Appendix A: Determination of the galaxies' structural parameters

Appendix A.1: Method used by Bruce et al. (2012)

Bruce et al. (2012) study the morphological properties of a sample of 215 galaxies at $1 < z < 3$ in the CANDELS-UDS field with the 2D morphology fitting code `galfit` (Peng et al. 2002, 2010) from WFC3/IR H_{160}-band images (Grogin et al. 2011; Koekemoer et al. 2011). They first run `Sextractor` both to determine initial estimates of the position, total magnitude, axis ratio and effective radius of the galaxies and to produce a segmentation map, before carrying out single-Sérsic fits. As in Häussler et al. (2007), they show that the `galfit` output depends strongly on the background subtraction in the input image, the weight map and the PSF. They notably compare the modelled TinyTim PSF (Krist 1995) with en empirical PSF constructed from a median stak of seven stars and find that the TinyTim PSF overpredicts sizes by 5-10% compared to the empirical PSF for the single Sérsic fits. Their weight image is first taken from the CANDELS rms map but then adapted to account for the Poisson noise contribution of the objects. For the two-component bulge disk fits, they lock the centers of both components to that obtained from the single Sérsic fit and impose $n = 1$ for the disk and $n = 4$ for the bulge. They further explore adding a point source at the center to account for a potential AGN or central star cluster. When the secondary component contributes to less than 10% of the total luminosity, they revert to the corresponding bulge- or disk-only model as the fainter component was often found to be unphysical. To avoid being trapped in a local χ^2 minimum, they construct a grid of 231 different starting values for the total magnitudes (11 steps) and effective radii (21 steps) and then adopt the output minimizing χ^2 as the best-fitting model.

Appendix A.2: Method used in Lang et al. (2014)

Lang et al. (2014) propose a slighly different systematic method to determine structural parameters of CANDELS/3D-HST galaxies from HST images, optimized for H_{160}-band two-component bulge disk modeling and the determination of the bulge-to-total luminosity ratio B/T. Their method also uses `galfit` and is divided in the following steps in order to mitigate the risk of being trapped in a local χ^2 minimum when doing the two-component fits:

1. **Preliminary study of the single-component Sérsic indices of noise-free two-component bulge disk systems.** As presented in their Appendix A, Lang et al. (2014) first carry out single Sérsic fits to idealized, noise-free bulge ($n = 4$) plus disk ($n = 1$) models with bulge-to-total luminosity ratio B/T between 0 to 1 and half-light radius ratio R_B/R_D between the two components from 0.1 to 1, both by steps of 0.1. This preliminary study is meant to infer initial guesses for the two-component fits from single Sérsic fits, which are precisely less sensitive to the initial guesses. For a given bulge disk system, the best-fit global Sérsic index increases both with increasing B/T and with decreasing size ratio R_B/R_D between bulge and disk half-light radii. Lang et al. (2014) further show that the Sérsic indices of the observed CANDELS galaxies are in good agreement with those expected from the idealized, noise free systems.

2. **The point spread function (PSF).** The PSF used by Lang et al. (2014) is a combination of stacked nearby stars and a TinyTim (Krist 1995) model, as in van der Wel et al. (2012).

3. **Masking out neighbording sources.** Lang et al. (2014) use an automated scheme to mask out neighboring sources and do initial guesses for `galfit` parameters, including size, total magnitude and center.

4. **Single Sérsic fits to the observed galaxies.** The parameters are the half-light radius R_{Sersic}, the Sérsic index n_{Sersic}, the galaxy position, integrated magnitude m_{Sersic}, axis ratio AR_{Sersic} and position angle PA_{Sersic}. The Sérsic index is only allowed to vary with the range $0.2 < n_{Sersic} < 8$.

5. **One-component pure disk and pure bulge fits to the galaxies.** The Sérsic indices are set to $n = 1$ and $n = 4$ for the disk and the bulge, respectively. The other parameters are left free as in the single Sérsic models.

