</sup> shape of the overall model  $m(\theta)_i$  defined over every pixel  
<sup>244</sup>  $i$  and convolved with the PSF:

$$m(\theta)_i = \alpha G(\phi)_i \otimes \text{PSF}_i \quad (1)$$

<sup>245</sup> The flux for a single isolated point source is essentially  
<sup>246</sup> computed as a mean of the input image  $x_i$  and the model  
<sup>247</sup> image  $m(\theta)_i$  normalized to unity and inversely weighted  
<sup>248</sup> by pixel variance  $\sigma_i^2$ . In other words, flux is the value re-  
<sup>249</sup> quired to scale a unit-normalized model image of a point  
<sup>250</sup> source to describe the real point source. **The Tractor**  
<sup>251</sup> attempts to maximize the likelihood  $\mathcal{L}(x|\theta)$  of the data  
<sup>252</sup>  $x$  given the free parameters  $\theta$ , and uses the quadrature  
<sup>253</sup> addition of the weighted residual image (i.e.  $\chi$ ), which  
<sup>254</sup> is analogous to a  $\chi^2$  minimization as  $\mathcal{L} \sim \exp(-\frac{1}{2}\chi^2)$   
<sup>255</sup> but in two spatial dimensions, ignoring pixel-pixel co-  
<sup>256</sup> variances:

$$\mathcal{L}(x|\theta) = \exp \left( - \sum_i \frac{(m(\theta)_i - x_i)^2}{2\sigma_i^2} \right) \quad (2)$$

<sup>257</sup> One immediate advantage of this approach is that it  
<sup>258</sup> avoids the need for PSF homogenization as the PSF is  
<sup>259</sup> included in convolution with the source profile. Another  
<sup>260</sup> advantage is that as long as the model is normalized to  
<sup>261</sup> unity including the wings, it may be truncated in nu-  
<sup>262</sup> merical processing without biasing the estimated flux.  
<sup>263</sup> Therefore while an aperture over the model realized in  
<sup>264</sup> some restricted image dimensions will return a flux less  
<sup>265</sup> than the true flux, the flux determined by scaling the  
<sup>266</sup> unit normalized (but truncated) model will remain ac-  
<sup>267</sup> curate. This is especially useful when considering nu-  
<sup>268</sup> mercial and computational limitations.

Of perhaps equal importance are parameter un-  
<sup>269</sup> certainties. The uncertainty estimates produced by  
**The Tractor**, reproduced here based on documentation  
<sup>270</sup> provided with the code, are related to the Cramér–Rao  
<sup>271</sup> bound which is a lower bound on the variance of any  
<sup>272</sup> unbiased estimator  $\hat{\theta}$ :

$$\text{var}(\hat{\theta}) \geq \frac{1}{I(\theta)} \quad (3)$$

where  $I(\theta)$  is the Fisher Information,

$$I(\theta) = E_x \left[ \left( \frac{\partial \log \mathcal{L}(x|\theta)}{\partial \theta} \right)^2 \right] \quad (4)$$

$$= -E_x \left[ \frac{\partial^2 \log \mathcal{L}(x|\theta)}{\partial \theta^2} \right]. \quad (5)$$

The log likelihood is therefore

$$\log \mathcal{L}(x|\theta) = - \sum_i \frac{(m(\theta)_i - x_i)^2}{2\sigma_i^2} \quad (6)$$

with first derivative

$$\frac{\partial}{\partial \theta} \log \mathcal{L}(x|\theta) = - \sum_i \frac{m(\theta)_i - x_i}{\sigma_i} \frac{1}{\sigma_i} \frac{\partial}{\partial \theta} m(\theta)_i , \quad (7)$$

<sup>269</sup> which should equal zero when the likelihood has been  
<sup>270</sup> maximized.

<sup>271</sup> The second derivative is

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} \log \mathcal{L}(x|\theta) &= \\ -\sum_i \left[ \frac{m(\theta)_i - x_i}{\sigma_i} \frac{1}{\sigma_i} \frac{\partial^2}{\partial \theta^2} m(\theta)_i + \left( \frac{1}{\sigma_i} \frac{\partial}{\partial \theta} m(\theta)_i \right)^2 \right] & , \end{aligned} \quad (8)$$

<sup>272</sup> where the first term is zero at the optimum. Returning  
<sup>273</sup> to the Cramér–Rao bound, we have

$$\frac{1}{\text{var}(\hat{\theta})} \leq -E_x \left[ \frac{\partial^2 \log \mathcal{L}(x|\theta)}{\partial \theta^2} \right] \quad (9)$$

<sup>274</sup> and since our second derivative (equation 9) is indepen-  
<sup>275</sup> dent of  $x$ , the expectation collapses and we get

$$\frac{1}{\text{var}(\hat{\theta})} \leq \sum_i \left( \frac{1}{\sigma_i} \frac{\partial}{\partial \theta} m(\theta)_i \right)^2 , \quad (10)$$

<sup>276</sup> which is the inverse-variance estimate reported by  
**The Tractor**. In the important case of estimating flux  
<sup>277</sup> where  $\theta \equiv \alpha$ , the derivative of the model with respect  
<sup>278</sup> to flux is just the profile of the model. Hence, the un-  
<sup>279</sup> certainty estimate on flux for point-like sources is based  
<sup>280</sup> entirely upon the PSF and the per-pixel error estimates  
<sup>281</sup>  $\sigma_i$  from the weight map.

<sup>282</sup> We can gain a better understanding of **The Tractor**,  
<sup>283</sup> both its functionality and limitations, through progres-  
<sup>284</sup> sively complex examples.

<sup>285</sup> The simplest example is an isolated, point-like galaxy.  
<sup>286</sup> **The Tractor** is supplied with the image, a weight  
<sup>287</sup> map, a PSF, and a known position for the source;  
<sup>288</sup> **The Tractor** does not provide means to detect sources,  
<sup>289</sup> and so a list of initial source positions is required before-  
<sup>290</sup> hand. While the data input (image, weight map, PSF,  
<sup>291</sup> optionally sky) must be kept fixed, we may also fix the  
<sup>292</sup> position parameter so that only the flux is allowed to  
<sup>293</sup> vary. This one parameter optimization is linear in the  
<sup>294</sup> case of a single source. However, profile-fitting photom-  
<sup>295</sup> etry is sensitive to offsets in source positions requiring  
<sup>296</sup> greater precision than is typically needed for accurate  
<sup>297</sup> aperture photometry. One can address this by simply al-  
<sup>298</sup> lowing the model position to also vary, and **The Tractor**

300 has built-in functionality to deal with this. This three  
 301 parameter (i.e. x, y, flux) optimization is a non-linear  
 302 procedure, although the degeneracy between the position  
 303 and flux parameters should be virtually zero. The  
 304 result is not only an estimate of the flux, but also the  
 305 source position. The source may also be photometered  
 306 in many bands in a single joint optimization where the  
 307 shape and position are shared but flux is now a vector  
 308 with an element for each band.

309 A more complicated example is an isolated, resolved  
 310 source. **The Tractor** includes a library of discrete para-  
 311 metric models which include but are not limited to, in  
 312 order of simplicity, point source profiles taken from the  
 313 PSF stamp (as assumed in the previous example), re-  
 314 solved models with exponential or de Vaucouleurs pro-  
 315 files (de Vaucouleurs 1948), full Sérsic profiles, and com-  
 316 posite profiles made by superimposing exponential and  
 317 de Vaucouleurs profiles. As before, the properties of the  
 318 input data (i.e. image pixel values) are kept fixed. We  
 319 also may fix the position, for simplicity, leaving the  
 320 source shape and flux free to vary. The question then  
 321 is how to decide which shape parameterization to use?  
 322 **The Tractor** does not provide an answer; rather it is  
 323 up to the user to **determine** a model type ahead of the  
 324 optimization. A resolved model type is appropriate in  
 325 this case, and so now our optimization returns source  
 326 fluxes and shapes (e.g., effective radius, axis ratio, and  
 327 position angle). Photometry of other images taken with  
 328 different filters is usually of interest and so by fixing the  
 329 model shape we can perform “forced photometry”. Al-  
 330 though it is possible to allow the shape to vary with each  
 331 band, this comes at the cost of potentially overfitting our  
 332 model.

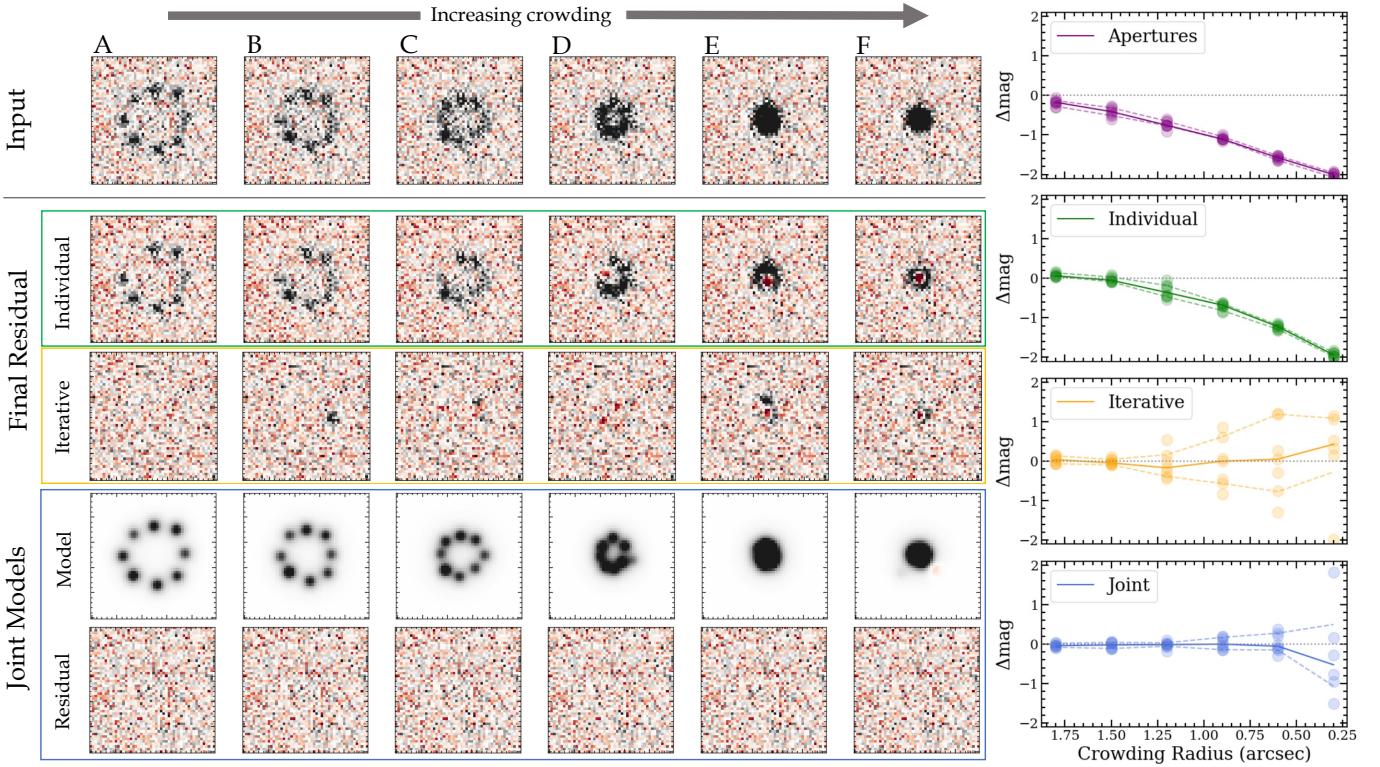
333 An even more complicated example is an image con-  
 334 taining many sources of various morphological presen-  
 335 tations and crowding. This is typically what is encoun-  
 336 tered in deep galaxy surveys and presents a serious chal-  
 337 lenge. We have already understood that **The Tractor**  
 338 does not provide source detection, and so the degree  
 339 to which the photometry succeeds is dependent on  
 340 the performance of some external detection procedure.  
 341 Once we have somehow supplied source centroids to  
 342 **The Tractor**, we are still left to determine the appro-  
 343 priate model type for each source. Although it may be  
 344 feasible to assign model types manually for small regions  
 345 of interest occupied by a small number of sources, this is  
 346 typically not practical for large surveys containing thou-  
 347 sands or even millions of sources. Assuming this can be  
 348 done in some way, **The Tractor** will optimize all source  
 349 models simultaneously on that given image to produce  
 350 optimized shapes which can then be fixed to performed  
 351 forced photometry on other bands of interest. Alterna-

352 tively, one can use all the bands of interest to optimize  
 353 the model and simultaneously obtain measurements of  
 354 fluxes, although this adds significant complexity that  
 355 may cause the optimization to not converge.

356 The situation does not improve much even if there is  
 357 only one source of interest amongst a crowded field. Al-  
 358 though one may try to instantiate a single model at that  
 359 source position, **The Tractor** uses information from ev-  
 360 ery pixel in the image that has non-zero weight. That  
 361 means that the presence of every other source in the  
 362 image counts against the likelihood. One option is to  
 363 restrict the weight map to only the pixels belonging to  
 364 that source. However, deciding the extent of such a  
 365 region is non-trivial. Regions that are too large may  
 366 include flux from a neighbor which are unaccounted for  
 367 by our one source model, and may bias the photome-  
 368 try typically towards higher fluxes. Having too small an  
 369 region is suboptimal, and ill-defined as you would need  
 370 to know the extent any neighbors beforehand. Another  
 371 option is to continue instantiating models (defined by  
 372 centroids and model types) for all nearby sources until  
 373 it is possible to cleanly define a contiguous region whose  
 374 boundaries do not contain light from other sources (i.e.,  
 375 an isolated group of sources). Such a manual approach  
 376 may work, but only in limited cases where the user is  
 377 heavily involved, severely limiting reproducibility. Even  
 378 if this can be done, it is **not immediately obvious**  
 379 how best to fit this potentially large group of nearby  
 380 sources. Should they be fit simultaneously? This ap-  
 381 proach is straight-forward but computationally expen-  
 382 sive. Perhaps they should be fit one by one, subtracting  
 383 the best-fit model each time? This is usually compu-  
 384 tationally faster, but induces hysteresis that can bias  
 385 photometry.

386 A generalized version of this dilemma is useful in prov-  
 387 ing this point. In Figure 1, eight point sources are in-  
 388 jected into a Gaussian noise field at signal strengths  
 389 ranging from  $\sim 3 - 10\sigma$  and arranged in a circle. A  
 390 total of six cases are constructed (A, B, C, D, E, F) by  
 391 varying the radial distance to each source such that at  
 392 one extreme they are separated (A) and overlapping at  
 393 the other (F).

394 As a baseline, fluxes are summed in 2'' apertures that  
 395 do not overlap in case A and so recover accurate fluxes.  
 396 However, a bias grows towards case F where the aper-  
 397 tures become confused and eventually include the flux  
 398 from all eight sources in each aperture. This highlights  
 399 the limitations of apertures in pathologically crowded  
 400 fields, after which one must appeal to statistical miti-  
 401 gation strategies afterwards to re-scale fluxes. We move  
 402 on to profile-fitting photometry in the subsequent rows.  
 403 The most direct approach is to model each galaxy indi-



**Figure 1.** Eight similarly bright point sources are injected into a simulated noise field over six scenarios of increasing degrees of circularly symmetric crowding. Assuming source positions are known beforehand, fluxes are measured in four ways: 2'' apertures, profile-fitting each source independently, profile-fitting each source with successive model subtraction, and jointly fitting all models simultaneously. The degree of success of each method is shown on the right measured in the difference in magnitude  $\Delta\text{mag}$  between the input and measured magnitudes as a function of source crowding, with a median  $\Delta\text{mag}$  and 68% ranges indicated for each scenario. Only joint fitting provides both precise and accurate recovery of crowded sources.

ividually in series (allowing the uncertain positions to necessarily vary in the fit), but by case C succumbs to the same confusion as the apertures and multiply counts each source per model. An attractive solution is to also iteratively subtract each model one by one, in series. **While this approach is certainly more successful in that the average flux is accurate, that of most individual sources is catastrophically wrong. This poor performance is also evidenced by significant residual flux in all but the least crowded case.**

The optimal way is to model each source simultaneously, which allows the *joint* model to recognize that there are neighbors that it can describe. This approach does not suffer from the drawbacks of fixed apertures, or of fitting models individually or with subtraction. **It recovers unbiased photometry in cases A, B, C and D, failing in only the most crowded cases (E and F).** Yet, this level of crowding is pathological as it is unlikely that (in the absence of higher resolution data) a source detection procedure would be able to separate the signal into even two centroids, let alone all eight. Therefore the most extreme cases remain a

problem, but one which will have to be addressed by innovations in source detection and associated de-blending techniques. Although fitting multiple nearby sources simultaneously is clearly the optimal approach, it is also the most computationally expensive one, and for that reason it cannot be so readily scaled up to large area surveys without first developing efficient algorithms that can be utilized successfully by high performance computing facilities.

As we can see, The Tractor is a powerful tool for determining best-fit values corresponding to parametric models of sources, but it requires significant manual attention in all but the simplest cases. Therefore there is a considerable gap between the function of The Tractor and what is required for front-to-back catalog pipelines. Developing such pipelines is not only time consuming, but independently developed pipelines perform differently (e.g., that of Nyland et al. 2017 is different than the pipeline of Dey et al. 2019). While each implementation may be optimized for a certain task, the overwhelming success of software like Source Extractor is that they are immediately accessible, flexible, and easy to use. However, the matters of source detection, model type

450 decisions, which groups of sources to model and how  
 451 best to model them, as well as computational efficiency  
 452 are challenges that must be solved if we are to construct  
 453 such a generalized pipeline that applies **The Tractor** to  
 454 the incredibly deep, crowded fields to be explored by the  
 455 next generation of galaxy surveys.

### 456 3. THE FARMER: A GENERAL DESCRIPTION

457 **The Farmer** is a generalized, flexible, and repro-  
 458ducible framework that uses the model library from  
 459 **The Tractor**, its optimization engine, and several  
 460 helper routines to photometer detected sources, mea-  
 461sure their shapes, produce catalogs and ancillary images,  
 462 as well as provide supporting diagnostics. **The Farmer**  
 463 overcomes the issue of how to assign model types by  
 464 identifying natural groups of nearby sources and deter-  
 465mines the best model type of each source using a decision  
 466 tree in a time efficient, optimal way whilst mitigating  
 467 related pathological situations. It includes a significant  
 468 organizational capacity such that images can be divided  
 469 up into sections for massively parallelized computation.  
 470 Here we walk though the process of **The Farmer** from  
 471 image preparation to the output catalogs.

#### 472 3.1. *Image Preparation*

473 At bare minimum, **The Farmer** requires a single sci-  
 474ence image containing sources of interest. A correspond-  
 475ing inverse variance weight map is ideal, but not re-  
 476quired. Lacking weight information, **The Farmer** can  
 477utilize the Pythonic **Source Extractor** code **SEP** by  
 478 [Barbary \(2016\)](#) to measure noise directly from the im-  
 479ages or simply assume equal weights.

480 In this basic case, **The Farmer** will detect sources,  
 481model them, and perform forced photometry all on the  
 482same monochromatic image. In more typical, complex  
 483cases it is desirable to produce a separate detection im-  
 484age that combines multiple bands. For surveys of faint  
 485sources, the **CHI-MEAN** approach ([Szalay et al. 1999](#);  
 486[Bertin 2010](#)) has been widely adopted (e.g. [Laigle et al.](#)  
 4872016; [Weaver et al. 2022](#)), or a similar signal-to-noise  
 488image co-add (e.g. [Whitaker et al. 2011](#)).

489 Masking is especially important in profile-fitting pho-  
 490tometry for the reason that it is inadvisable to attempt  
 491to model large, saturated stars, nebulae, or nearby  
 492galaxies which are essentially nuisance foreground con-  
 493tamination. While apertures have the advantage of be-  
 494ing able to efficiently sum fluxes in whatever regions of  
 495an image are of interest, models must attempt to de-  
 496scribe the image as it is. Attempting to model such  
 497nuisance sources, which lie outside the reach of our para-  
 498metric models, will never achieve a satisfactory fit even  
 499after several hundred central processing unit (CPU)

500 hours, if at all. That being said, we note without ex-  
 501tensive sky background modelling, sources within bright  
 502star halos (e.g.) will not be photometered accurately  
 503with apertures either.

504 A useful recipe is to stack all bands that will used to  
 505detect sources, and mask out the full extent of such nui-  
 506sance foreground objects, and possibly also the edges of  
 507the mosaic or detector. **The Farmer** can be configured  
 508to apply a mask before or after source detection. The  
 509latter is preferred in virtually all cases, as mask edges  
 510can produce spurious sources. Applying a mask after  
 511source detection simply removes sources from the cata-  
 512log and their corresponding segments are zeroed out.

513 **The Farmer** includes several ways to measure image  
 514backgrounds and per-pixel noise based on **SEP**, and this  
 515can be configured by the user. Backgrounds can be  
 516measured as global medians or spatially varying (fol-  
 517lowing the methods of **Source Extractor**; see [Bertin](#)  
 518& [Arnouts 1996](#)), with per-pixel noise being estimated  
 519directly from the RMS of the image. The background  
 520and per-pixel noise estimates can be produced with and  
 521without the mask in order to mitigate the adverse ef-  
 522fects of bright stars and foreground galaxies. Although  
 523currently all detected sources will be modelled, the abil-  
 524ity to identify and remove spurious sources on-the-fly is  
 525expected to be included in a future update.

#### 526 3.2. *PSF creation*

527 With the images and weights in hand, **The Farmer**  
 528needs a PSF for each band of interest. There are many  
 529way of generating PSF stamps, including as realizations  
 530of spatially varying models, and **The Farmer** can be sup-  
 531plied with several PSF types.

532 The most common is a constant PSF stamp sampled  
 533at the same pixel scale as the its corresponding im-  
 534age; these can be readily produced by packages such as  
 535**PSFEx** ([Bertin 2013](#)). One may also use **PSFEx** to gener-  
 536ate spatially varying PSF models, all flavours of which  
 537(e.g., Gauss-Laguerre or pixel bases) are understood by  
 538**The Farmer** (and importantly also by **The Tractor**).  
 539 While this can be achieved through using **PSFEx** by itself,  
 540**The Farmer** is able to run **PSFEx** in a semi-automatic  
 541way using built-in functions. First **The Farmer** runs  
 542**Source Extractor** to identify bright sources and pro-  
 543duce a catalog including vignette stamps that **PSFEx** can  
 544read in (**SEP** do not produce such output files). Candi-  
 545date point sources are then selected either automatically  
 546by **PSFEx**, or more directly by a pre-selection by the user  
 547based on source FWHM and brightness. The user can  
 548also declare which bands should use a constant PSF and  
 549which should be spatially varying, and **The Farmer** will  
 550automatically reconfigure **PSFEx** in each case.

In some cases the PSF varies too quickly across an image to be accurately characterized by a smoothly varying surface as used by PSFEx. It is possible therefore for the user to supply a set of PSFs and a file which maps each one to a coordinate so that **The Farmer** can use the nearest sampled PSF for a given source. The assumption of a smoothly varying PSF can thereby be avoided, and the user is free to choose the grid geometry according to their requirements. This ‘PSF Grid’ approached was developed in Weaver et al. (2022) to characterize the photometry of the Subaru Suprime-Cam mosaics in COSMOS.

The images of *Spitzer*’s Infrared Array Camera (IRAC) feature a highly variable PSF which is generally triangular in shape. The PRFMap package<sup>3</sup>) attempts to characterize this highly irregular PSF by mapping the pixel of each stacked image back to the locations on the CCD of the constituent images. It then uses the spatially-dependent calibration PSFs to construct a combined PSF for the stacked image. Similar to the PSF Grid approach, PRFMap produces a library of individual PSFs corresponding to a fixed grid of sampling coordinates. This output can be used with **The Farmer** to measure IRAC photometry.

One important caveat to note is that in all cases the PSF must be measured into its wings and not be truncated. This is for two reasons. Firstly, profile-fitting models generally benefit from the wings of the PSF being in tact. This can be immediately appreciated in the case of unresolved sources fit with point-source models for which **The Tractor** uses the PSF stamp for the model profile: if the wings of the point-source model do not describe the full spatial profile of the source of interest (i.e. all pixels within the source segment that contribute to its  $\chi^2$ ) then the residual will always have leftover signal in the wings and the measured flux may be biased. Secondly, the pixel values of a PSF which has been truncated and then normalized to unity will be larger than those of the full PSF normalized to unity, and so its optimal normalization coefficient (i.e. flux) will be smaller for the same source, introducing a bias. Therefore it is strongly advised to sample the entire PSF profile out to radii where the wings are indistinguishable from noise, in most cases corresponding to a radius of several arcseconds.

### 3.3. Source Detection

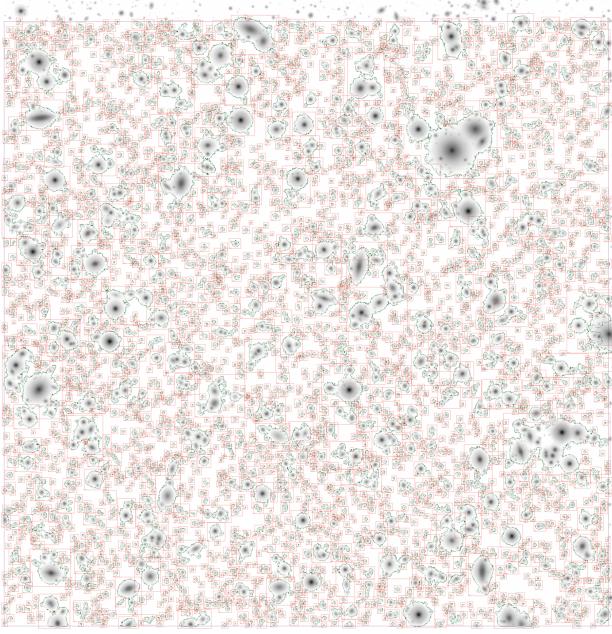
The first step in catalog creation is source detection. **The Farmer** utilizes SEP (Barbary 2016) to pro-

vide source detection, segmentation maps, background, and noise estimation with near identical performance as classical Source Extractor. As with any detection software, the performance of SEP as measured by source de-blending and segmentation, e.g., is entirely dependent on the configuration of the detection parameters set by the user (see Bertin & Arnouts 1996; Szalay et al. 1999; Holwerda 2005). Segmentation of blended sources in both SEP and SourceExtractor relies on multiple thresholds to determine which pixels belong to which galaxy (see Section 2.3.1 of Haigh et al. 2021 for details). Generally, detection strategies vary between catalogs, are typically driven by science objectives, and are often tuned by eye. The performance of **The Farmer** described in this work is no different; for the sake of comparison to COSMOS2020 we adopt the same detection configuration as described in Weaver et al. (2022).

We stress that although SEP is responsible for identifying individual galaxies, the deblending of their light is entirely determined by **The Farmer**. However, if SEP incorrectly groups separate galaxies together, **The Farmer** cannot deblend them afterwards (see Section A.4). Detection parameters for SEP can be configured directly with **The Farmer**, and related diagnostic images are supplied indicating source centroids on the detection image. It is also possible to hand **The Farmer** a catalog of source coordinates and a corresponding segmentation map from e.g., Source Extractor, or any other similar detection software.

**The Farmer** performs all functions on discrete sections of the total mosaic called “bricks” (following Dey et al. 2019). An example is shown in Figure 2. Each brick is cut out of the total mosaic image, weight, and mask with equal dimensions, and includes a overlap region on each side. Sources detected with centroids in the overlap region are removed from the source catalog of the brick, and the pixels belonging to their segment are set to zero (like background pixels). They are not lost, however, as they are found again in the main region of a neighboring brick. This “fuzzy boundary” approach means **The Farmer** can construct unique source catalogs for each brick which have no overlap with neighboring bricks, thus accounting for every source without duplication or loss. Although the overlap regions of the segment map are also set to zero, **The Farmer** keeps segment pixels in the overlap region of sources whose centroids are in the main region of the brick. This behavior allows sources which are near the overlap zone to be modelled with all of their pixels, as opposed to a strict cut-off

<sup>3</sup> <https://github.com/cosmic-dawn/prfmap>.



**Figure 2.** Example of groups detected over a brick. The brick lies at the edge of a mosaic and so has boundaries with two adjacent bricks. Source centroids are shown by green crosses. Groups are outlined by red boxes and group footprints of owned pixels by the red borders.

at the overlap boundary where their profiles would be truncated.

Following the creation of the brick’s preliminary source catalog provided by SEP and a cleaning of the overlap regions, The Farmer attempts to identify natural groups of detected sources which would benefit from being simultaneously modelled. Groups are identified by dilating the original segments to form contiguous non-zero regions. Sources which are not in crowded areas form singularly occupied groups, whereas sources in crowded regions end up members of larger groups to be modelled simultaneously. See Section A.1 for further discussions.

### 3.4. Model Type and Shape Determination

A model must now be determined for each source in a given group. The goal is to not only determine the most suitable model for each source, but also its best-fit parameters. While the number of possible decision tree architectures is virtually infinite, The Farmer relies on a balanced architecture consisting of five discrete models to describe resolved and unresolved, stellar and extragalactic sources:

1. **PointSource** models are taken directly from the PSF used. They are parameterized by flux and centroid position and are appropriate for unresolved sources.

2. **SimpleGalaxy**<sup>4</sup> models use an exponential light profile with a fixed user-defined effective radius such that they describe marginally resolved sources and mediate the choice between PointSource and a resolved galaxy model. They are parameterized also by flux and centroid position.

3. **ExpGalaxy** models use an exponential light profile. They are parameterized by flux, centroid position, effective radius, axis ratio, and position angle.

4. **DevGalaxy** models use a de Vaucouleurs light profile. They are parameterized by flux, centroid position, effective radius, axis ratio, and position angle.

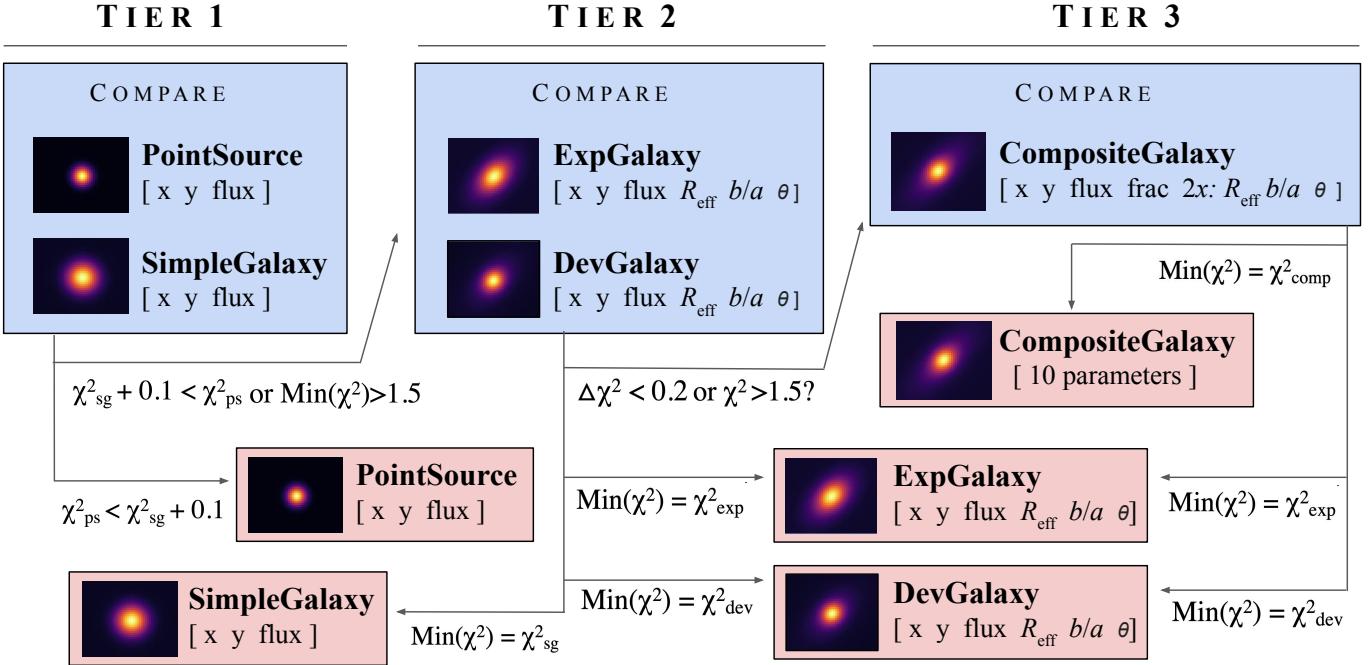
5. **CompositeGalaxy** models use a combination of ExpGalaxy and DevGalaxy models. They are concentric, and hence share one centroid. There is a total flux parameter as well as a parameter for the fraction of total flux assigned to the DevGalaxy component<sup>5</sup>. Each component has their own effective radius, axis ratio, and position angle.

In practice, these spatially-resolved models are optimized using sigmoid-softened ellipticities provided by The Tractor (i.e. the `EllipseESoft` class), which allows for an unbounded parameter space more suitable for numerical computation. **Also note that although ExpGalaxy and DevGalaxy models can be generalized under a single Sérsic model with variable Sérsic index, we purposely forgo this additional free parameter as it is generally under-constrained by the relatively low resolution imaging of COSMOS2020.**

These five models form The Farmer’s decision tree, whose goal is to both determine the most suitable model for a given source, and provide an optimized set of parameters to describe the shape and position of that source. To ensure that crowded regions do not suffer from poor modelling as a result of the model of a particular source being constrained by light from neighboring source, the models are determined simultaneously at each stage of the decision tree. The values for the decision tree parameters quoted here are examples taken from Weaver et al. (2022) but can and should be tuned by the user for other data sets. An example of a group

<sup>4</sup> SimpleGalaxy models are not a standard The Tractor model, see <https://github.com/legacysurvey/legacypipe>.

<sup>5</sup> CompositeGalaxy models assume the `FixedCompositeGalaxy` model class in The Tractor.



**Figure 3.** Schematic of the three-tiered decision tree used by **The Farmer** to determine the most suitable model type for a given source. The five models are tested in order of increasing number of parameters. Values shown are examples, and should be configured by the user to suit their dataset.

722 containing two sources progressing through the decision  
 723 tree is shown in Figure 4.

724 **The Tractor** uses a likelihood cost function to score  
 725 the performance of the joint model containing all of the  
 726 individual models of the sources in a group. All weight  
 727 pixels outside the group footprint are set to zero (**i.e.**  
 728 **no weight**) such that nearby sources which are not part  
 729 of the group cannot influence the likelihood. However,  
 730 we also need to be able to assess the performance of  
 731 an individual model for a given source in our group.  
 732 **The Farmer** adopts  $\chi^2_N$  as its goodness of fit statistic,  
 733 which is calculated by quadrature addition of the residual  
 734 image pixels belonging to a particular source by its  
 735 original segment and then reduced by dividing by the  
 736 number of **degrees of freedom**  $N$ , taken as the differ-  
 737 ence between the number of pixels in its segment and  
 738 the number of free parameters.