6. **A series of two-component bulge disk fits to the galaxies.** As in the previous step, the Sérsic indices are set to be $n = 1$ for the disk and $n = 4$ for the bulge. The centers are left free but within 2 pixels of each other, other parameters being allowed to vary independently. Lang et al. (2014) carry out ten such runs with different initial guesses, each of them motivated by a certain value of the size ratio R_B/R_D between bulge and disk half-light radii: for each R_B/R_D taken between 0.1 to 1 by steps of 0.1, the initial guess for B/T is deduced from the single-component n_{Sersic} according to the empirial relation obtained during the preliminary study. Similarly, the initial magnitudes and absolute values of the sizes are set such that the total magnitude and half-light radius of the composite model match those of the single Sérsic fit.

7. **Selection of the best-fit model.** The best-fit model is that with the lowest reduced χ^2 between the different two-component fits, the pure disk fit and the pure bulge fit. Cases where GALFIT yields bulge that are too small ($R_B < 0.1$ pixel) or disks smaller than the bulge (i.e., $R_B/R_D > 1$) are cast away.

Appendix A.3: Method used in Contini et al. (2016)

To study 28 emission line galaxies observed with MUSE in the HDFS field, Contini et al. (2016) carry out two-component bulge disk fits from HST WFPC2 F814W I-band images with `galfit`. In their study, they use a circular Moffat model for the PSF, derived from stacking seven non-saturated stars, and consider a $n = 1$ disk with a $n = 4$ circular bulge (of axis ratio equal to unity). They

investigate the dependence of the results to various PSF models, finding a negligible difference between the TInyTim model and the stacked PSF (<2%), and to a free Sérsic index for the bulge but do not address the risk of being trapped in a local χ^2 minimum.

Appendix A.4: Method followed in this paper

The method followed in this paper to determine the galaxies' structural parameters and their B/T ratio builds upon those mentioned above, generally following that used by Lang et al. (2014) but with some improvements to compensate for `galfit`'s extreme sensitivity to the initial guesses and the subsequent degeneracy of the solutions when carrying two-components fits. Indeed, the main disadvantage of the least-square fitting algorithm that is `galfit` is that it can converge to a local χ^2 minimum instead of the global minimum (e.g., Häussler et al. 2007; Peng et al. 2010; Bruce et al. 2012). The improvements implemented here are enabled by our relatively small sample size compared to Lang et al. (2014) and include: (i) accounting for the most luminous nearby galaxies or satellites by fitting them simultaneously with a single Sérsic distribution; (ii) imposing a bulge axis ratio greater or equal to that of the disk and the same position angle for the two components; and (iii) adding initial guesses built from the single disk and bulge fits. These additional initial guesses are particularly important when the single Sérsic fits lead to high Sérsic indices not accounted for in the preliminary study of noise-free two-component bulge disk systems. We use a uniform weight map for simplicity, which is a good approximation of the weight map deduced from the HST rms map given the redshift of the sources but which does not account for potential Poisson noise.

1. **Empirical study of the single-component Sérsic indices of noise-free two-component bulge disk systems.** We reproduce the study carried out by Lang et al. (2014) in their Appendix A, i.e., we generate a grid of noise-free two-component bulge disk systems with B/T between 0 and 1 and R_b/R_d between 0.1 and 1 by steps of 0.1 and carry out single Sérsic fits without PSF with `galfit`, using weights corresponding to the Poisson noise. Fig. A.1 shows how the resulting single-component Sérsic index n_{Sersic} depend on B/T and R_b/R_d as in Appendix A of Lang et al. (2014). Similarly, Fig. A.2 shows how the single-component half-light radius R_{tot} depends on B/T and R_B/R_D. For both B/T and R_d, we derive 2D interpolations as a function of n_{Sersic} and R_b/R_f for the former and as a function of the resulting B/T and R_b/R_d for the latter. We use these functions to set initial guesses for the two-component fits. We also tried to use a decision tree regression instead of the 2D interpolation function, but we favored the smooth interpolation function instead of the discrete decision tree.