739 **The Farmer** begins by considering PointSource mod-  
 740 els for every source in a group, using centroids and fluxes  
 741 estimated by **SEP** as initial conditions. **The Tractor**  
 742 then performs an optimization to maximize the com-  
 743 bined likelihood of the entire joint model, after which  
 744 **The Farmer** computes the  $\chi^2_N$  for each source in the  
 745 group. Next, SimpleGalaxy models are considered for  
 746 all sources in a group with the same initial conditions  
 747 as before. The models are optimized and the  $\chi^2_N$  per  
 748 source is computed considering pixels within each seg-  
 749 ment. **The Farmer** then tries to place each source into

750 one of three categories: either the source is well fit by  
 751 the PointSource and is fixed as a PointSource, it is fit  
 752 well by a SimpleGalaxy, or neither model is appropriate.  
 753 Satisfying either of the last two categories advances the  
 754 source down the decision tree towards more complex,  
 755 resolved model types. The role of the SimpleGalaxy  
 756 here is not to be a commonly chosen model, but rather  
 757 a fast-to-compute indicator of a resolved source. Unlike  
 758 comparing a PointSource model to a more complex Exp-  
 759 Galaxy model, the comparison with the SimpleGalaxy is  
 760 not only computationally faster but is statistically fair  
 761 since the number of parameters for both PointSource  
 762 and SimpleGalaxy are the same, as are the number of  
 763 data points. Sources which are best fit with PointSource  
 764 models will be assigned a PointSource model thereafter,  
 765 which in the case of a one source “group” will con-  
 766 clude the decision tree. A source that is better fit by  
 767 a SimpleGalaxy model by only a slim margin is typ-  
 768 ically sufficiently modelled by a PointSource also. It  
 769 is desirable therefore to prefer a PointSource in these  
 770 cases as a better fit. However, a source well fit by a  
 771 SimpleGalaxy model triggers the more complex tiers of  
 772 the decision tree, meaning that the overall group model  
 773 becomes more complex which requires even longer com-  
 774 putational times. **The Farmer** therefore penalizes the  
 775 SimpleGalaxy models in  $\chi^2_N$  by 0.1 such that a Sim-  
 776 pleGalaxy model must have a lower  $\chi^2_N$  by a margin  
 777 of 0.1 or better in order to not choose a PointSource

(these values again being examples suitable for the COSMOS2020 catalog as empirically determined in Weaver et al. 2022). A PointSource will also not be chosen (at this stage) if produces a bad fit, assessed by  $\chi_N^2 > 1.5$ , whereupon the source continues to the next level of the decision tree. However, a PointSource or SimpleGalaxy may still be chosen in the end, but only if it is still favored after the assessment of more complex models.

The next stage of the decision tree determines the general Sérsic light profile of resolved sources whose model types remain unfixed, choosing between ExpGalaxy or DevGalaxy. At this stage, fixed sources can only have been assigned PointSource models. The Farmer starts by considering ExpGalaxy models for all other unfixed sources, performs the optimization, and determines  $\chi_N^2$  for each. Initial guesses for shape parameters are initialized borrowing from SEP measurements (e.g.  $a$ ,  $b$ ,  $\theta$ ) estimated at detection. Then The Farmer performs the same computation but with DevGalaxy models on all unfixed sources. Again the  $\chi_N^2$  is a fair comparison as the number of degrees of freedom are identical between the two model types. The Farmer allows the model parameters to remain variable for all sources, regardless of whether they have been assigned a final model type, at each stage of the decision tree (e.g., fixed PointSource models still re-optimize their flux). Sources whose ExpGalaxy and DevGalaxy models both fail to achieve a lower  $\chi^2$  than the SimpleGalaxy are fixed as SimpleGalaxy models, unless the SimpleGalaxy also fails to achieve a  $\chi_N^2$  of 1.5 in which case that source advances down the decision tree to the third tier. The choice between ExpGalaxy and DevGalaxy models is determined by the lowest  $\chi_N^2$ , without any penalties. However, if the absolute difference in  $\chi_N^2$  between the two models is less than 0.2, or neither ExpGalaxy or DevGalaxy achieves a  $\chi_N^2$  of 1.5, the source also advances to the third tier.

All sources have typically been assigned a fixed model by this stage, especially those that have smooth light profiles or are unresolved, and the decision tree ends without trying more complex, time intensive models that The Farmer has already determined are not required for a sufficient fit. However, highly spatially resolved sources that have reached the third tier without an assigned model are fit assuming the most complex CompositeGalaxy models. If the CompositeGalaxy model fails to achieve a better  $\chi_N^2$  than either ExpGalaxy or DevGalaxy, the source is assigned the model type that achieved the lowest  $\chi_N^2$  overall.

Now that models for all sources belonging to a given group are assigned, The Farmer optimizes the models

a final time. This is an important step as it is possible for an otherwise pathological case to arise whereby two assigned models were never optimized at the same time and their fits may influence each other. By computing this final optimization, the overall likelihood of the model set for the group of sources tends to improve.

### 3.4.1. Forced Photometry

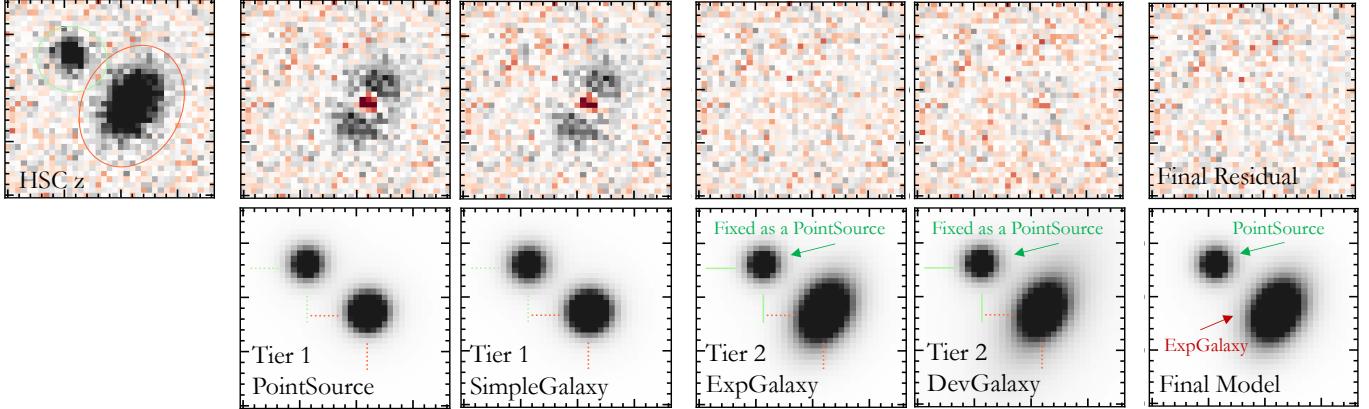
Now that models types have been assigned and their parameters optimized for each source in a given group, it is straight forward to apply these parametric models to photometer the sources in other bands of interest. We can do this via “forced” photometry is to measure fluxes and their uncertainties for already known (detected) sources, fixing the model shape parameters and only allow flux ( $\alpha$  in Equation 1) to vary. However, The Tractor provides the flexibility to allow shapes and positions to vary as well; they can be unbounded or limited by a Gaussian prior. For example, it may be desirable to allow the shape to change in the presence of morphological differences between the model bands and the forced photometry band, or allow the position to vary if there are significant astrometric offsets (see discussions in Section A.3). The Farmer enables the user to choose which parameters (if any) are fixed during the forced photometry stage.

As before, fitting proceeds on a group-by-group basis so that the forced photometry can benefit from the same advantage as in the model stage by simultaneously optimizing all models belonging to a given group. Each model is convolved with the PSF of the band of interest and realized into the frame of the image, including images of different pixel scales to that of the detection image<sup>6</sup>. The group models are then simultaneously optimized until their joint likelihood converges, or until some maximum iteration set by the user. Figure 5 shows the results of forced photometry using the same sources from Figure 4. While this procedure is generally faster than the model stage, forcing photometry on dozens of images may approach a similar computational expense. For consistency, it is advisable to perform forced photometry for all bands, even if they were used in the modelling stage. Computational strategies are discussed in Section A.6.

### 3.4.2. Catalogs and Other Output

After the modeling stage, The Farmer produces an intermediate catalog containing the source IDs, including their brick and group numbers, followed by the detection

<sup>6</sup> Currently, The Farmer requires pixel scale homogenization, but this restriction will be removed in a future update.



**Figure 4.** Example of a decision tree process for a group containing two identified sources. The input  $z$ -band image is shown in the top left with colored ellipses around the two detected objects beside the final residual constructed from subtracting the best model determined by **The Farmer**. Shown rightward are the residuals and models for two of the three tiers in the decision tree from which the best model types were determined. This particular pair of sources satisfied the decision tree before reaching the CompositeGalaxy in tier 3. While model images are scaled by  $\log_{10}$  to highlight morphologies, the science image and residuals are scaled to  $\pm 3\sigma$  to highlight faint signal and any oversubtraction.

parameters from SEP. For each source, the best-fit model type (e.g., PointSource or ExpGalaxy) are recorded, as well as their best-fit parameters and associated uncertainties. Shapes and sizes are not measured for sources assigned unresolved models (e.g. PointSource and SimpleGalaxy). Fluxes and flux uncertainties are also measured for each source in every band used in the modelling stage.

A number of residual statistics are also included that provide valuable insight into the goodness-of-fit of a given model for a given source and band. In order to minimize contamination with neighbours, we consider only the pixels belonging to the source segment in the computing these estimates (same as in the decision tree). The primary statistic is  $\chi^2$ , already discussed in Section 3.4. Three other related statistics are produced by measuring the moments of the inverse variance weighted  $\chi_i$  images where each  $i$  pixel value indicates the significance of the residual in units of per-pixel uncertainty  $\sigma_i$ : the median  $\mu(\chi)$ , standard deviation  $\sigma(\chi)$ , and D'Agostino's  $K^2$  test which measures the normality of the residual by combining estimates of skew and kurtosis<sup>7</sup> (D'Agostino 1970; D'Agostino & Belanger 1990). These statistics can also be combined to separate reliable models from poor fits and blends, as shown in Figure 6.

Once forced photometry is completed, **The Farmer** appends the measurements to (a copy of) the existing model catalog. This can be done on a band by band basis, or for all bands simultaneously. Output includes fluxes, as well as other parameters including

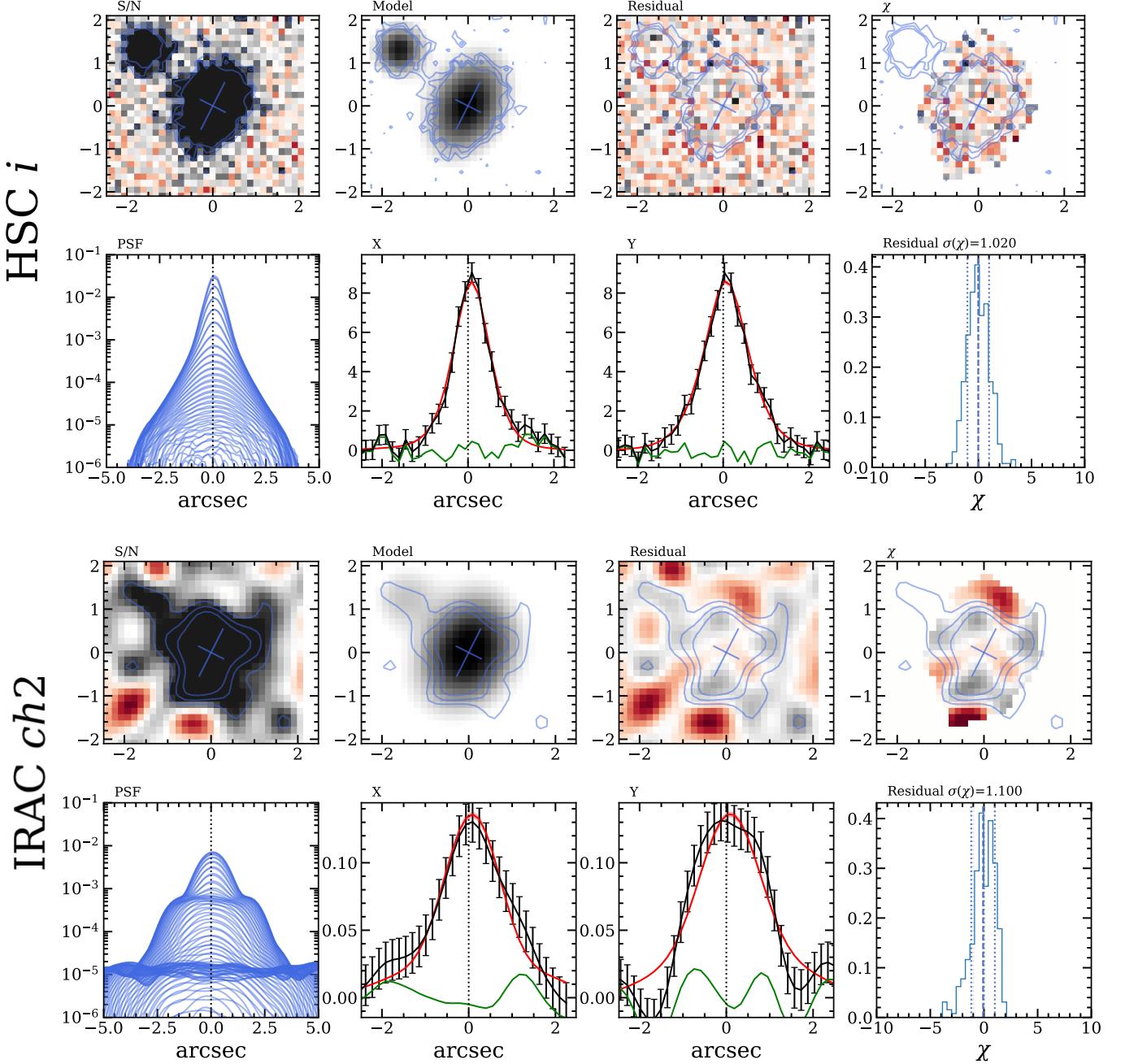
band-specific positions and shapes if the user has allowed them to vary. Residual statistics are also included for every source in each measurement band.

**The Farmer** has an additional diagnostic ability to measure photometry of these known, detected sources with concentric circular apertures of various diameters specified by the user. This is especially useful for constructing profiles of the radial flux growth. Aperture photometry can be measured on the science images (to get basic comparisons with the profile-fitting measurements), and it can go further by forcing the same apertures on the residual image and weight images. Most interestingly, these apertures can be forced on model images constructed by realizing the entire group of models into pixel-space. The aperture fluxes can then be readily compared with fluxes measured on the same apertures on the science image. Similarly, apertures can be forced on single sources realized into the pixel-space of the image in complete isolation; measurements in large apertures be compared with the total flux reported by **The Tractor**<sup>8</sup>. Together, these aperture measurements can help diagnose model inaccuracies and bias, providing an effective means to internally validate the results of **The Farmer**.

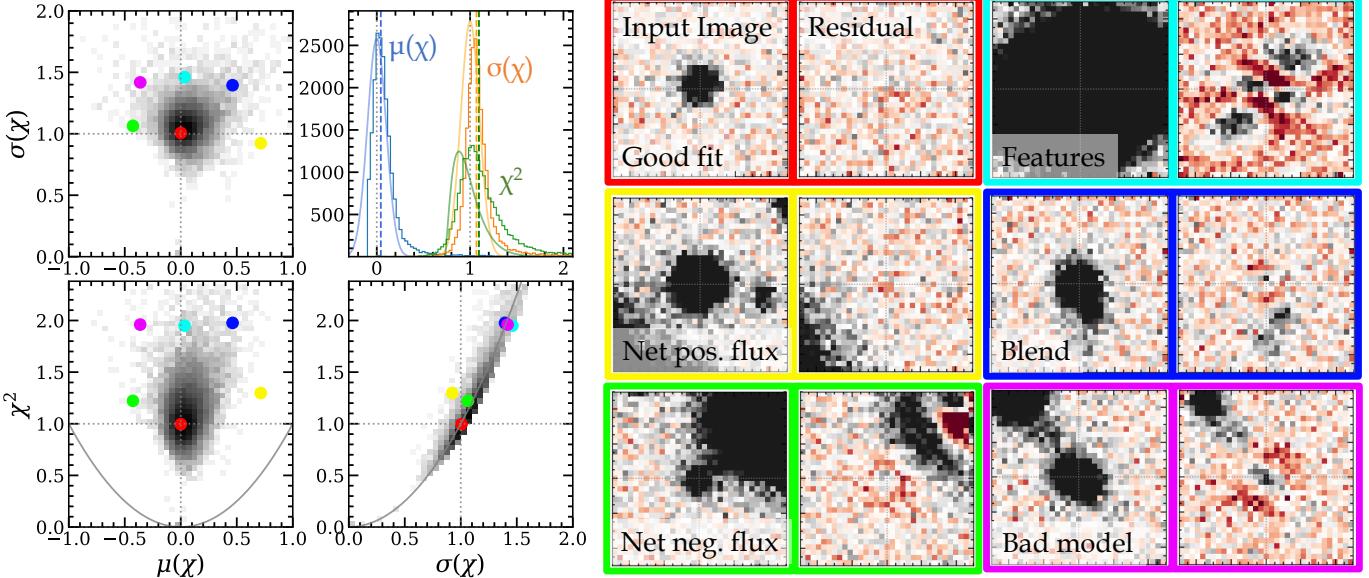
Diagnostic images can be incredibly useful. **The Farmer** can be configured to produce pixel-level background and RMS maps in addition to source and group segmentation maps. Importantly, **The Farmer**

<sup>7</sup> The  $K^2$  test is generally stable only for sources which have more than 8 pixels in their segment.