2. **The PSF.** As PSF, we use a TinyTim model corresponding to the ACS camera since Bruce et al. (2012) and Contini et al. (2016) show that is yields radii within 10% of those obtained with a stack of stars.

3. **Single Sérsic fits to the observed galaxies.** As in Lang et al. (2014), we only impose $0.2 < n_{\rm Sersic} < 8$. The initial guesses for the center, the magnitude and the radius of the single Sersic fits are taken from the ACS image for the center and otherwise taken to be $m_{Sersic} = -5$, $R_{Sersic} = 10$ pixels, $n_{Sersic} = 1$, $\rm AR_{Sersic} = 1$, $\rm PA_{Sersic} = 0$. As already noted by Lang et al. (2014), the single Sérsic fits are indeed rather insensitive to the initial guess.

4. **One-component pure disk and pure bulge fits to the galaxies**, with $n = 1$ and $n = 4$.

5. **A series of two-component bulge disk fits to the galaxies.** As in Lang et al. (2014), we impose $n = 1$ for the disk and $n = 4$ for the bulge and leave the centers free but within 2 pixels of each other. However, we additionally impose the position angles of both components to be equal and the axis ratio of the bulge to be higher or equal to that of the disk to avoid fitting potential bars separately. We do not impose any contraints on the radius ratio since such a constraint would potentially lead to converging to a false minimum that would not be physically motivated. As in Lang et al. (2014), we carry out ten runs with different initial guesses, each of them motivated by a certain size ratio R_B/R_D between 0.1 to 1 in steps of 0.1: for each value of R_B/R_D, we deduce B/T and R_D according to the interpolation functions of step 1. The initial guesses for the axis ratio and the position angle are taken from the single Sérsic fits. Noting that single Sérsic fits with $n_{\rm Sersic} > 5$ fall outside of the range covered by the empirical study of step 1, we also build initial guesses from the single bulge and single disk fits: on one side we carry out runs with B/T = 0.1 and 0.5 with the center, the disk radius and the total magnitude assumed to be those of the single disk fit; on the other size runs with B/T = 0.5 and 0.9 with the center, the bulge radius and the total magnitude assumed to be those of the single bulge fit.The initial axis ratio and position angle are those of the single Sérsic fit. In all cases, the initial radius ratio R_B/R_D is taken between 0.1 to 1 by steps of 0.1, hence leading to a total of 53 `galfit` runs per galaxy including the single Sérsic, single disk and single bulge fits instead of 13 in Lang et al. (2014). The addition of initial guesses stemming from the single disk and single bulge fits is further motivated by the fact that a high Sérsic index often correspond to an unphysically large half-light radius, which prevent the two components fits to converge when used as initial guess. In such cases, adding inital guesses stemming from the single bulge and single disk fits enables to start from a more physically meaningful half-light radius.

6. **Best-fit model.** As in Lang et al. (2014), we take the model with the lowest reduced χ^2 between the two-component fits and the pure disk and pure bulge fits. Cases where GALFIT yields implausibly small bulge sizes ($R_B < 0.1$ pixel) or disks smaller in size than the bulge (i.e., $R_B/R_D > 1$) are cast away, except for galaxy `GN4-36596`. Indeed, this galaxy appears to be a lenticular galaxy for which solutions with $R_B/R_D > 1$ are not satisfactory given the low surface brightness of the stellar halo in which the disk is embedded.