<sup>8</sup> However, if the model is severely truncated by being realized into an image whose dimensions are much smaller than the full extent of the model then the integrated flux in large apertures will underestimate the total, correct flux measured by the normalization coefficient.



**Figure 5.** Example of results from forced photometry for the brightest source in Figure 4 measured in *i* and channel 2. The image of each band is shown next to the best-fit model from The Farmer. Brightness contours and principal axes are overlaid on the model in blue. The residuals are shown  $\pm 3\sigma$  (same as the image) to highlight faint signal and any over subtraction. The rightmost panel shows the per-pixel  $\chi$  image scaled  $\pm 3$  computed within the bounds of source segment. Bottom rows for each band show a wiremesh representation of the PSF profile; slices though the source (black), best-fit model (red), and residual (green); and the distribution of  $\chi$  values over the group pixels which on expectation should be normally distributed. The median and 68% range of the distribution is shown for illustration.



**Figure 6.** The **Farmer** provides detailed statistics to readily select problematic sources such as blends, poor model fits, and artifacts that are difficult to identify from  $\chi^2$  alone. The red, yellow, and green sources are well-fit with a small variation in negative and positive flux. The cyan source contains features not well described by our smooth model profiles, the blue source is an unidentified blend, and the purple source is a pointsource model assigned to a resolved galaxy. Each of these six sources is selected from unique regions of a 3D statistic space (left). Each statistic is measured within the residual segment pixels of each source and include the  $\chi^2$ , the median  $\chi$  distribution  $\mu(\chi)$ , and the standard deviation of the  $\chi$  distribution  $\sigma(\chi)$ .  $\chi^2$  and  $\sigma(\chi)$  are closely related; grey curves indicate the allowed regions of the joint distributions. The colored histograms show their observed distributions while solid curves indicate expectation:  $\langle \chi^2 \rangle = 1$ ,  $\mu(\chi) = 0$ , and  $\sigma(\chi) = 1$ .

can realize the entire model library of a brick as a reconstructed pixel-level model image from which responding residual and weighted significance  $\chi$  images can be produced. Since catastrophic failures can result in models spanning large regions of the reconstructed model images, The **Farmer** allows the user to automatically filter models based on  $\chi_N^2$  or axis ratio such that they are not included in the reconstructed model, residual, or  $\chi$  images (especially useful for cleaning residuals when searching for undetected signal). Also, models with negative fluxes will create positive flux in residual images; these can also be automatically filtered. Although removing sources at this level introduces incompleteness, it is likely that the measurements of these problematic sources are not scientifically useful anyways. To account for the missing area, The **Farmer** also provides an effective mask image which flags pixels belonging to removed sources according to their segment ownership and computes the effective area of that mask. Although laborious, this is an optimal system for precisely determining the effective area from which a cleaned sample has been selected. Caveats regarding these reconstructed images are discussed in Section A, below.

#### 4. BENCHMARK AND VALIDATION

In this section we test and validate the performance of The **Farmer** using a set of simulated deep images with COSMOS-like properties.

##### 4.1. Construction of Mock Images

The construction of the mock images used here follows the approach presented in L. Zalesky et al. (in prep.). Images are created to include a number of realistic features. The noise in each filter is matched to the RMS measured on real images used in Weaver et al. (2022). Galaxy-type sources are included with random positions and orientations using the open-source code **GalSim** (Rowe et al. 2015) via the **RealGalaxy** class, which allows the user to inject images of real sources observed by *HST* in the COSMOS field. Unfortunately, the morphology of these sources is only available at the resolution of *HST* in one filter (F814W). In order to simulate wavelength-dependent profiles, we use parametric model representations of these galaxies (bulge+disk composites), and give red spectral energy distributions to bulge components and blue spectral energy distributions to disk components; this is handled internally within **GalSim** by the **RealGalaxy** class. To ensure a realistic colors for each galaxy, we have cross-matched the *HST* catalog internal to **GalSim** to the COSMOS2020

catalog, and re-scaled the flux in each band that we simulate to that of the matched source.

The shape of the galaxy number counts is fixed by the internal *GalSim* catalog, and all we modify is the normalization, such that resolved galaxies comprise  $\sim 2/3$  of all sources at intermediate magnitudes ( $20 < m_i < 24.5$ ). **The GalSim counts are incomplete beyond  $m_i \gtrsim 25$ , and so we inject PSF-models with The Tractor, assuming a constant PSF in each band. This is reasonable since Weaver et al. 2022 showed that objects fainter than 24.5 AB in COSMOS2020 are generally unresolved. This means that the injected point sources are a fair test of The Farmer’s ability to identify resolved and unresolved sources: had we somehow injected realistically sized galaxy models they would appear as unresolved sources anyway and so The Farmer would have rightly fit them as such.** The fluxes of these point sources are tuned such that together with the galaxy sources, the total sample yields a complete sample in the HSC-*i* band to 28.5 magnitude. Colors of point sources are assigned by randomly selecting sources of similar flux (within  $\pm 0.1$  mag) from the COSMOS2020 catalog and scaling the flux in a given filter to match the color. Finally, the number counts are calibrated and scaled according to the number counts of the COSMOS2020 catalog and to those in the empirical mock catalog of Girelli et al. (2020).

**Our optical and NIR images are simulated at the same scale as the images used in Weaver et al. (2022) (0.15"/pixel). Likewise, we also simulate the mid-IR *Spitzer*/IRAC images at their native resolution of 0.6"/pixel and then use Swarp (Bertin 2010) to resample them to 0.15". This step introduces correlated noise that affects the effective degrees of freedom of the model fits (see Section 4.3).**

Although the galaxies in our simulated images are parametric representations, it should be noted that real galaxies feature structures such as spiral arms and starbursting regions that are not captured by these models. As such, the performance of The Farmer for the brightest sources assessed on this simulation is likely overestimated compared to real galaxy images.

## 4.2. Procedure

We follow the general procedure outlined in Section 3. For simplicity and to ease the interpretation of our tests, we adopt the input PSFs used to produce the simulated images. No backgrounds are subtracted. These two aspects of our procedure are functionally equivalent to perfect knowledge of the PSF and of the image background.

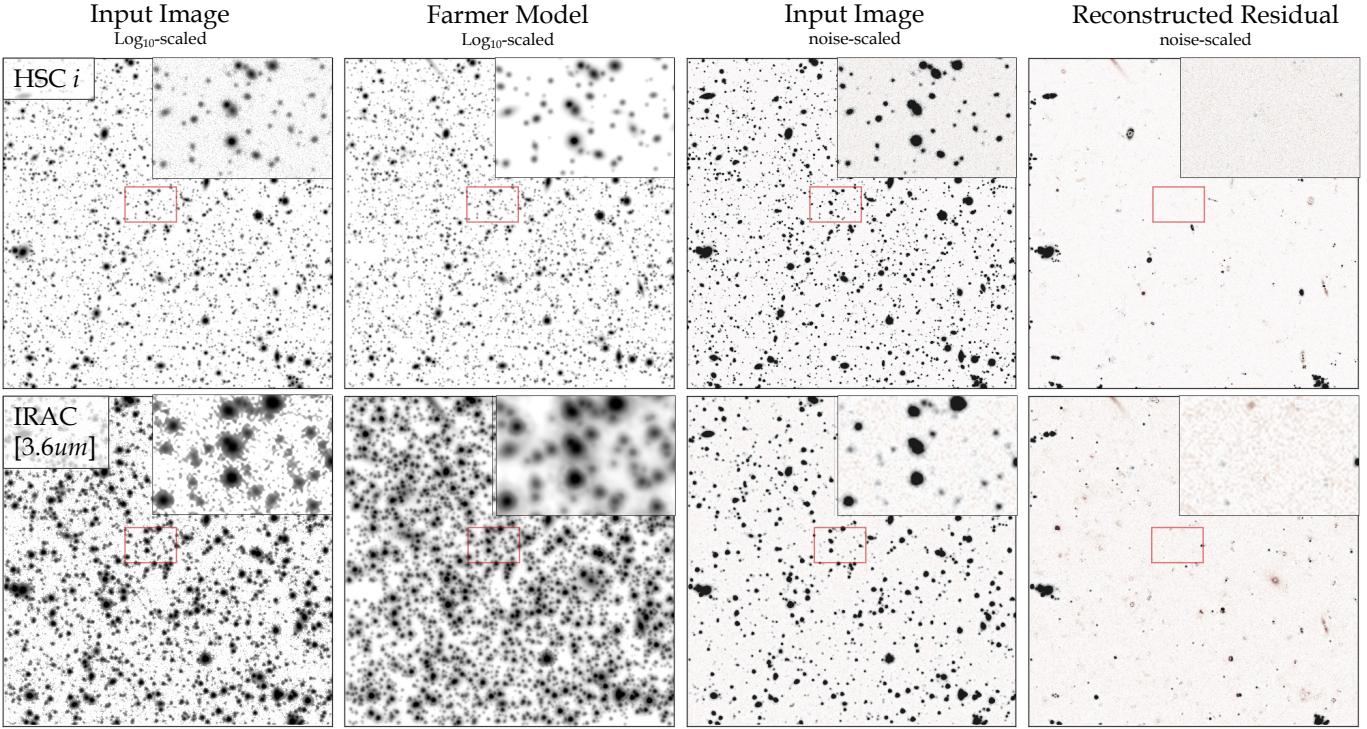
Sources are detected on a  $izK_s$  CHI-MEAN image created using Swarp (Bertin 2010). **We do not model on the detection image due to the combined, chromatically dependent PSF. For this reason, and to preserve our selection function, we determine our best-fit model types, positions, and shapes by jointing fitting the same bands that constitute our detection image (*i*, *z*, and  $K_s$ ) using their respective PSFs.** Models are assigned according to the same decision tree structure as described in Figure 3. Sub-optimal model assignments such as assigning resolved models to many point-like sources tends to produce a photometric bias which manifests as a plateau or sharp rise in source magnitude distributions (i.e. number counts) at the threshold in magnitude beyond which sources are generally unresolved. Therefore we tune the decision tree to produce smoothly increasing number counts, and then tune further by spot checking residuals of individual sources; **determining in  $\chi^2_N$  a Pointsource penalty of 0.1 and a resolved model similarity threshold of 0.2.** The modelling stage is run, which assigns models and optimizes their parameters on a group basis.

**We perform force photometry by re-fitting the models on the bands of interest: *r*, *i*, *z*,  $K_s$ , channel 1, and channel 2. Positions and shapes are fixed for each object, with only the five independent fluxes free to vary.** Figure 7 shows the reconstructed model images and residuals produced by The Farmer over a region of the simulated *i* and channel 1 mosaics. The vast majority of sources are well modelled with only a handful of failed fits which are left in the residual map. While the value of visual inspection of residuals cannot be understated, a rigorous statistical analysis can provide powerful quantitative insight.

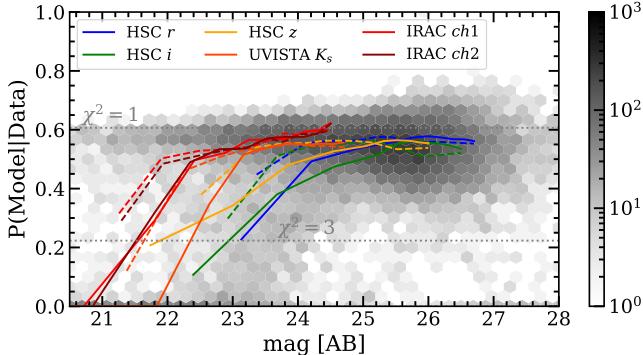
## 4.3. Model and Decision Tree Performance

Now we use the suite of statistics provided by The Farmer to assess the performance of the models and decision tree.

As demonstrated by Figure 8, the probability of the model given the data (inversely proportional to  $\chi^2_N$ ) is greatest for faint sources across all bands. For images of high spatial resolution (e.g. *r*, *i*, *z*), the model performance degrades for both resolved and unresolved models at magnitudes brighter than  $\sim 24$  AB, although with considerable variance. These bright sources are smooth in our simulations, however they are still more complex than the models supplied by The Tractor. Additionally, brighter sources usually subtend a larger area and so reside in more complex groups where blending makes accurate photometry more challenging. **A notable ex-**



**Figure 7.** Results of forced photometry by **The Farmer** on simulated fields of depths similar to COSMOS. Sources are  $izK_s$ -detected and modelled on  $i$ ,  $z$ , and  $K_s$  jointly, then forced on other bands including channel 1. Models can be compared to input images in the two leftmost panels in  $\log_{10}$  scaling where morphology is visible. Insets show a zoom-in of a smaller region. Residuals can be compared to input images in the two rightmost panels in  $\pm 3\sigma$  scaling to highlight faint signal above (black) and below (red) the background. While some sources were skipped as they were too complex, other notable over-subtractions are driven by blends and/or injected sources that are not well-described by the five model types.



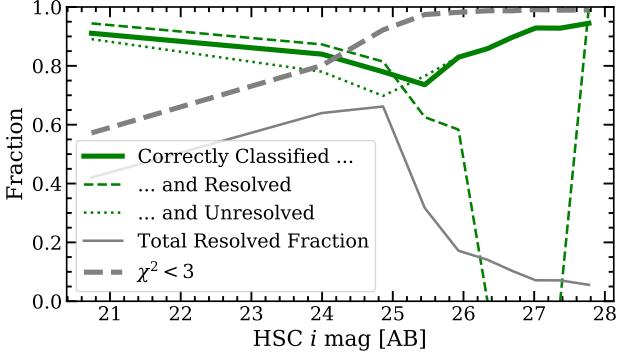
**Figure 8.** Probability of models as a function of apparent magnitude. Results from  $i$  are shown in grey histograms, summarized by binned medians of unresolved (solid) and resolved models (dashed). Other band are similarly summarized by binned medians.

ception are bright point-like sources which are typically well-fit by the PointSource model type.

The NIR and IR bands ( $K_s$  and IRAC) have slightly better performance at bright magnitudes. This is because their resolution threshold is at a brighter magnitude and so these particular bands contain a higher

fraction of bright sources which appear unresolved. Whether or not **The Farmer** assigned resolved or unresolved models to these sources, the resolution is low enough that they are effectively unresolved. Photometry is then made easier because there is little dependence on accurate model shapes. The key insight therefore is that the effectiveness of profile-fitting photometry is not dependent on source magnitude directly, but rather on the size of the source and whether or not it is resolved, with some lesser dependence on the resolution of the bands used to derive the models.

As shown in Figure 9, **The Farmer** is generally able to correctly assign resolved models to sources which are injected as resolved galaxies, and unresolved models to those which are injected as unresolved point sources (which could be stars or galaxies – **The Farmer** does not try to separate them). As alluded to earlier, the resolution threshold averaged over the modelling bands ( $\sim 25$  for  $izK_s$ ) is where it is most difficult for **The Farmer** to distinguish between resolved and unresolved sources and so ultimately the fine tuning of the decision tree is aimed at improving performance in this

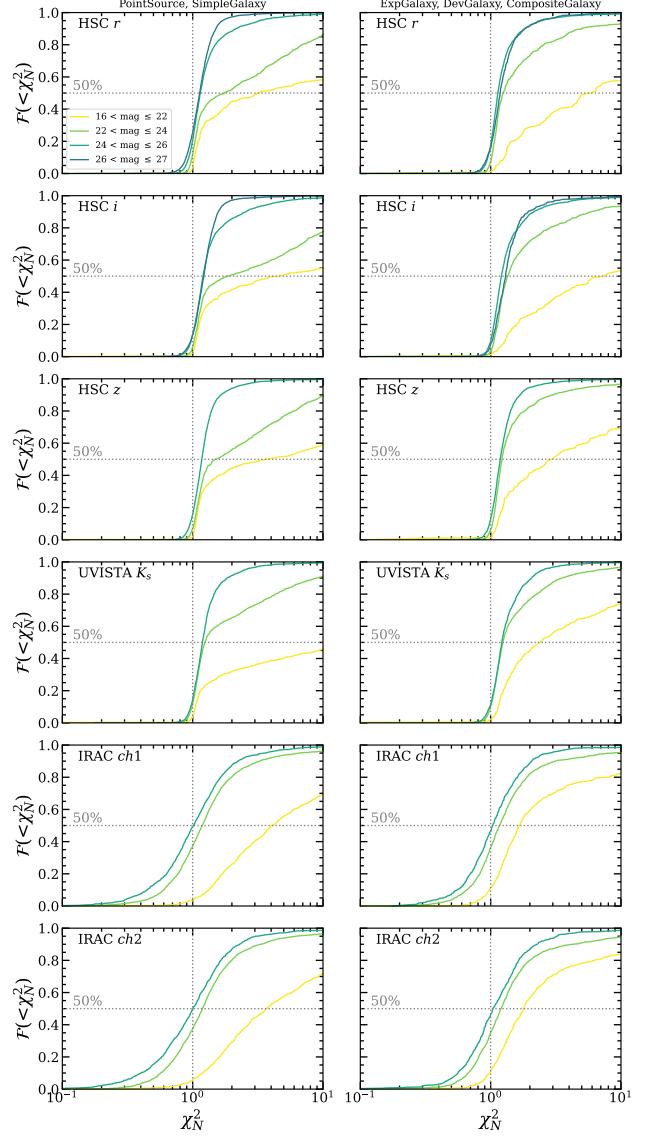


**Figure 9.** Fraction of unresolved (resolved) sources correctly assigned an unresolved (resolved) model by **The Farmer** as a function of apparent  $i$ -band magnitude is shown by the solid green curve, and broken down into resolved and unresolved subsets by the dashed and dotted curves, respectively. The in recovery fraction at  $i \sim 25$  is expected as sources at that brightness are typically only marginally resolved and so challenge the decision tree. We consider only sources with  $\chi^2 < 3$  in the  $i$ -band as they are considered reliable, the fraction of which is shown by the grey dashed curve. Our simulated field at  $i > 25$  uses mostly point-like sources to reflect real conditions in a COSMOS-like survey; the total resolved fraction is shown by the solid grey curve.

regime. Based on our tuning, **The Farmer** correctly assigns  $> 75\%$  of marginally resolved sources.

While it appears that **The Farmer** is not able to correctly assign resolved models to injected resolved galaxies at  $i > 25$ , this is almost certainly because these sources actually appear unresolved in our simulated images. It should be noted, therefore, that while a given source in the simulated images corresponds to either a resolved galaxy or unresolved point source model, the former may be effectively unresolved in the image if it is smaller than the PSF. Identifying such cases in the  $i$ ,  $z$ , and  $K_s$  bands is therefore of interest as **The Farmer** should not be expected to assign them a resolved model. These cases cannot be cleanly identified beforehand, nor is it possible to identify them afterwards with full confidence. As a result, the performance of **The Farmer** may be expected to be better than it appears in Figure 9 around the  $i \sim 25$  resolution threshold.

The performance of models optimized in forced photometry is also generally better at faint magnitudes where sources are typically unresolved. Figure 10 shows the fraction of sources below a given reduced  $\chi^2_N$  in four ranges of magnitude for each band separated into unresolved and resolved model types. A sample which is  $\chi^2$  distributed reduced by  $N$  degrees of freedom should have an expectation value of unity. Its cumulative distribution should therefore be approximately evenly divided



**Figure 10.** Fraction of sources below a certain  $\chi^2$  as a function of band and magnitude for unresolved (left) and resolved models (right).

around  $\chi^2_N \approx 1$ . It should be noted that  $\chi^2$  is a measurement of significance and is therefore dependent on accurate per-pixel errors.

The performance of models for the well-resolved bands ( $r$ ,  $i$ ,  $z$ ,  $K_s$ ) is better for faint sources irrespective of resolved or unresolved models. Overall these distributions seem slightly shifted towards larger values of  $\chi^2_N$ . Inspection of the residuals suggest these models are well fit, and so this shift may be due to inaccurate per-pixel errors, or pixel covariance which is not accounted for by  $\chi^2$  which assumes independent, Gaussian distributed data. For bright sources, a tail develops at  $\chi^2_N > 10$  which also suggests an increased fraction of bad models.

1154 This is expected as any imperfection in the model  
 1155 will add some term proportional to the square of  
 1156 the source flux.. By inspection, we confirm that the  
 1157 complexity of the injected galaxies is not always well-  
 1158 captured by the smooth models from **The Tractor** (as  
 1159 would happen in real images). Source crowding may  
 1160 also play a role for these typically large, bright sources  
 1161 that may have fainter sources near their wings that if  
 1162 not detected may cause a photometric bias.

1163 The two infrared bands (channel 1 and 2) appear to  
 1164 have slightly better performance at faint magnitudes.  
 1165 There does not seem to be a shift, which relative to  
 1166 the bluer bands may be due to greater degree of signal  
 1167 covariance relative to the bluer bands (from the larger  
 1168 PSF) whereby a good fit in one pixel means one can  
 1169 expect to achieve a good fit in the adjacent pixels. A  
 1170 tail does not develop for bright sources, which instead  
 1171 are shifted towards higher  $\chi^2$ . This systematic behav-  
 1172 ior suggests that **The Farmer** has the greatest difficulty  
 1173 modelling the bright IRAC sources in general. **This**  
 1174 **is not a surprise given that the IRAC images**  
 1175 **have worse resolution, meaning that light from**  
 1176 **neighboring (bright) objects can impact sources**  
 1177 **in a given group. Because this extra light is not**  
 1178 **expected by the group model, it may lead to a**  
 1179 **photometric bias.**

1180 **The Farmer** also provides accurate shape measure-  
 1181 ments for all resolved sources. Figure 11 demonstrates  
 1182 the recovery of axis ratio and position angle of the sim-  
 1183 ulated galaxies, finding agreement within 1 per cent.  
 1184 There are no obvious biases in any parameter, whether  
 1185 compared to itself, source magnitude, Sérsic index, or  
 1186 local source density. The only notable deviations are  
 1187 expected: circular sources with  $b/a \sim 0$  where the axis  
 1188 ratio signal is very weak and small sources where  $R_{\text{eff}}$   
 1189 approaches the pixel scale of the image ( $0.15''/\text{px}$ ). The  
 1190 insensitivity to local source density gives **The Farmer**  
 1191 a considerable advantage over shapes estimated from  
 1192 **Source Extractor**.

#### 1193 4.4. Counts and photometric accuracy

1194 Credible survey science ultimately rests on a founda-  
 1195 tion of complete samples and accurate photometry. We  
 1196 characterize the relevant performance of **The Farmer** in  
 1197 the following assessments.

1198 Source number counts not only diagnose issues in sam-  
 1199 ple selection and incompleteness, but are also sensitive  
 1200 to photometric accuracy. The number counts of injected  
 1201 sources in our simulated images are shown alongside  
 1202 those recovered by **The Farmer** in Figure 12. The re-  
 1203 covery of number counts is generally excellent. They  
 1204 are complete up the limiting magnitude of each band,

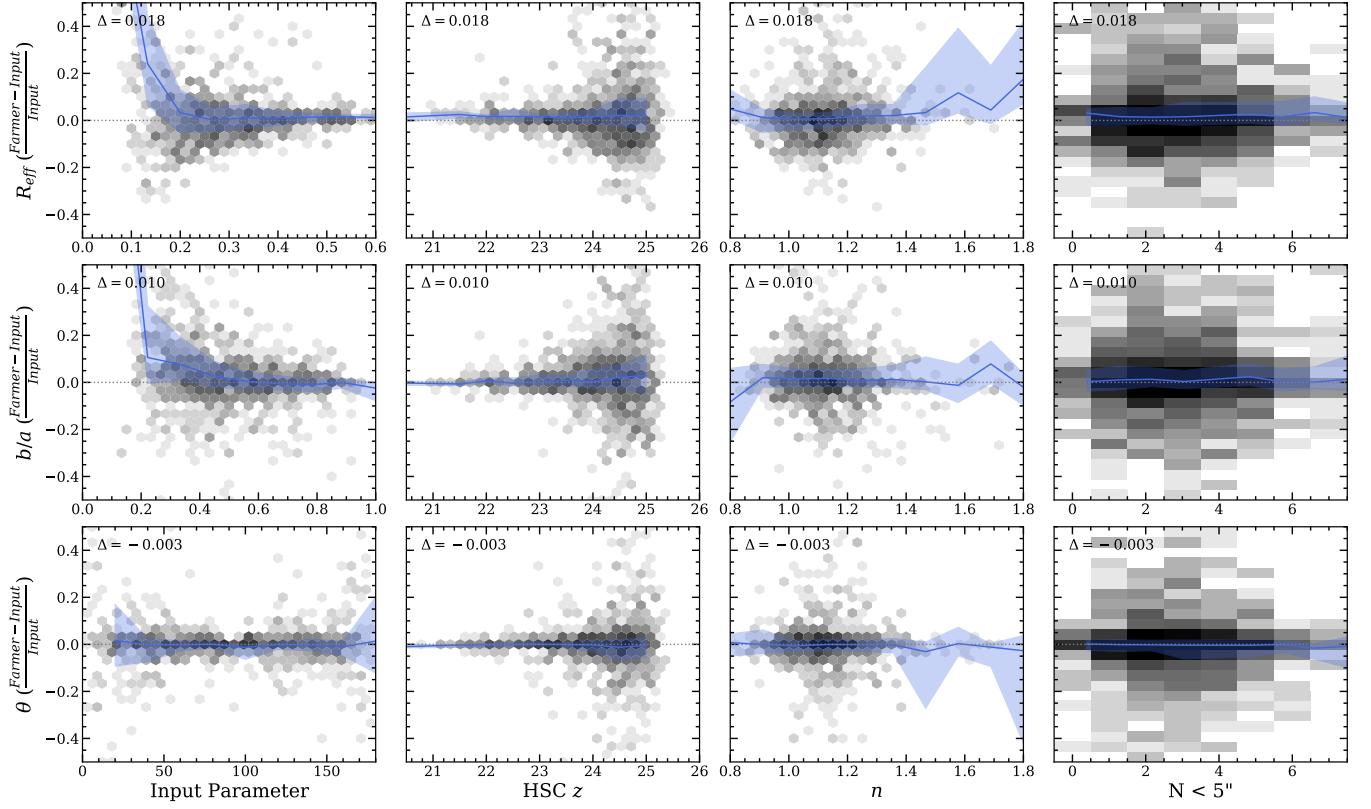
1205 which is most important for the  $i$ ,  $z$ ,  $K_s$  bands used in  
 1206 sample selection as incompleteness in other bands may  
 1207 be driven by selection effects. For instance, a small frac-  
 1208 tion of faint  $r$ -band sources is missing from our sample  
 1209 as expected given the simulation includes real galaxy  
 1210 colors and these predominantly blue sources are likely  
 1211 faint in our redder detection image. We can trust that  
 1212 **The Farmer**'s decision tree is performing well given that  
 1213 there are no extended plateaus or sharp rises present  
 1214 anywhere in the number counts, in combination with  
 1215 other available diagnostics (e.g., residuals,  $\chi^2$ , etc.).

1216 The most important measurement is ultimately pho-  
 1217 tometry. As shown in Figure 13, the photometry mea-  
 1218 sured by **The Farmer** is seen on median expectation to  
 1219 be accurate below 0.05 AB in all bands, including IRAC.  
 1220 There are no significant systematic biases, with only  
 1221 a small trending towards overestimated fluxes for faint  
 1222 sources in  $K_s$ . The 68% scatter is similar to the typical  
 1223 magnitude uncertainty at a given magnitude for  $r$ ,  $i$ ,  $z$ ,  
 1224 and  $K_s$ . For IRAC bands, the scatter is about three  
 1225 times larger than the typical magnitude uncertainty,  
 1226 suggesting that the photometric uncertainties may be  
 1227 underestimated. This may be expected given the high  
 1228 spatial covariance of noise in IRAC images.

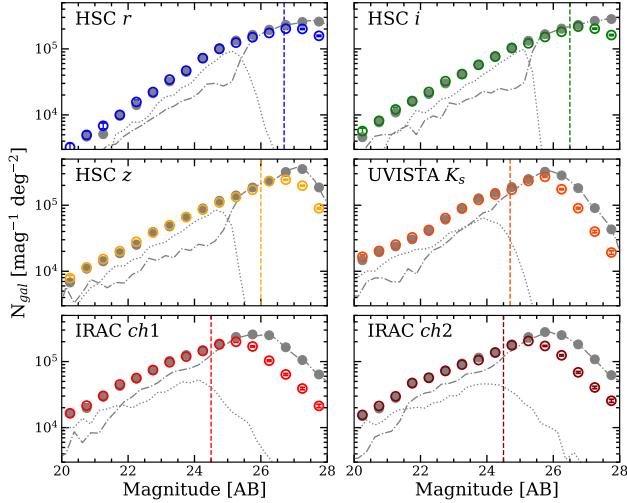
1229 Photometric measurements are more appropriately as-  
 1230 sessed by directly examining the cumulative distribu-  
 1231 tions (CDF) of relative error  $\epsilon = |f_{\text{input}} - f_{\text{Farmer}}|/\sigma$ .  
 1232 These are shown in Figure 14 broken down by band  
 1233 and separated into resolved and unresolved model types.  
 1234 On expectation, 68% of sources should be contained by  
 1235  $|\epsilon| \leq 1$ . Given the lack of bias in our photometry, devia-  
 1236 tions of the  $\epsilon$  CDFs from this expectation can be directly  
 1237 attributed to inappropriate flux uncertainties resulting  
 1238 from miscalibrated weights and/or spatially covariant  
 1239 noise.

1240 We see a similar picture to the  $\chi^2$  CDFs in Figure 10  
 1241 whereby photometry of faint sources measured in the  
 1242 high spatially-resolved bands ( $r$ ,  $i$ ,  $z$ , and  $K_s$ ) better  
 1243 follows expectation compared to photometry of bright  
 1244 sources. The distribution of  $\epsilon$  for bright point sources  
 1245 has a tail as even the smallest biases are expected to  
 1246 yield large  $\epsilon$  values as the typical flux uncertainties are  
 1247 small. However, the same is not true for the resolved  
 1248 models which are systematically shifted towards larger  
 1249  $\epsilon$  with increasing brightness. This may suggest poor  
 1250 modelling performance of the brightest sources, in ac-  
 1251 cord with previous results.

1252 The  $\epsilon$  CDFs for the IRAC bands are significantly  
 1253 shifted towards higher values in agreement with the re-  
 1254 sults from Figure 13. This is further evidence that the  
 1255 weights from our IRAC mocks may produce underes-  
 1256 timated photometric uncertainties. This is not an im-



**Figure 11.** Recovery of effective radius ( $R_{\text{eff}}$ ; top row), axis ratio ( $b/a$ ; middle row) and position angle ( $\theta$ ; bottom row) as a function of input parameter,  $z$ -band magnitude, Sérsic index ( $n$ ), and the local source density (number of sources within  $5''$ ). Greyscale distributions are summarized by binned medians (colored curves) with 68% range indicated by the envelope.

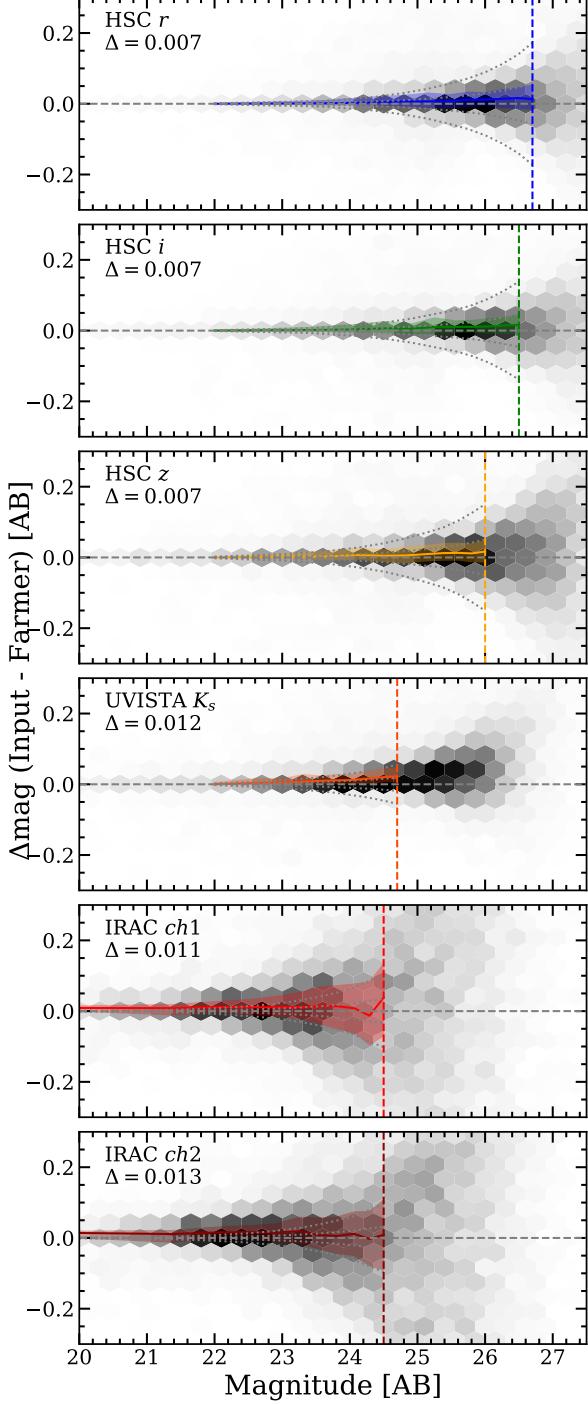


**Figure 12.** Number counts are shown for each band corresponding to the simulation input for all sources (filled grey points), resolved sources (grey dotted curve), and unresolved sources (grey dash-dot curve). This is compared to output from The Farmer (unfilled colored points with Poisson uncertainties) for an  $iz$ -selected sample. Nominal depths are shown by the vertical colored lines.