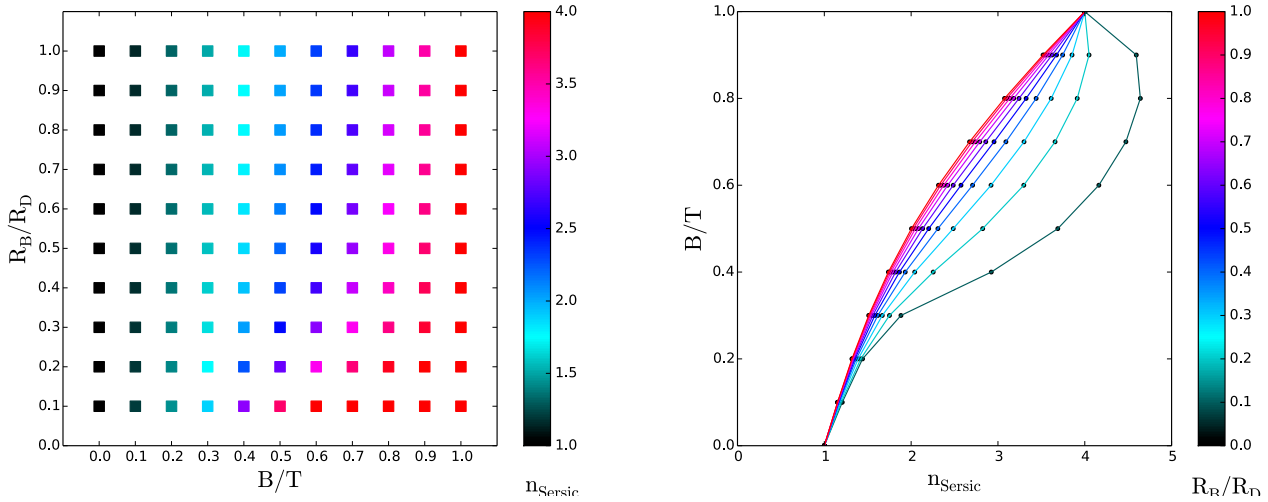

Fig. A.1: *Left:* Size ratio R_B/R_D between the bulge and disk component versus B/T together with the resulting single-component Sersic index n_{Sersic} inferred from noise-free idealized bulge disk models. *Right:* B/T as a function of n_{Sersic}, colors indicating the size ratio.

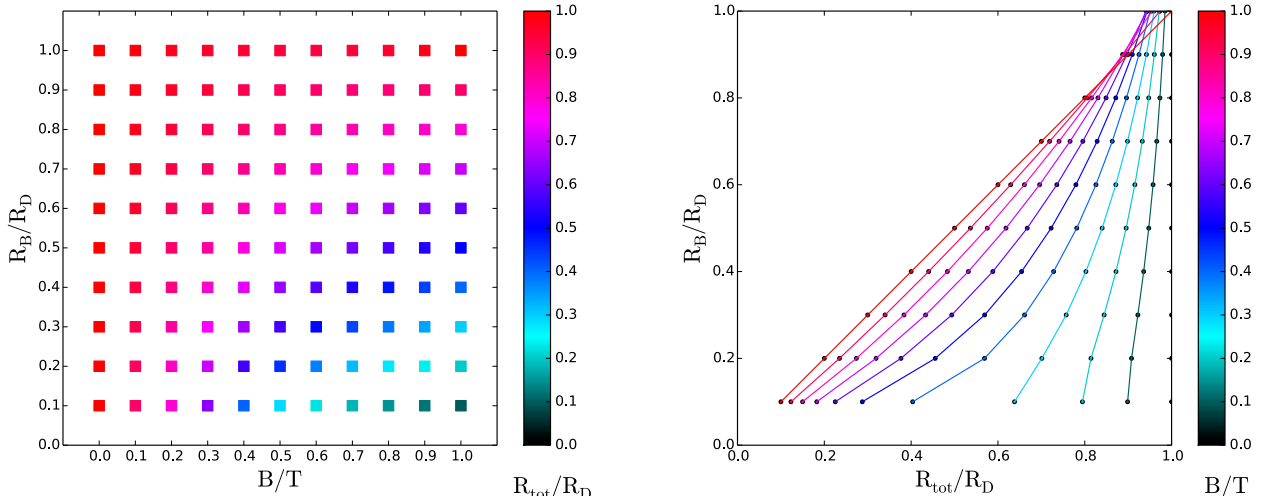

Fig. A.2: *Left:* Size ratio R_B/R_D between the bulge and disk component versus R_{tot}/R_D inferred from noise-free idealized bulge disk models together with B/T. *Right:* R_{tot}/R_D as a function of B/T, colors indicating the size ratio R_B/R_D.