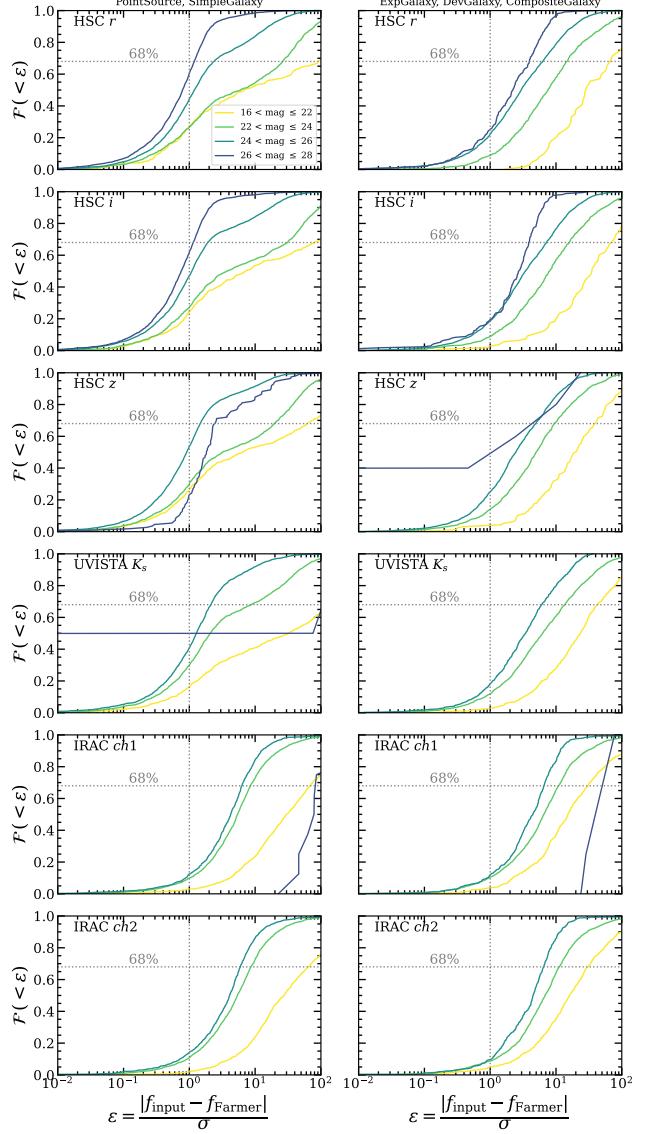
mediate confirmation, however, because both  $\chi^2$  and  $\epsilon$

assume independent, Gaussian distributed data which  
1258 may not be the case in instances of significant pixel co-  
1259 variance; e.g. as in the case of IRAC as it has been  
1260 up-sampled such that the PSF is correlated across more  
1261 pixels. While this is treated to some degree by  
1262 SWarp, the resulting weights seem to still be over-  
1263 estimated.

1264 Another way to investigate typical model accuracy  
1265 is demonstrated in Figure 15. As described in Sec-  
1266 tion 3.4.2, The Farmer can be configured to extract flux  
1267 in circular concentric apertures at every source pos-  
1268 ition. We have measured fluxes in several aperture sizes  
1269 with sub-arcsecond steps for both resolved and unre-  
1270 solved models computed on the group images, models,  
1271 and residuals. Fluxes are also measured consistently  
1272 for each source individually, such that they are real-  
1273 ized in isolation of other sources (the ‘isomodel’). The  
1274 largest aperture is  $6''$  in diameter which likely captures  
1275 flux from neighboring sources in the  $i$ -band image used  
1276 here. Expectedly, while the ‘image’ and ‘model’  
1277 flux grow beyond the input source flux due to  
1278 the presence of neighbors, that of the ‘isomodel’  
1279 tends towards agreement with the input source



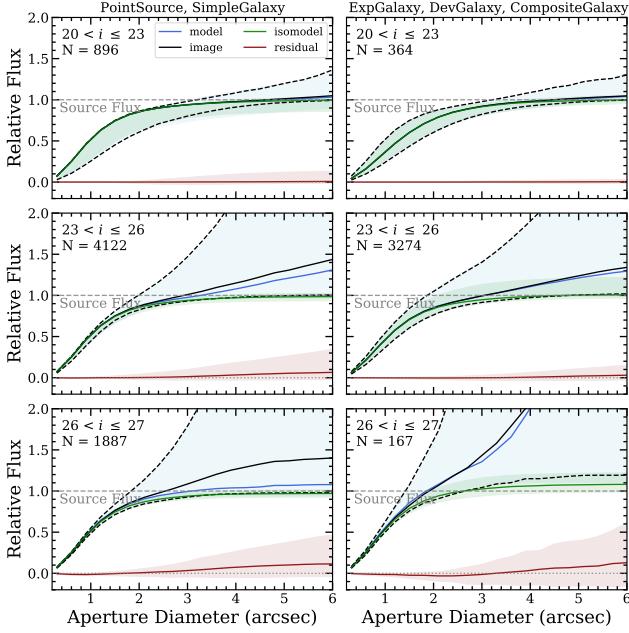
**Figure 13.** Photometry produced by The Farmer is compared with true fluxes of simulated sources for all bands. Differences in magnitude as a function of input magnitude (grey histograms) are summarized by binned medians (colored curves) with 68% ranges indicated by the colored envelope out to the nominal depth limit of each band (vertical colored lines). Expected  $\pm 1$  and  $3\sigma$  uncertainties on  $\Delta\text{mag}$  are computed as medians from the The Farmer uncertainties (grey dotted curves).



**Figure 14.** Fraction of sources whose relative photometric error  $\varepsilon$  is less than a certain value, broken down by resolved (left panels) and unresolved models (right panels) for each band. On expectation,  $|\varepsilon| < 1$  for 68% of sources where a significant departure may indicate under- or over-estimation of photometric uncertainties.

flux (i.e. 1), and that of the residuals generally tends towards zero.

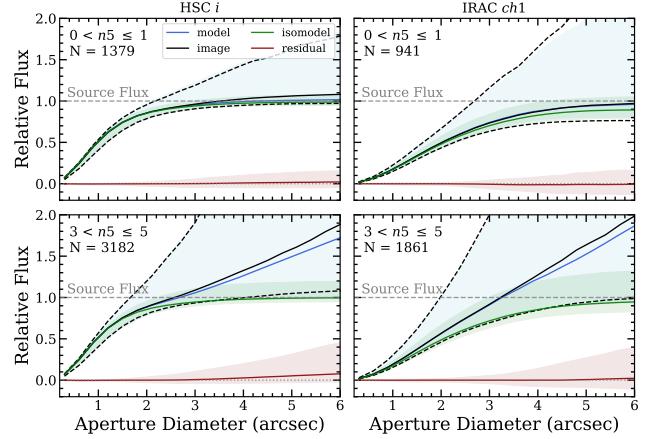
Bright sources are typically large on the sky such that the largest apertures are dominated by the bright source with insignificant contributions from faint neighbors. The apertures measured on the image, model, and isomodel agree well for both resolved and unresolved bright sources, and tend towards agreement with the true input flux at large radii (a value of 1 on the y-axis). Interestingly, the flux at small radii varies



**Figure 15.** The accuracy of the models in various magnitude regimes can be assessed by measuring fluxes are in circular apertures of increasing diameter on simulated input *i*-band image (black), The Farmer joint model with neighbors (blue), their residual (red), and The Farmer model without neighbors (green). **For easy comparison, the y-axis shows each aperture flux measurement normalized to the true input flux** and are summarized by a median and 68% range, which are shown as colored regions except for the input image which is shown as black dashed lines. While faint sources contain higher fractions of blends at a given radius (the black rises above unity in the lower panels), The Farmer is still able to recover the input source flux (green tends to unity at large radii).

significantly. This is driven by the variation in light profiles (i.e. Sérsic index) that's more visible for bright, well-resolved sources. In the case of sources fit with PointSource or SimpleGalaxy models, the variation is driven entirely by the different light profiles. As one might expect, including only sources fit with PointSource models results in almost no variation whatsoever as all point sources have the same curve of growth.

The behavior is different for fainter sources. While their image and model fluxes continue growing even at large apertures, the flux of the isomodel stops growing around  $3''$  as no new flux is captured by the apertures and agrees with the true input flux. The situation changes again for the faintest sources where on average there is blending at radii smaller than  $3''$  as shown by the divergence of the black image and blue model flux growth curves from that of the isomodel in green that on average agrees with the true input flux. Hence, while



**Figure 16.** Summary of de-blending power of The Farmer. Similar layout to Figure 15, but for sources photometered in *i*-band (left) and channel 1 (right) broken down by local density  $n_5$  defined by the number of sources within  $5''$ .

there is blending of sources within even  $2\text{-}3''$  apertures in *i*-band, the approach used by The Farmer produces fluxes which are not typically affected by blending<sup>9</sup>.

#### 4.5. Deblending in IRAC

Here we assess the de-blending performance of The Farmer more thoroughly in the context of our simulated IRAC images in Figure 16.

Similar to Figure 15, photometry is measured in apertures forced on source positions computed on the images, models, isomodels, and residuals. As a baseline, growth of flux for sources measured in *i* are in agreement between the images, model, and isomodel, as well as the true total flux for isolated sources. However, for sources in crowded regions the flux measured on the image and model continues to grow whereas that of the isomodel flattens out around  $4''$  in agreement with the true flux.

Although IRAC images have very different properties compared to HSC's *i* band, the behavior for isolated sources is similar. The only difference being that larger apertures are required to encompass the total flux of IRAC sources. Aperture photometry measured in crowded regions of IRAC images, however, quickly become contaminated by the flux of neighbors so that no aperture diameter can cleanly measure the total flux of the central source. While the encompassed flux from both the image and model apertures grows exponentially, that of the isomodel finds good agreement with the true total flux of the simulated source. What is incredible is that the flux growth curve of the isomodels

<sup>9</sup> This will not be true in cases where blended sources are not separated by detection, see Section A.

in green deviates from that of the total group images in black and their joint models in blue already below 2'', meaning that de-blending is typically significant in our IRAC images even on these small scales. As such, the only tenable way to obtain accurate, high signal-to-noise photometry of IRAC sources is with a profile-fitting approach which, crucially, provides for joint modelling with neighboring sources as employed by **The Farmer**.

## 5. SUMMARY AND OUTLOOK

While deep galaxy surveys from space-based facilities offer exquisitely resolved images, ground-based surveys are capable of efficiently obtaining similar depths over significantly larger areas where searches for rare populations can be conducted, although at the cost of resolution. Already such survey images contain source densities that demand increasingly smaller aperture photometry to avoid crowding, which results in more uncertain measurements (Laigle et al. 2016; Weaver et al. 2022). As we have demonstrated, **aperture photometry will grow less reliable as extragalactic fields deepen and become more crowded**. Investments in deep, ground-based surveys will continue in the coming decade and so it should be expected that the magnitude of these challenges will only increase. Profile-fitting methods have been a longstanding technique for measuring low-resolution infrared images as they are less susceptible to source crowding. However, their advantages are now needed in the optical and near-infrared regimes. **The Farmer** attempts to answer this call.

We have explored the methodology of **The Tractor** whose photometry does not require that images be PSF homogenized, and total fluxes are reported solely based on the scaling of the model profile; avoiding the need for often ill-posed aperture corrections. However, we highlighted several obstacles preventing us from directly applying **The Tractor** to deep, crowded galaxy fields. These problems were solved by developing **The Farmer** which leverages an efficient albeit complex decision tree to assign models to sources in an optimal and less pathological way compared to simpler approaches. The decision tree is shown to be more than a useful algorithm, but indeed a required development in overcoming challenges related to blending in deep fields. **The Farmer** is also a means by which to organize survey data so that one can utilize massively parallelized computing facilities to streamline computational time from potentially years down to only a few weeks. Profile-fitting photometry is, however, more complicated than apertures and comes with its own drawbacks and considerations ranging from selection functions to image resolution, and from deblending capabilities to computational limits.

In a series of validation tests, we examined the ability of **The Farmer** to photometer sources in realistically simulated images. We found no significant biases in photometry in any band. Furthermore, we illustrated the unique advantage of **The Farmer** in de-blending sources in low-resolution images like IRAC. Still, bright and potentially resolved sources will continue present a limitation when employing smooth model profiles. On the other extreme, **The Farmer** has been shown to provide incredibly accurate photometry of the faintest unresolved sources, and in this sense it helps open the door to the distant universe.

Still, challenges in profile-fitting photometry remain and many difficult problems are yet unsolved. While we have demonstrated that **The Farmer** will provide accurate photometry for the next generation of deep, crowded fields, we must continue to innovate as we move towards deeper and more complex surveys promising even greater discoveries.

## ACKNOWLEDGMENTS

The authors thank Dustin Lang, Charles Steinhardt, Keith Horne, and Ranga-Ram Chary for helpful discussions.

The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. ST, GB and JW acknowledge support from the European Research Council (ERC) Consolidator Grant funding scheme (project ConTExt, grant No. 648179). OI acknowledges the funding of the French Agence Nationale de la Recherche for the project “SAGACE”. HJMcC acknowledges support from the PNCG. ID has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 896225. This research is partially funded by the Joint Survey Processing effort at IPAC/Caltech through NASA grant NNN12AA01C. The HST COSMOS program was supported through NASA grant HST-GO-09822. More information on the COSMOS survey is available at <http://www.astro.caltech.edu/cosmos>. This work used the CANDIDE computer system at the IAP supported by grants from the PNCG and the DIM-ACAV and maintained by S. Rouberol.

**Software:** `numpy` (van der Walt et al. 2011), `matplotlib` (Hunter 2007), `astropy` (Astropy Collaboration et al. 2013, 2018), `Source Extractor` (Bertin & Arnouts 1996), `PSFEx` (Bertin 2013), `SWarp` (Bertin 2010), `GalSim` (Rowe et al. 2015), and `The Tractor` (Lang et al. 2016)

## REFERENCES

- 1435 Aihara, H., AlSayyad, Y., Ando, M., et al. 2019,  
 1436 Publications of the Astronomical Society of Japan, 71,  
 1437 114, doi: [10.1093/pasj/psz103](https://doi.org/10.1093/pasj/psz103)
- 1438 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J.,  
 1439 et al. 2013, A&A, 558, A33,  
 1440 doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- 1441 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M.,  
 1442 et al. 2018, AJ, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- 1443 Barbary, K. 2016, The Journal of Open Source Software, 1,  
 1444 58, doi: [10.21105/joss.00058](https://doi.org/10.21105/joss.00058)
- 1445 Bertin, E. 2010, SWarp: Resampling and Co-adding FITS  
 1446 Images Together. <http://ascl.net/1010.068>
- 1447 Bertin, E. 2013, Astrophysics Source Code Library,  
 1448 ascl:1301.001.  
<http://adsabs.harvard.edu/abs/2013ascl.soft01001B>
- 1450 Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393,  
 1451 doi: [10.1051/aas:1996164](https://doi.org/10.1051/aas:1996164)
- 1452 Bertin, E., Schefer, M., Apostolakos, N., et al. 2020, in  
 1453 Astronomical Society of the Pacific Conference Series,  
 1454 Vol. 527, Astronomical Data Analysis Software and  
 1455 Systems XXIX, ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de  
 1456 Plaa, & H. Verkouter, 461
- 1457 Bigourdan, G. 1888, Bulletin Astronomique, Serie I, 5, 303
- 1458 Cramer, H. 1946, Mathematical methods of statistics  
 1459 (Princeton University Press)
- 1460 D'Agostino, R. B. 1970, Biometrika, 57, 679,  
 1461 doi: [10.1093/biomet/57.3.679](https://doi.org/10.1093/biomet/57.3.679)
- 1462 D'Agostino, R. B., & Belanger, A. 1990, The American  
 1463 Statistician, 44, 316.  
<http://www.jstor.org/stable/2684359>
- 1465 de Vaucouleurs, G. 1948, Annales d'Astrophysique, 11, 247
- 1466 Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168,  
 1467 doi: [10.3847/1538-3881/ab089d](https://doi.org/10.3847/1538-3881/ab089d)
- 1468 Ding, X., Birrer, S., Treu, T., & Silverman, J. D. 2021,  
 1469 arXiv e-prints, arXiv:2111.08721.  
<https://arxiv.org/abs/2111.08721>
- 1471 Faisst, A. L., Chary, R. R., Fajardo-Acosta, S., et al. 2021,  
 1472 arXiv e-prints, arXiv:2103.09836.  
<https://arxiv.org/abs/2103.09836>
- 1474 Ferrari, F., de Carvalho, R. R., & Trevisan, M. 2015, ApJ,  
 1475 814, 55, doi: [10.1088/0004-637X/814/1/55](https://doi.org/10.1088/0004-637X/814/1/55)
- 1476 Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, A&A,  
 1477 634, A135, doi: [10.1051/0004-6361/201936329](https://doi.org/10.1051/0004-6361/201936329)
- 1478 Haigh, C., Chamba, N., Venhola, A., et al. 2021, A&A, 645,  
 1479 A107, doi: [10.1051/0004-6361/201936561](https://doi.org/10.1051/0004-6361/201936561)
- 1480 Häußler, B., Vika, M., Bamford, S. P., et al. 2022, arXiv  
 1481 e-prints, arXiv:2204.05907.  
<https://arxiv.org/abs/2204.05907>
- 1483 Holwerda, B. W. 2005, arXiv e-prints, astro,  
 1484 doi: [10.48550/arXiv.astro-ph/0512139](https://doi.org/10.48550/arXiv.astro-ph/0512139)
- 1485 Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, The  
 1486 Astrophysical Journal Supplement Series, 203, 23,  
 1487 doi: [10.1088/0067-0049/203/2/23](https://doi.org/10.1088/0067-0049/203/2/23)
- 1488 Hubble, E. 1929, Proceedings of the National Academy of  
 1489 Science, 15, 168, doi: [10.1073/pnas.15.3.168](https://doi.org/10.1073/pnas.15.3.168)
- 1490 Hubble, E. P. 1926, ApJ, 64, 321, doi: [10.1086/143018](https://doi.org/10.1086/143018)
- 1491 Hunter, J. D. 2007, Computing in Science Engineering, 9,  
 1492 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- 1493 Jin, S., Daddi, E., Magdis, G. E., et al. 2022, arXiv e-prints,  
 1494 arXiv:2206.10401. <https://arxiv.org/abs/2206.10401>
- 1495 Kauffmann, O. B., Ilbert, O., Weaver, J. R., et al. 2022,  
 1496 A&A, 667, A65, doi: [10.1051/0004-6361/202243088](https://doi.org/10.1051/0004-6361/202243088)
- 1497 Kokorev, V., Brammer, G., Fujimoto, S., et al. 2022, ApJS,  
 1498 263, 38, doi: [10.3847/1538-4365/ac9909](https://doi.org/10.3847/1538-4365/ac9909)
- 1499 Kümmel, M., Bertin, E., Schefer, M., et al. 2020, in  
 1500 Astronomical Society of the Pacific Conference Series,  
 1501 Vol. 527, Working With the SourceXtractor++ Software,  
 1502 ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de Plaa, &  
 1503 H. Verkouter, 29
- 1504 Laidler, V. G., Papovich, C., Grogin, N. A., et al. 2007,  
 1505 PASP, 119, 1325, doi: [10.1086/523898](https://doi.org/10.1086/523898)
- 1506 Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS,  
 1507 224, 24, doi: [10.3847/0067-0049/224/2/24](https://doi.org/10.3847/0067-0049/224/2/24)
- 1508 Lang, D., Hogg, D. W., & Mykytyn, D. 2016, Astrophysics  
 1509 Source Code Library, ascl:1604.008.  
<http://adsabs.harvard.edu/abs/2016ascl.soft04008L>
- 1511 Lang, D., Hogg, D. W., & Schlegel, D. J. 2016a, AJ, 151,  
 1512 36, doi: [10.3847/0004-6256/151/2/36](https://doi.org/10.3847/0004-6256/151/2/36)
- 1513 —. 2016b, AJ, 151, 36, doi: [10.3847/0004-6256/151/2/36](https://doi.org/10.3847/0004-6256/151/2/36)
- 1514 Lesser, M. 2015, Publications of the Astronomical Society  
 1515 of the Pacific, 127, 1097.  
<http://www.jstor.org/stable/10.1086/684054>
- 1516 Mancone, C. L., Gonzalez, A. H., Moustakas, L. A., &  
 1517 Price, A. 2013, PASP, 125, 1514, doi: [10.1086/674431](https://doi.org/10.1086/674431)
- 1519 McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al.  
 1520 2012, Astronomy and Astrophysics, 544, A156,  
 1521 doi: [10.1051/0004-6361/201219507](https://doi.org/10.1051/0004-6361/201219507)
- 1522 Merlin, E., Fontana, A., Ferguson, H. C., et al. 2015, A&A,  
 1523 582, A15, doi: [10.1051/0004-6361/201526471](https://doi.org/10.1051/0004-6361/201526471)
- 1524 Merlin, E., Bourne, N., Castellano, M., et al. 2016, A&A,  
 1525 595, A97, doi: [10.1051/0004-6361/201628751](https://doi.org/10.1051/0004-6361/201628751)
- 1526 Nyland, K., Lacy, M., Sajina, A., et al. 2017, ApJS, 230, 9,  
 1527 doi: [10.3847/1538-4365/aa6fed](https://doi.org/10.3847/1538-4365/aa6fed)
- 1528 Oke, J. B. 1974, ApJS, 27, 21, doi: [10.1086/190287](https://doi.org/10.1086/190287)
- 1529 Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002,  
 1530 AJ, 124, 266, doi: [10.1086/340952](https://doi.org/10.1086/340952)
- 1531 —. 2010, AJ, 139, 2097, doi: [10.1088/0004-6256/139/6/2097](https://doi.org/10.1088/0004-6256/139/6/2097)

- 1532 Portillo, S. K. N., Speagle, J. S., & Finkbeiner, D. P. 2020,  
1533 AJ, 159, 165, doi: [10.3847/1538-3881/ab76ba](https://doi.org/10.3847/1538-3881/ab76ba)
- 1534 Rao, C. R. 1945, in Bulletin of the Calcutta Mathematical  
1535 Society, Vol. 37, 81–89
- 1536 Robotham, A. S. G., Taranu, D. S., Tobar, R., Moffett, A.,  
1537 & Driver, S. P. 2017, MNRAS, 466, 1513,  
1538 doi: [10.1093/mnras/stw3039](https://doi.org/10.1093/mnras/stw3039)
- 1539 Rowe, B. T. P., Jarvis, M., Mandelbaum, R., et al. 2015,  
1540 Astronomy and Computing, 10, 121,  
1541 doi: [10.1016/j.ascom.2015.02.002](https://doi.org/10.1016/j.ascom.2015.02.002)
- 1542 Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de  
1543 Astronomia La Plata Argentina, 6, 41
- 1544 Stetson, P. B. 1987, PASP, 99, 191, doi: [10.1086/131977](https://doi.org/10.1086/131977)
- 1545 Stevans, M. L., Finkelstein, S. L., Kawinwanichakij, L.,  
1546 et al. 2021, ApJ, 921, 58, doi: [10.3847/1538-4357/ac0cf6](https://doi.org/10.3847/1538-4357/ac0cf6)
- 1547 Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002,  
1548 AJ, 123, 485, doi: [10.1086/324741](https://doi.org/10.1086/324741)
- 1549 Szalay, A. S., Connolly, A. J., & Szokoly, G. P. 1999, AJ,  
1550 117, 68
- 1551 van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011,  
1552 Computing in Science Engineering, 13, 22,  
1553 doi: [10.1109/MCSE.2011.37](https://doi.org/10.1109/MCSE.2011.37)
- 1554 Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022,  
1555 ApJS, 258, 11, doi: [10.3847/1538-4365/ac3078](https://doi.org/10.3847/1538-4365/ac3078)
- 1556 Whitaker, K. E., Labb  , I., van Dokkum, P. G., et al. 2011,  
1557 ApJ, 735, 86, doi: [10.1088/0004-637X/735/2/86](https://doi.org/10.1088/0004-637X/735/2/86)
- 1558 Wright, A. H., Robotham, A. S. G., Bourne, N., et al. 2016,  
1559 MNRAS, 460, 765, doi: [10.1093/mnras/stw832](https://doi.org/10.1093/mnras/stw832)

1560

## APPENDIX

1561     A. CONSIDERATIONS, ASSUMPTIONS, AND  
1562        LIMITATIONS

1563     Although **The Farmer** effectively extends the func-  
1564     tionality of **The Tractor** to include source detection  
1565     and grouping, model assignments, catalog creation, and  
1566     computational efficiency, these advantages come with  
1567     considerable limitations that are discussed below.

1568     A.1. *Image Preparation & Source Grouping*

1569     Several aspects of the image preparation and group  
1570     identification stages are unique to **The Farmer**.

1571     *How should one determine how many bricks should a  
1572     mosaic be broken into?* This is primarily a computa-  
1573     tional concern. We can understand why by considering  
1574     the combined perimeter of all the bricks in a mosaic of  
1575     a fixed size; the perimeter is large when bricks are small  
1576     and vice versa. The larger the perimeter, the greater the  
1577     chance that the brick will split across a group of sources  
1578     which should be ideally modelled simultaneously in the  
1579     same brick. In general this should be avoided, and so  
1580     bricks should be made as large as possible. It is possible  
1581     for each brick to be operated on by **The Farmer** inde-  
1582     pendently, which means they can be parallelized across  
1583     computational nodes or even processed by different com-  
1584     puting facilities altogether. The brick also needs to be  
1585     read into memory at runtime, and so should be sized  
1586     appropriately for the memory capacity of a given com-  
1587     putational facility. Bricks from forced photometry are  
1588     typically the largest files as they contain all bands of in-  
1589     terest, their weight maps, and masks and so can become  
1590     tens of gigabytes for even modest dimensions.

1591     *What about sources near the edges of bricks that ex-  
1592     tend into the next brick?* It is up to the user to de-  
1593     termine how large the brick overlap should be. In gen-  
1594     eral, the overlap should be large enough that the largest  
1595     sources of interest, placed at the brick edge, would not  
1596     extend beyond the overlap. While one can set a large  
1597     brick overlap, doing so comes at the cost of memory  
1598     and computational overhead. Excessive brick overlaps  
1599     should be avoided where possible.

1600     *How should one assess if groups are correctly iden-  
1601     tified?* As discussed in Section 3.3, groups of sources  
1602     are identified by joining source segments that have been  
1603     dilated by some morphological structure whose size dic-  
1604     tates the extent of the dilation. The segments are con-  
1605     structed from the detection stage, and so one should only  
1606     consider the detection image when assessing the iden-  
1607     tification of groups. Dilation is necessary because the  
1608     segmentation extent in **Source Extractor** (and **SEP**) is

1609     tied to the significance level set for the detection. In  
1610     some cases the segments may not capture the full ex-  
1611     tent of two neighboring sources such that they should  
1612     be simultaneously modelled, but their segments do not  
1613     touch. Hence the size of the dilation structure should be  
1614     set so that these kinds of nearby sources are correctly  
1615     assigned into one group. This is most easily assessed by  
1616     inspection, and tuned in successive trials.

1617     It is important to note that morphological dilation can  
1618     destroy segments nearby to larger ones. For this reason  
1619     the dilation is carried out on a copy of the segmenta-  
1620     tion image which has been made binary such that pixels  
1621     assigned to the background are set to zero, and those  
1622     active pixels assigned to sources are set to 1. Segments  
1623     which are already touching are now indistinguishable,  
1624     and the dilation simply enlarges the footprint of con-  
1625     tiguous regions of active pixels. The small segments re-  
1626     main identifiable from the segmentation image. This is  
1627     important, because the group pixels belong to the group  
1628     itself; no one group pixel belongs to a single source.  
1629     That ownership is retained in the original segmentation  
1630     only. This is essential because while the joint likelihood  
1631     maximized by **The Tractor** is computed over the group  
1632     pixels, **The Farmer** can still judge the fit of individual  
1633     sources from the  $\chi^2$  computed over their uniquely owned  
1634     pixels.

1635     In some cases the segments produced at detection may  
1636     be too large and so over-group sources. While this is not  
1637     a problem scientifically, it increases the computational  
1638     complexity of the fit which can lead to poor model per-  
1639     formance, or worse, the joint model may even fail to  
1640     converge altogether. However, unlike morphological di-  
1641     lation which cannot destroy groups of pixels, morpho-  
1642     logical erosion can destroy the smallest segments typi-  
1643     cally containing one source. This is a limit that must be  
1644     avoided in order for **The Farmer**'s decision tree to work.  
1645     More work is required to address this case.

1646     A limitation of this approach is that groups are defined  
1647     based on the detection image, its effective resolution,  
1648     as well as the depths and properties of its constituent  
1649     bands. A group determined from well-resolved optical  
1650     images will likely miss pixels with significant flux when  
1651     applied to sources in drastically lower resolution images.  
1652     This can be overcome by further dilating each group on  
1653     a band-by-band basis such that all of the relevant pixels  
1654     are now constraints on the model. This comes with a  
1655     dilemma, however, as crowding is worse in low resolu-  
1656     tion images of the same depth and so light from sources  
1657     not originally included in the group may now contribute.

1658 Yet these nearby sources are not described by the existing group model, and so leaving their flux unaccounted  
 1659 for may instigate a bias in the photometry. The only  
 1660 tractable option seems to be to join these groups and  
 1661 perform the forced photometry in a simultaneous op-  
 1662 timization. However, the shapes of these models were  
 1663 never determined together, and so it is uncertain how  
 1664 well the new group of models would perform. Worse,  
 1665 most sources in the deepest IRAC images are blended to  
 1666 some degree and so strictly keeping to this philosophy  
 1667 of joint optimization of all overlapping sources would  
 1668 require every source to be simultaneously fit. **While**  
 1669 **not impossible** (e.g. [Lang et al. 2016b](#)), it is po-  
 1670 **tentially computationally expensive.** Alternative  
 1671 **strategies will be explored in future work.**

1672 *What sets the overlap sizes for groups?* Although  
 1673 groups of sources are limited to their footprint whose  
 1674 pixels are identified by dilating source segments, the  
 1675 groups themselves are saved in memory as rectangular  
 1676 arrays whose dimensions are set by the maximum ex-  
 1677 tent of the group footprint. Although pixels inside of  
 1678 the rectangular array but outside the group footprint  
 1679 (which can often be fractal-like in shape) do not provide  
 1680 any constraining power as their weight is set to zero,  
 1681 the models are still realized onto the larger array dur-  
 1682 ing the optimization. It is generally best if these models  
 1683 are not truncated whatsoever, and so **The Farmer** en-  
 1684 larges the dimensions of the group array by a set number  
 1685 of pixels. This is not only for numerical reasons inter-  
 1686 nal to **The Tractor**, but also is a requirement if post-  
 1687 processing apertures are to measure the full extent of  
 1688 the joint model image. Truncation of that joint model  
 1689 will mean that the wings will not be realized and so  
 1690 the largest apertures will underestimate the true flux.  
 1691 Thankfully, if the models are correctly normalized then  
 1692 the truncation will *not* affect the best-fit normalization  
 1693 coefficient from which the source flux from **The Tractor**  
 1694 is derived. Nonetheless it is advisable that the group ar-  
 1695 ray size is large enough so that the PSF stamp would  
 1696 not be truncated for a source placed near the edge of  
 1697 main group footprint.

#### 1699 A.2. Selection Functions and Image Depth

1700 One must be cognisant regarding which band should  
 1701 be used to determine the models and their best-fit pa-  
 1702 rameters. In fact, this is not a free choice. Using a  
 1703 band outside of those used in the detection image is in-  
 1704 appropriate because there may be sources identified in  
 1705 the detection image which do not have flux in the cho-  
 1706 sen modelling image. If one is to maintain the selection  
 1707 function constructed by the detection strategy, then it  
 1708 must be guaranteed that a detected source has sufficient

1709 signal to constrain its model. Otherwise sources with-  
 1710 out models cannot be photometered, and so the selec-  
 1711 tion function changes in a non-trivial way. For the same  
 1712 reason, it is also inadvisable to use only one band of a  
 1713 multi-band detection image, or even the bands which de-  
 1714 fine the spectral domain of the detection image. Nor is  
 1715 it advisable to attempt to model sources in a co-added  
 1716 image as the effective PSF is not easily characterized,  
 1717 and the FWHM of the constituent PSFs can produce  
 1718 additional variation in the surface brightness profiles.  
 1719 Therefore, it is strongly recommended that the models  
 1720 be produced from precisely the same bands and images  
 1721 that were used or combined to make the detection im-  
 1722 age.

1723 Measuring photometry of a source in an image which  
 1724 contains additional sources outside the selection func-  
 1725 tion (by virtue of not being detected) presents another  
 1726 often encountered dilemma, although common also to  
 1727 aperture-based methods. This is because signal from  
 1728 an additional, undetected source is not described by the  
 1729 set of models assigned to a group. For instance, a red  
 1730 source which is undetected in a predominantly blue se-  
 1731 lection function may in a red band appear next to a  
 1732 known blue source. Although fixing model shapes helps  
 1733 avoid contamination, it is possible that the likelihood  
 1734 will be maximized by increasing the flux parameter of  
 1735 the blue model such that some of the flux from the new,  
 1736 red source is inadvertently accounted for, thus biasing  
 1737 the photometry for the blue source in that red band. Of-  
 1738 ten times these cases can be identified afterwards from  
 1739 diagnostics provided by **The Farmer**, although not guar-  
 1740 anteed.

1741 A similar situation is encountered when forcing pho-  
 1742 tometry onto deeper bands of the same wavelength as  
 1743 the detection, and although such images typically can  
 1744 provide better photometric constraints, they may at the  
 1745 same time introduce bias by introducing new, unde-  
 1746 tected sources. This means that ideally all sources in  
 1747 a forced photometry image should be modelled, which  
 1748 requires that they were detected. However, identifying  
 1749 these new sources automatically ahead of photometry  
 1750 is not practical as lists of detected sources will differ  
 1751 due the blends; the two catalogs must then somehow  
 1752 be reconciled and segmentation maps merged. Doing so  
 1753 in limited numbers is possible with careful supervision,  
 1754 typically with the assumption that new sources are un-  
 1755 resolved to avoid re-processing the decision tree (as used  
 1756 to photometer optically dark sources in [Jin et al. 2022](#)).  
 1757 This potentially pathological issue will be addressed in  
 1758 future work.