Appendix A.5: Results

Fig. A.3 compares the new values of the bulge-to-total ratio B/T obtained with the method presented in section A.4 with those of the previous artisanal method and shows the B/T distribution in both cases, compared to that of the parent sample in the F160W I-band. Fig. **??** displays the HST map, best-fit model, residuals and the radial profile of the galaxies of the sample. We note that as in Lang et al. (2014), two-component decompositions are preferred over single Sérsic fits for about 2/3 of the sample, namely for 37 galaxies out of 61 (61%).

Appendix A.6: Discussion

Advantages. As in Bruce et al. (2012) and Lang et al. (2014), the main advantage of the method used in this paper is the construction of a grid of initial guesses for the two-component fits to avoid converging to local χ^2 minima. This grid is in part based on the preliminary analysis of noise-free two-component bulge disk models together with single Sérsic fits as in Lang et al. (2014), but also includes initial guesses derived from single bulge and single disk fits, assuming different values of the bulge-to-total luminosity ratio B/T and of the radius ratio R_B/R_D. These additional starting points notably compensate for the cases where the single Sérsic fit yield large Sérsic indices that are not accounted for in the preliminary analysis (L14EG008, L14EG012, L14GN033, XB54, XD53,XG54, XR53 - i.e., G4-21351, G4-6449, GN4-1964, G3-5038, COSMOS 822965, G3-169, COSMOS 816955). Contrarily to Bruce et al. (2012) the initial grid is motivated physically from the single component fits rather than being fixed for all galaxies. Imposing the bulge size to be smaller than the disk as in Lang et al. (2014) enables us to account for faint components, contrarily to Bruce et al. (2012). Indeed, this condition notably removes cases where the faint component is unphysical, such as extremely elongated disks

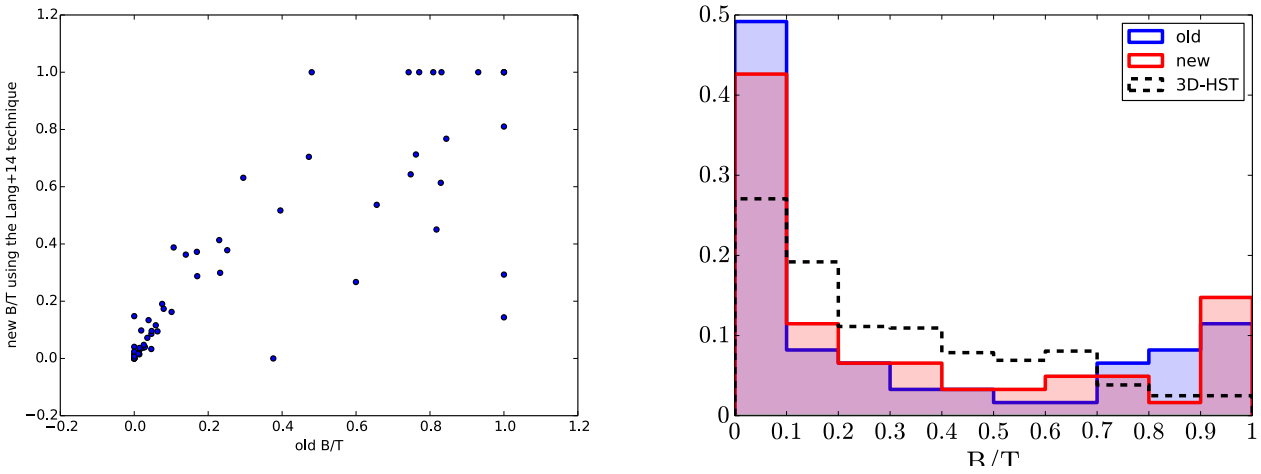

Fig. A.3: *Left:* Comparison between the F814W I-band B/T obtained with the new method derived from Lang et al. (2014) and that derived manually as in the previous version of the paper *Right:* Distribution of the F814W I-band B/T both in the new and old method for the PHIBSS2 sample at $z = 0.5 - 0.8$, compared to the F160W H-band B/T in the CANDELS/3D-HST parent sample in the same redshift, stellar mass and δMS range.

or bulges extending far away in the intergalactic medium. Masking out neighboring sources by carrying out individual Sérsic fits enables us to account for their luminosity distribution, which is particularly crucial when the galaxies are close to each other (cf. for example COSMOS 811360, GN3-4568 or GN4-21683). As in Lang et al. (2014), the centers are allowed to vary up to 2 pixels instead of being locked to the single Sérsic value (Bruce et al. 2012). Finally, imposing the same position angle and a bulge that is similarly or less elongated than the disk enables to remove many unphysical fits, in particular when the bulge overfits the inner structures of the galaxies. We however do note that this condition may be a drawback in certain cases, such as XN53 (COSMOS 839268) which is modeled as a bulge while harboring a very strong bar.

Potential drawbacks. The main drawback of the method used in this article may be the use of a uniform weight map within `galfit`, which are motivated by the almost uniform HST rms images at the redshifts of interest. In their study, Lang et al. (2014) use weight maps obtained from the HST rms images, while Bruce et al. (2012) additionally account for Poisson errors in the high flux regions. By taking a uniform weight and hence neglecting the Poisson errors, the fits potentially higlight regions of higher fluxes. The second drawback of our method is the use of a generic TinyTim PSF instead of a stack of nearby stars (Bruce et al. 2012), a fit to this stack (Contini et al. 2016) or a combination of nearby stars and the TinyTim model (van der Wel et al. 2012; Lang et al. 2014). The ACS is indeed an off-axis camera, which leads to position-dependent PSF not accounted for in the TinyTim PSF used here. Nevertheless, we expect the TinyTim PSF to induce errors well below 10% (Bruce et al. 2012; Contini et al. 2016). Contrarily to Bruce et al. (2012), we do not account for point sources at the centers. Imposing $R_B < R_D$ also lead to unphysical fits for the lenticular galaxy GN025 (GN4-36596), leading us to drop this condition in one of the cases.

Comparison with the parent sample. As can be seen in Fig. A.3, the B/T distribution from the F814W I-band images differ from the average F160W H-band distribution of the parent sample, both with more pure disks and pure bulges. The enhancement of disks most likely has a physical origin, namely that the F814W I-band images are bluer than the parent sample's F160W H-band measurements. However, the enhancement of bulges may be due to the uniform weight matrix that neglects Poisson errors and highlights regions of higher flux. Indeed, amongst the 9 cases where $B/T = 1$, 4 harbor strong spiral arms or a ring. The condition on the axis ratio may prevent fitting the bar, while the rings or strong spiral arms are not accounted for in the bulge disk decomposition we carry out. We decide not to include the $B/T = 1$ cases in the morphological study.

References

Bruce, V. A., Dunlop, J. S., Cirasuolo, M., et al. 2012, MNRAS, 427, 1666
Contini, T., Epinat, B., Bouché, N., et al. 2016, A&A, 591, A49
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Häussler, B., McIntosh, D. H., Barden, M., et al. 2007, ApJS, 172, 615
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Krist, J. 1995, in Astronomical Society of the Pacific Conference Series, Vol. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes, 349
Lang, P., Wuyts, S., Somerville, R. S., et al. 2014, ApJ, 788, 11
Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266
Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193
van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24