1759 A.3. *Models, Morphological Corrections, and Drifting*

1760 One significant complication with The Farmer is that  
 1761 the decision tree needs to be tuned. Because the cen-  
 1762 tral operation of the decision tree is to separate resolved  
 1763 and unresolved sources, its parameters are most sensi-  
 1764 tive to the resolution of the image. Subtended size corre-  
 1765 lates strongly with apparent brightness, and so sources  
 1766 in deep images typically become unresolved around a  
 1767 certain magnitude threshold. In order to succeed, the  
 1768 decision tree needs to be tuned such that it correctly as-  
 1769 signs unresolved models to essentially all sources fainter  
 1770 than this limit, in addition to bright point-sources. A  
 1771 photometric bias can develop if instead the decision  
 1772 tree assigns resolved models to unresolved sources, or  
 1773 vice versa. This can be readily diagnosed from num-  
 1774 ber counts which should be smooth and increase mono-  
 1775 tonically with decreasing brightness. If the decision  
 1776 tree is not providing adequate model type assignments,  
 1777 the number counts of the detection bands will either  
 1778 contract towards a sharp rise or flatten into a plateau  
 1779 around the resolution threshold. An unresolved model  
 1780 assigned to a resolved source tends to produce an un-  
 1781 derestimated flux, thereby moving these typically bright  
 1782 but comparatively rare sources towards fainter magni-  
 1783 tudes thus creating a plateau. In this case it is likely that  
 1784 the decision tree poorly tuned such that PointSource  
 1785 models are too easily assigned, and so the  $\chi^2$  penalty to  
 1786 the PointSource models should be lowered. A resolved  
 1787 model assigned to an unresolved source tends to pro-  
 1788 duce overestimated flux, thereby moving these typically  
 1789 faint but abundant sources towards brighter magnitudes  
 1790 thus creating a sharp rise in counts. In this case the  $\chi^2$   
 1791 penalty of PointSource models should be increased so  
 1792 that it is easier for sources to be assigned an unresolved  
 1793 model. Number counts are not as sensitive to which  
 1794 resolved model is assigned to a resolved source (e.g.,  
 1795 ExpGalaxy or DevGalaxy) and so the corresponding pa-  
 1796 rameters are most easily tuned by examining residuals  
 1797 of bright, resolved sources.

1798 It may not be possible to assign a simple paramet-  
 1799 ric model to a particular source. It might be that the  
 1800 source is actually two blended together. Meanwhile, the  
 1801 brightest sources tend to be resolved and have features  
 1802 such as spiral arms, bars, and starbursts that are not  
 1803 described by the smooth models from The Tractor. As  
 1804 such, model performance tends to decrease for bright,  
 1805 resolved sources (e.g. spirals). **While the presence**  
**1806 of poorly fit morphological features will be indi-**  
**1807 cated in the residual statistics (e.g.  $\chi^2$ ), the asso-**  
**1808 ciated photometry will likely be biased in some**  
**1809 way. This is especially true for space-based imag-**  
**1810 ing (e.g., HST) where the space spanned by mod-**

1811 **els from The Tractor are divorced from the real**  
 1812 **space spanned by highly resolved galaxies. While**  
 1813 **aperture photometry should be less biased, they**  
 1814 **are unable to inform about the presence of mor-**  
 1815 **phological features.**

1816 Chromatic changes in morphology presents a chal-  
 1817 lenge for The Farmer. The model for a given source  
 1818 during the modelling stage may be simultaneously con-  
 1819 strained by multiple bands, but The Tractor allows  
 1820 only one shape shared between the bands. Therefore  
 1821 the shapes reported by The Farmer from the modelling  
 1822 stage are most appropriate for the modelling bands with  
 1823 the largest weights. Forced photometry in regular opera-  
 1824 tion proceeds by only allowing the flux to vary with the  
 1825 shape fixed, meaning that changes in morphology are  
 1826 not accounted for by the model. However, The Farmer  
 1827 makes it possible to perform forced photometry on each  
 1828 band separately so that the shape parameters can be al-  
 1829 lowed to vary in each case with or without priors, albeit  
 1830 at greater computational expense and runs the danger  
 1831 of over-fitting. It is important to note that forcing mod-  
 1832 els derived from well-resolved bands onto images of lower  
 1833 resolution is typically successful as the larger PSF of the  
 1834 forced photometry band makes the photometric mea-  
 1835 surement less sensitive to morphology. However, forcing  
 1836 models derived from low resolution images onto bands at  
 1837 high spatial resolution typically results in a poor per-  
 1838 formance as the band of interest contains more information  
 1839 than the model can describe.

1840 As discussed in Portillo et al. (2020), flux and shape  
 1841 estimates can suffer from biases introduced from in-  
 1842 adequate centroiding. Given the great number of  
 1843 multi-wavelength images and facilities involved in mod-  
 1844 ern surveys, even small astrometric offsets can impact  
 1845 the measurements derived from model fitting. Hence,  
 1846 The Farmer allows the user to unfix the centroid pos-  
 1847 ition of each model and introduces a Gaussian prior on  
 1848 its position, on a band-by-band basis. This prior acts  
 1849 to penalize the likelihood of the model fit if the model  
 1850 obtains a centroid that is beyond the distance set by  
 1851 the prior (i.e. ‘drifts’). This drifting can be especially  
 1852 prevalent in the case of a known faint source next to  
 1853 an undetected bright neighbor which because it is not  
 1854 accounted for by the model will cause the model of the  
 1855 faint source to move towards the bright source, whose  
 1856 unaddressed presence counts against the likelihood more  
 1857 than the original, fainter source. Priors can be set on the  
 1858 position, although their widths are usually determined  
 1859 by successive trials.

1860 It is important to appreciate that the grouping of  
 1861 sources imparts a significant advantage over fitting in-  
 1862 dividual sources. Because groups of sources are pho-

tometered separately from other groups, a failure of the model in one group does not affect any other group. Let us consider the unfortunate example in which a galaxy is assigned an inaccurate model whose large axis ratio results in wings extending well beyond the group. While those wings will be a problem for the source in question, and perhaps its group members as well, they will not affect any other group in the image. Hence, while this is an issue for the residual map, there is no reason to be concerned about the photometry of the other group as they were fitted in an entirely separate optimization in isolation of the problematic source. However, this advantage effectively decouples the reconstructed brick-level residual image from the photometry and so complicates searches for sources in residuals. As mentioned in Section 3.1, The Farmer has built-in functions to filter out these problematic models.

#### A.4. Source De-blending

While profile-fitting photometry can be used to deblend two sources, they first must be identified as separate objects, which in turn depends on the original source detection. As such, de-blending sources at the detection stage is not a problem which profile-fitting photometry can (or should) solve and instead is well-suited to address the related, but distinct issue of accurately measuring the flux of two *identified* but blended sources. It is essential, therefore, to understand that if two nearby sources are not successfully de-blended at detection, then profile-fitting techniques should not be expected to reliably de-blend them either.

This concept is demonstrated in Figure 17 whereby a point source is placed in the vicinity of a bright resolved galaxy and appropriate models are assumed to be known. Several cases are set up by varying their distance and relative brightness. Attempting to photometer both of them with only one model produces expectedly poor fits in several cases. The system is then evolved by allowing it to subtract the first source, find the brightest residual source, and re-fit the original using two centroids which in turn improves the performance in cases where the residual source can be identified. However, in practice this is dangerous if one does not know beforehand whether there is another source or whether the model for the one source was simply a poor fit. Lastly, the two sources are fit by two appropriate models which results in accurate photometry at all distances and relative brightness. Hence, de-blending with profile-fitting photometry requires the correct number of models (and centroids) to optimize for a given group of sources.

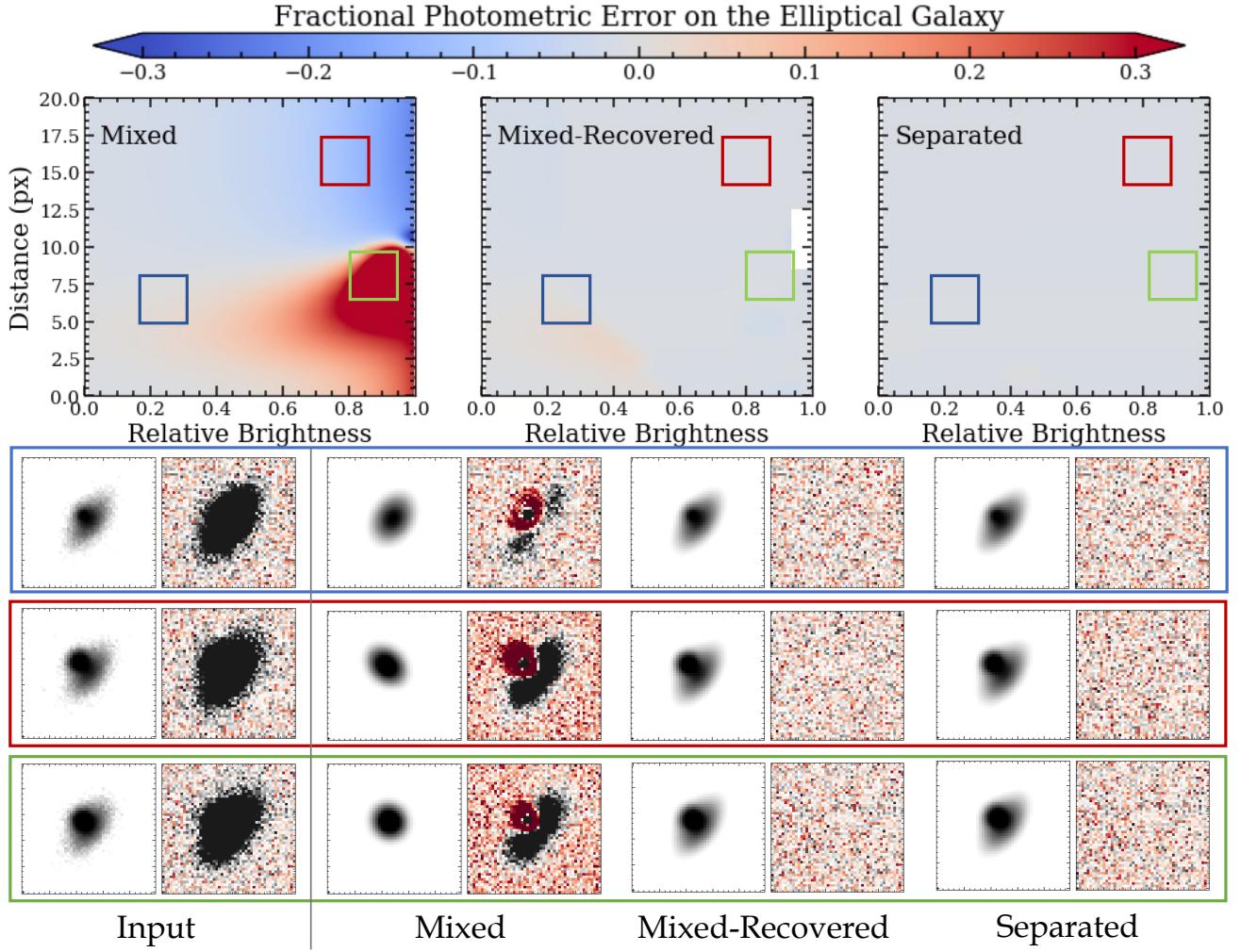
What is little appreciated, however, is that this behavior is undoubtedly an advantage. While aperture

techniques do not make any assumption about source morphology and are hence extremely powerful in the face of resolved structural features in galaxies, their ability to identify cases of sources blended at detection or quantify contamination in photometry of photometrically blended sources is severely limited. Parametric profile-fitting techniques suffer from neither of these drawbacks. So long as intrinsically blended sources are not well-described also by a single profile<sup>10</sup>, then the optimized model will not achieve a satisfactory fit. These cases may be confidently identified a posteriori using statistics such as those discussed in Section 3.4.2.

#### A.5. Comparison to Similar Methods

These advantages and limitations hold mainly for purely parametric models. There exists another, related class which uses a high resolution cutout of a given source as its model that can be used to photometer other bands by first convolving it with an appropriate kernel to translate its native PSF to that of the band of interest, and then scaling the unit normalized model to match the source in that band. These ‘stamps’ have a distinct advantage over parametric models in that they can exploit the resolution of the cutout image to capture structural features not describable by typically smooth parametric models. Without shape parameters to constrain, these stamps can be extremely efficient in measuring fluxes as essentially a scaling factor between the PSF-transformed stamp and the source in question. While simpler than purely parametric models, this approach requires a deep high-resolution image which contains every detected source (if not the same image as that used for detection). More so, the PSF must be well-understood to provide kernels to map the original PSF to those of the lower resolution images, a drawback not shared by parametric models. The stamp must also be resampled to match the pixel scale of the image to be photometered. For example, an *HST*-derived stamp of a marginally resolved source applied to *Spitzer* provides no significant advantage over a parameteric model. Worse, the morphology described by the stamp is assumed to be constant, and so there can arise significant effects between the wavelength of the stamp image and that of the image to be photometered. Such stamp-based profile-fitting software include TFIT (Lai-dler et al. 2007), T-PHOT (Merlin et al. 2015, 2016), PyGFIT (Mancone et al. 2013), Morfometryka (Ferrari et al. 2015), LAMBNDAR (Wright et al. 2016), and GOLFIR (Kokorev et al. 2022). Each one takes its own approach

<sup>10</sup> In such cases sources cannot be identified as separate objects without higher resolution ancillary data anyways.



**Figure 17.** A point source is simulated in the vicinity of a large, central elliptical galaxy. Models are fitted for three cases: the sources are blended and have only one centroid (Mixed), the sources are blended, modelled, and then the missing source is recovered from the residual image and modelled (Mixed-Recovered), and lastly both sources are a priori known and simultaneously fit (separated). Each measurement is repeated over a grid of relative brightness (0 meaning that the point source is negligible) and the distance from the central elliptical to the point source (20 means the point source is in the top left corner). While grey areas indicate successful recovery of the input (true) flux of the elliptical source, red areas indicate that the flux of the elliptical is underestimated. White areas in the middle panel indicate where the point source is not detected in the mixed residual. The model and residual of three situations are shown for each measurement strategy.

to the problem of flux estimation in terms of available models, parametrization, algorithm speed, flexibility, and accessibility. As discussed in Section A.3, purely parametric models can overcome the limitations of these stamp-based codes by freely fitting the shape of the model, possibly with some prior constraints (e.g., GALFIT, Peng et al. 2002, 2010, ProFit, Robotham et al. 2017; SExtractor++, Bertin et al. 2020; Kümmel et al. 2020; GALAPAGOS-2, Häußler et al. 2022).

One of the most similar photometry frameworks to The Farmer is HSCPipe, in part because they both inherit the profile-fitting approach of SDSS (Stoughton et al. 2002). As discussed in Aihara et al. (2019),

HSCPipe provides model-based photometry by fitting both point-source (PSF) and composite galaxy (cModel) profiles to each galaxy individually. Even though both resolved and point-like models are tried, unlike The Farmer they are tried for each source independent of their neighbors, which for blended sources can lead to inconsistencies (as demonstrated in Figure 1). Furthermore, HSCPipe does not choose a best-fit model type for each source and instead provides fluxes measured from each profile assuming independence from neighbors. While this is computationally faster than a decision tree, it is also inefficient to fit unresolved sources with highly parameterized composite models (which risk

overfitting). As of version 8 of `HSCPipe`<sup>11</sup>, only likelihood of the CModel fits are reported and so a consistent statistical comparison with the PSF models is not possible, which leaves only a binary extendedness flag to indicate a resolved source. `The Farmer` provides not only a best-fit model type for each source, but also suite of statistics from which the reliability of that model can be assessed.

Although limited to low resolution IRAC images, the `IRACLEAN` software (Hsieh et al. 2012) measures photometry by iteratively subtracting PSFs at detected source centroids until the residual is clean of signal to some user defined level. Although broadly similar to `The Farmer`, `IRACLEAN` does not perform model-fitting in a classical sense as an unbounded iterative subtraction of the PSF is equivalent to a model with effectively unlimited parameters. Furthermore, the order in which sources are processed can introduce hysteresis in crowded regions. There is also the danger of overfitting, as `IRACLEAN` will continue subtracting a scaled PSF stamp until a given segment has no more signal, which in the case of a blend will combine the flux of the two sources into one photometric measurement. `The Farmer`'s parametric models act as a prior which can, in some cases, ignore the flux of a neighbor which is left in the residual, and report statistics flagging the problem to the user. Further discussions and comparisons with `IRACLEAN` are presented in Weaver et al. (2022).

#### A.6. Computational Considerations

Computation of sources scales with the number of sources fit simultaneously as well as the number of free parameters, meaning that these techniques require significantly longer runtimes compared to aperture photometry. In the context of modern deep surveys containing millions of sources (many of which are resolved), fitting all sources simultaneously would be enormously complex requiring significant computational resources. However, a high degree of parallelization can be achieved so long as the source density and resolution allow for distinct groups of sources to be identified and fit separately. A practical approach is to process each brick independently. Source groups are constructed at runtime and kept in memory only, so they are ideal for being run in parallel, e.g. across many CPUs of a given cluster node. However, computational time still increases with the number of free parameters. As such, the modelling stage is not only more complicated than forced photometry because of the several trials of the decision tree, but also

because shapes are left to vary in some stages. It is for this reason that the decision tree starts with simple models and moves towards complexity, or in other words, computational expense. If the conditions of the decision tree are satisfied for every source, then the models are assigned without moving to the next stage. For example, an isolated point source should only be tried out with a `PointSource` and `SimpleGalaxy` model whereupon it should satisfy the `PointSource` criterion and stop. Each of these model types have three parameters (two for position and one for flux) and so are incredibly quick compared to a `CompositeGalaxy` with ten parameters.

Unfortunately, computational time also increases with source crowding given the stronger covariance between neighboring models. Many separate sources can be modelled independently and in parallel without a loss of accuracy. However, because deep images of crowded fields are best photometered when groups of nearby sources are simultaneously modelled, the complexity and computational expense is greater than if the same number of sources were fit separately. As such, it is strongly advised that typical source groups contain as few members as possible without breaking across two blended sources. Unfortunately, the situation of source crowding will only become more difficult as surveys grow deeper. While apertures will eventually hit a limit, profile-fitting photometry can forge ahead, albeit with a greater computational cost.

<sup>11</sup> [https://hsc.mtk.nao.ac.jp/pipedoc/pipedoc\\_8\\_e/index.html](https://hsc.mtk.nao.ac.jp/pipedoc/pipedoc_8_e/index.html)